Article

# Uniform Free-Energy Profiles of the P−O Bond Formation and Cleavage Reactions Catalyzed by DNA Polymerases $\beta$ and $\lambda$

Martin Klvaňa,[†,‡] Urban Bren,[†,§] and Jan Florián*,[‡]

†Laboratory of Physical Chemistry and Chemical Thermodynamics, Faculty of Chemistry and Chemical Technology, University of Maribor, Smetanova ulica 17, 2000 Maribor, Slovenia

‡Department of Chemistry and Biochemistry, Loyola University Chicago, 1032 W. Sheridan Road, Chicago, Illinois 60660, United States

§Laboratory for Molecular Modeling, National Institute of Chemistry, Hajdrihova ulica 19, 1001 Ljubljana, Slovenia

Ⓢ Supporting Information

**ABSTRACT:** Human X-family DNA polymerases $\beta$ (Pol$\beta$) and $\lambda$ (Pol$\lambda$) catalyze the nucleotidyl-transfer reaction in the base excision repair pathway of the cellular DNA damage response. Using empirical valence bond and free-energy perturbation simulations, we explore the feasibility of various mechanisms for the deprotonation of the 3′-OH group of the primer DNA strand, and the subsequent formation and cleavage of P−O bonds in four Pol$\beta$, two truncated Pol$\lambda$ (tPol$\lambda$), and two tPol$\lambda$ Loop1 mutant (tPol$\lambda\Delta$L1) systems differing in the initial X-ray crystal structure and nascent base pair. The average calculated activation free energies of 14, 18, and 22 kcal mol$^{-1}$ for Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, respectively, reproduce the trend in the observed catalytic rate constants. The most feasible reaction pathway consists of two successive steps: specific base (SB) proton transfer followed by rate-limiting concerted formation and cleavage of the P−O bonds. We identify linear free-energy relationships (LFERs) which show that the differences in the overall activation and reaction free energies among the eight studied systems are determined by the reaction free energy of the SB proton transfer. We discuss the implications of the LFERs and suggest p$K_a$ of the 3′-OH group as a predictor of the catalytic rate of X-family DNA polymerases.
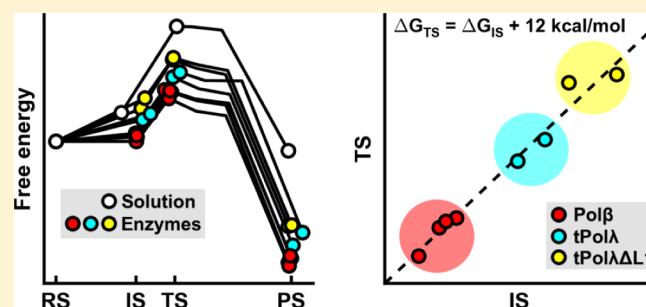
## INTRODUCTION

Preserving genetic and epigenetic information in DNA yet allowing for mutation and recombination, two essential properties of life, to occur, requires a proper balance between DNA mutagenesis and DNA repair processes.[1] Two X-family DNA polymerases,[2] $\beta$ (Pol$\beta$) and $\lambda$ (Pol$\lambda$), have been implicated in numerous DNA repair pathways including base excision repair (BER) of oxidative and alkylation damage.[3−16] Both Pol$\beta$ and Pol$\lambda$ possess 5′-deoxyribosephosphate (dRP) lyase and DNA polymerase activities[8,9,17,18] conferred by two different domains.[19−25] The lyase domain creates a single-nucleotide gap by removing dRP and the polymerase domain inserts a specific 2′-deoxyribonucleotide 5′-triphosphate (dNTP) into the gap according to the identity of the template nucleotide. The fidelity, determined by the ratio between correct and incorrect dNTP insertion events, depends on the proof-reading capacity and reflects the specific biological role of each DNA polymerase. Both Pol$\beta$ and Pol$\lambda$ are low-fidelity DNA polymerases, in part because they lack an intrinsic 3′ → 5′ exonuclease activity.[26−31] To advance our understanding of the structural basis of the catalytic power and the fidelity determinants of Pol$\beta$ and Pol$\lambda$, we need to explore the reaction mechanisms of these enzymes at the atomic level of detail, which can be achieved by complementing experimental (structural and kinetic) observations with the

insight obtained using theoretical calculations.[32] Further progress in this field will eventually lead to a thorough realization of the principles that govern the balance between the protective and harmful involvement of DNA polymerases in carcinogenesis.[33,34]

The lyase and polymerase domains of human Pol$\beta$ and Pol$\lambda$ are evolutionarily homologous, sharing 35.9% amino acid sequence identity (Clustal Omega[35]). Pol$\lambda$ contains an additional, N-terminally located, nuclear localization motif (Met1−Glu35), followed by a BRCT domain (Glu36−Arg132) and a proline-rich domain (Arg133−Pro244),[36] none of which is, however, essential for the lyase and polymerase activities of Pol$\lambda$.[22,29,37] X-ray crystal structures of the ternary complexes of Pol$\beta$ (Thr10−Glu335; PDB code: 2FMP) and truncated Pol$\lambda$ (tPol$\lambda$; comprising the lyase and polymerase domains, Ser245−Trp575; PDB code: 2PFO) with single-nucleotide gapped dsDNA, dNTP, and two active site cations superimpose with a root-mean-square deviation (RMSD)$_{C\alpha}$ of 1.4 Å (Figure 1). The structures differ in amino acid sequence, length, and geometry of the surface loop between $\beta$-strands 3 and 4 (Loop 1). This loop
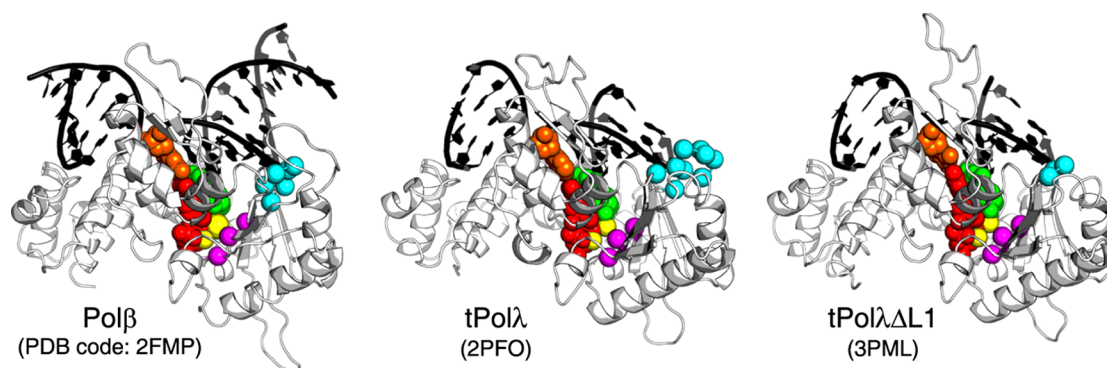
**Figure 1.** Crystal structures of the ternary complexes of human DNA polymerases $\beta$ and $\lambda$. White cartoon, protein; black cartoon, DNA; orange spheres, template nucleotide; green spheres, 3′-terminal primer nucleotide; red spheres, dNTP; yellow spheres, metal ions; magenta spheres, catalytic triad ($C_\alpha$ atoms); cyan spheres, Loop 1 ($C_\alpha$ atoms).

modulates the fidelity of tPol$\lambda$ by controlling the transition from the tPol$\lambda$·DNA binary complex to the tPol$\lambda$·DNA·dNTP·2 Mg$^{2+}$ ternary complex, which is accompanied by a conformational change in the template DNA strand.[38,39] A chimeric tPol$\lambda$ that contains a shorter, Pol$\beta$-like Loop1 (tPol$\lambda\Delta$L1) assumes the tPol$\lambda$-like catalytically competent conformation for the template DNA strand in the ternary complex (Figure 1) but differs from tPol$\lambda$ in the inactive conformation of the binary complex. In contrast, the transition from the inactive to the active conformation in Pol$\beta$ involves an open-to-closed conformational change in the polymerase domain, whereas the template strand remains in the reaction-ready conformation in both binary and ternary complexes.[21]

The rate-limiting nucleotidyl-transfer reaction[19,29,30,40,41] (Figure 2) is initiated in the active ternary complex by deprotonation of the 3′-OH group of the primer DNA strand. The loss of proton is facilitated by catalytic Mg$^{2+}$, which lowers p$K_a$ of the hydroxyl group and stabilizes the evolving negative charge of the nucleophilic oxygen (O$_{nuc}$). The proton acceptor can be either an active site group via a general base (GB) mechanism or a hydroxide anion in bulk water via a specific base (SB) mechanism. In the latter scenario, the proton transfer can be described by a pH-dependent equilibrium between the protonated and deprotonated O$_{nuc}$. The activated O$_{nuc}$ forms a covalent bond with P$_\alpha$ of dNTP and thus extends the primer DNA strand by one nucleotide. Concomitantly, the cleavage of the P$_\alpha$–O$_{lg}$ bond in the triphosphate group of dNTP generates an inorganic pyrophosphate (PP$_i$), the byproduct of the reaction. Despite broad consensus on the overall mechanism of the nucleotidyl-transfer reaction,[19] three important questions remain: (1) What is the preferred mechanism for the deprotonation of the 3′-OH group of the primer DNA strand? (2) What is the degree of concertedness of the P–O bond formation and cleavage reactions? (3) Which of the three reaction steps limits the overall reaction rate?

Several theoretical studies have already addressed these challenging questions by employing hybrid quantum mechanics/molecular mechanics (QM/MM), free-energy perturbation (FEP), or empirical valence bond (EVB) simulation methods to Pol$\beta$[30,42−48] and Pol$\lambda$,[49−51] but definitive answers are yet to be reached: (1) An ordered active site water molecule, two active site catalytic aspartates, a nonbridging oxygen of the P$_\alpha$-group, and bulk water have been proposed as the proton accceptors.[30,42−44,49,51] (2) Associative and concerted P–O bond formation and cleavage reactions have been suggested.[30,43,48,49] (3) Either the proton transfer or the P–O bond formation/
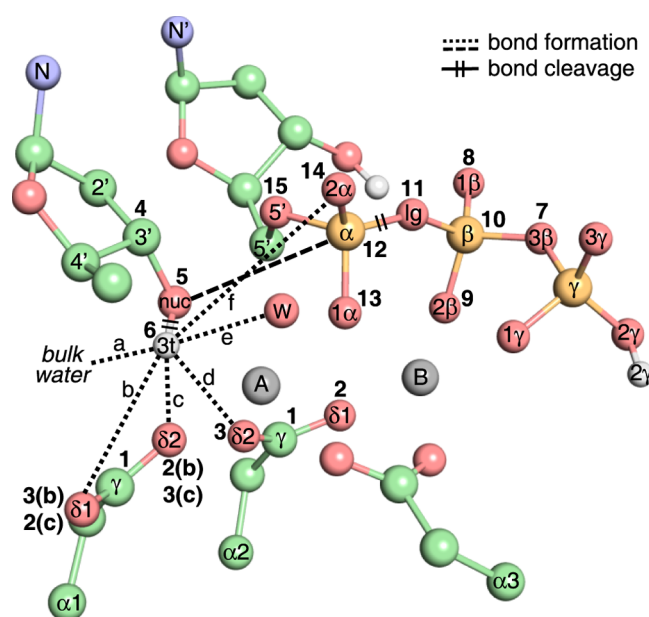


**Figure 2.** Nucleotidyl-transfer reaction catalyzed by Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1. Atoms and ions are colored by element: light green, carbon; light red, oxygen; white, hydrogen; light blue, nitrogen; light orange, phosphorus; light gray, magnesium ion. A, catalytic Mg$^{2+}$; B, structural Mg$^{2+}$; W, oxygen atom of water molecule; N, nitrogen atom of the nucleobase of the primer nucleotide; N′, nitrogen atom of the nucleobase of dNTP; $\alpha$1, C$_\alpha$ of Asp256/490/485 in Pol$\beta$/tPol$\lambda$/tPol$\lambda\Delta$L1; $\alpha$2, C$_\alpha$ of Asp192/429/429 in Pol$\beta$/tPol$\lambda$/tPol$\lambda\Delta$L1; $\alpha$3, C$_\alpha$ of Asp190/427/427 in Pol$\beta$/tPol$\lambda$/tPol$\lambda\Delta$L1; a–f, proton transfer options. The atoms assigned to the reactive region in MD simulations of the reaction in this study are labeled by atom identifiers (1−15); see the Methods section for details.

cleavage reaction has been reported as the rate-limiting step.[30,43−45,49]

This article is a contribution to this ongoing debate. Using FEP[52] and EVB[53,54] methods, we examine the nucleotidyl-transfer reaction in water, Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1 with the aim of (1) obtaining the free-energy profiles of the reaction; (2) calculating the catalytic rate constants ($k_{pol}$) using the transition-state (TS) theory approximation,[55] comparing them with available experimental values (Pol$\beta$ and tPol$\lambda$), and offering them as genuine predictions (tPol$\lambda\Delta$L1); (3) assessing the relative feasibility of the proton transfer to active site aspartates (Asp256 and Asp192 in Pol$\beta$ and their corresponding Asp490/485 and Asp429/429 in tPol$\lambda$/tPol$\lambda\Delta$L1) and to bulk water; (4)

exploring the concertedness and tightness of the P—O bond formation and cleavage reactions by employing a trigonal-bypyramidal geometry around the $P_\alpha$ atom of dNTP in the TS model;[56,57] and (5) determining the dependence of the calculated energetics on the initial atomic coordinates and on the identity of the nascent base pair (Table 1). Our free-energy

**Table 1. Model Systems**

| enzyme | system | XRC[a] | dN$_p$[b] | dNTP·dN$_t$[c] | A[d] | B[d] |
|--------|--------|--------|-----------|----------------|------|------|
| Polβ | Polβ1 | 2FMP | dC | dCTP·dG | Mg$^{2+}$ | Mg$^{2+}$ |
| | Polβ2 | 2PXI | dC | dCTP·dG | Mg$^{2+}$ | Mg$^{2+}$ |
| | Polβ3 | 2FMP | dC | dGTP·dC | Mg$^{2+}$ | Mg$^{2+}$ |
| | Polβ4 | 2PXI | dC | dGTP·dC | Mg$^{2+}$ | Mg$^{2+}$ |
| tPolλ | Polλ1 | 2PFO | dC | dCTP·dG | Mg$^{2+}$ | Mg$^{2+}$ |
| | Polλ2 | 2PFP | dC | dCTP·dG | Mg$^{2+}$ | Mg$^{2+}$ |
| tPolλΔL1 | Polλ3 | 3PML | dC | dGTP·dC | Mg$^{2+}$ | Mg$^{2+}$ |
| | Polλ4 | 3PML | dC | dGTP·dT | Mg$^{2+}$ | Mg$^{2+}$ |

[a]XRC, X-ray crystal structure (PDB code). [b]dN$_p$, 3′-terminal nucleotide of the primer DNA strand. [c]dN$_t$, templating nucleotide of the template DNA strand. [d]A and B, active site metal ions (see Figure 2).

calculations suggest the same reaction mechanism of the nucleotidyl-transfer reaction and similar energetics of the P—O bond formation and cleavage reactions in Polβ and Polλ and indicate linear free-energy relationships (LFERs) as potentially useful concept for further theoretical and experimental studies aimed at deepening our insight into DNA polymerase catalysis and fidelity.

## ■ METHODS

**Structural Models.** The structures of 12 water systems and 24 enzyme systems (Tables S1−S3) were prepared using PyMOL 0.99[58] and Coot 0.6.2[59] molecular graphics programs. The Qprep 5.04 module of the molecular dynamics (MD) package Q[60] was then used to immerse the structures in a sphere of TIP3P[61] water molecules. The sphere, which measured 48 Å in diameter, was centered at the $C_{4'}$ atom of dNTP. Amino acid residues and nucleotides located outside the sphere (Figure S1) were excluded from nonbonded interactions and restrained to their initial coordinates by a harmonic potential with a force constant of 200 kcal mol$^{-1}$ Å$^{-2}$. Each solvated structure (Supporting Files 1 and 2) was equilibrated by employing the ff94 AMBER force field[62] in a series of 12 MD simulations using the Qdyn 5.06 module of Q (Table S4). The equilibrating MD simulation protocol was identical to that of our previous studies:[63,64] Water molecules were subjected to the surface constraint all-atom solvent (SCAAS)[65] boundary condition. The SHAKE algorithm[66] was applied to all hydrogen atoms. Nonbonded interactions were evaluated explicitly and using the local reaction field method[67] within and beyond 10 Å cutoff distance, respectively. The values of stepsize, temperature, and external bath coupling[68] parameters were gradually increased from 0.01 to 2 fs, 5 to 298 K, and 0.1 to 40 fs, respectively. The equilibrated structures were subsequently used in (1) EVB simulations of the GB proton transfer, $O_{nuc}$—$P_\alpha$ bond formation, and $P_\alpha$—$O_{lg}$ bond cleavage reactions, and (2) FEP simulations of the SB proton transfer (Table S5).

**EVB Simulations.** The EVB method (eqs S1−S19) is based on the assumption that the most important effect of the environment (water vs enzyme) is the change in the relative energies of the covalent and ionic valence bond (VB) states.[53] The application of the EVB method thus comprises five steps: (1) Defining VB states. (2) Sampling the potential energy surface between adjacent VB states in water and enzyme systems by FEP simulations. (3) Obtaining free-energy profile (energy gap vs EVB free energy) of the reaction in the water system by superimposing and mixing the potential energy curves of adjacent VB states and reproducing the experimental or ab initio activation and reaction free energies. (4) Applying the water-derived superimposing and mixing parameters to VB potential energy curves of the enzyme system. (5) Calculating the quantitative catalytic effect of the enzyme by comparing its activation and reaction free energies to those of the water system.

Note that the FEP simulations in step 2 are practically performed on the linear combination of molecular mechanical potentials of the two states involved in the given reaction step. EVB free energies in step 3 are generated via the umbrella sampling technique. This technique combines a statistical ensemble, obtained in step 2, with EVB energies obtained by the diagonalization of the EVB Hamiltonian. Although the umbrella sampling step could be, in principle, avoided if the FEP simulations were performed using the EVB forces rather than the classical forces, such simulations would be slower. Perhaps more importantly, they would have to be repeated many times for the correct values of the mixing and superimposing terms in the EVB Hamiltonian to be obtained, so as to reproduce the experimental activation and reaction free energies of the reference reaction.

To model the nucleotidyl-transfer reaction in the water, Polβ, and Polλ systems, a reactive region was designated as consisting of 15 and 11 atoms in the GB and SB pathways, respectively (Figure 2). Four VB states of the reactive region were defined (Tables 2 and S6−S24): (1) the reactant state (RS) represented by the catalytically competent Polβ and Polλ systems, and the

**Table 2. EVB/FEP States for Modeling of the GB and SB Pathways of the Nucleotidyl-Transfer Reaction in the Water, Polβ, and Polλ Systems**[a]

| state | GB | SB |
|-------|-----|-----|
| RS | $O_\delta^{(-0.8514)} + H_{3t}-O_{nuc} + P_\alpha-O_{lg}$ | $H_{3t}-O_{nuc} + P_\alpha-O_{lg}$ |
| IS1 | $O_\delta-H_{3t} + O_{nuc}^{(-0.8549)} + P_\alpha-O_{lg}$ | $O_{nuc}^{(-0.8549)} + P_\alpha-O_{lg}$ |
| IS2a | $O_\delta-H_{3t} + O_{nuc}\cdots(1.90\ Å)\cdots P_\alpha\cdots(1.90\ Å)\cdots O_{lg}$ | $O_{nuc}\cdots(1.90\ Å)\cdots P_\alpha\cdots(1.90\ Å)\cdots O_{lg}$ |
| IS2b | $O_\delta-H_{3t} + O_{nuc}\cdots(2.28\ Å)\cdots P_\alpha\cdots(2.25\ Å)\cdots O_{lg}$ | $O_{nuc}\cdots(2.28\ Å)\cdots P_\alpha\cdots(2.25\ Å)\cdots O_{lg}$ |
| PS | $O_\delta-H_{3t} + O_{nuc}-P_\alpha + O_{lg}^{(-0.8100)}$ | $O_{nuc}-P_\alpha + O_{lg}^{(-0.8100)}$ |

[a]Each state is described schematically by the atoms or atomic groups most directly involved in the formation and cleavage ($\cdots$) of the covalent bonds (−). Atomic charges of selected atoms are indicated in superscripts, and distances between the oxygen and phosphorus atoms in IS2 and IS2b are given in parentheses. See Tables S6−S28 for the full description of the EVB/FEP states. The missing proton in the IS1, IS2, and PS states in the SB pathway was compensated by sodium and/or chloride ions in the water systems, and by a charged arginine or a sodium ion in the Polβ and Polλ systems (see Tables S1 and S3).

corresponding reference water systems; (2) the proton transfer intermediate state (IS1); (3) an artificial intermediate state (IS) along the concerted P−O bond formation and cleavage reactions (IS2); and (4) the product state (PS) with the primer DNA strand extended by one nucleotide and PP$_i$ fully formed. The artificial IS2 allowed us to explore the TS region of the free-energy surface in a thorough and controlled way.

To model the compact and loose associative P−O bond formation and cleavage reactions, two IS2 VB states, IS2a and IS2b, differing in O$_{nuc}$⋯P$_\alpha$⋯O$_{lg}$ distances and, consequently, in the charge distribution were defined. The atomic charges of RS, IS1, IS2a, and PS were obtained from Florián et al.[56] These charges are consistent with the atom-centered point charge model of the ff94 AMBER force field, which is based on molecular electrostatic potentials obtained using the Hartree−Fock method and the 6-31G* basis set. The atomic charges of IS2b were based on IS2a and adjusted to account for the difference between PCM/B3LYP/TZVP//HF/6-31G* Mulliken atomic charges in QM models of the compact and loose associative-like TS between IS1 and IS2.[69] To achieve net zero charge of the reactive region in IS1, IS2, and PS in the SB pathway, the atomic charge of C$_{3'}$ in 2′-deoxyribose (dRib; water systems) and dC (Pol$\beta$ and Pol$\lambda$ systems) was further adjusted (cf. Tables S6 and S7).

The FEP method was employed to sample the transition between adjacent VB states in both directions using Qdyn: RS ↔ IS1 ↔ (IS2a or IS2b) ↔ PS in the GB pathway and IS1 ↔ (IS2a or IS2b) ↔ PS in the SB pathway. The transition was carried out in 51 $k$ steps by gradually changing the value of the FEP mapping parameter $\lambda$ from 1 to 0; $\Delta\lambda = \lambda_{k+1} - \lambda_k = -0.02$. Each FEP step consisted of 10 ps MD simulation at 298 K with a stepsize of 1 fs; MD parameters were identical, except for the stepsize, to the equilibrating MD simulation XII (see Table S4). Mapping and perturbation potential energies of the reactive region with the surrounding environment were recorded every 10 fs with no cutoff applied to nonbonded interactions. The recorded potential energies, 51 × 999 = 50 949 for each FEP transition, were processed using the Qfep 5.01 module of Q: EVB free energies ($y$ coordinate) were calculated as a function of the energy gap ($x$ coordinate, consisting of 100 and 60 bins in the RS ↔ IS1 and IS1 ↔ IS2 ↔ PS reactions, respectively) using the umbrella sampling method. The coupling of adjacent VB$_i$ and VB$_j$ potential surfaces and difference in their minima were modulated using EVB parameters $H_{ij}$ and $\Delta\alpha_j$ to set the activation and reaction free energies in the water systems in accordance with Florián et al.[57] See eqs S1−S19 for a detailed description of the EVB method.

**FEP Simulations.** The SB proton transfer free energy, which corresponds to the free energy of IS1, was calculated from p$K_a$ = 12.67 for the 3′-OH group of dRib[70] and from FEP simulations of the hydrogen annihilation and creation in the water, Pol$\beta$, and Pol$\lambda$ systems. The reactive region consisted of the C$_{3'}$, O$_{nuc}$, and H$_{3t}$ atoms of dRib in the water systems or dC in the Pol$\beta$ and Pol$\lambda$ systems (Tables S25−S28). Two FEP states were defined: (1) RS with the 3′-OH group and (2) intermediate state 1 (IS1) with O$_{nuc}^-$ (Table 2). The transition between RS and IS1 was carried out in 51 $k$ FEP steps, each consisting of 10 ps MD simulation, using Qdyn. The sampling was performed in both directions (RS ↔ IS1); the hydrogen creation was initiated from the final coordinates of the hydrogen annihilation MD simulation. All MD parameters were identical to the sampling MD simulations in the EVB approach. The reaction free energy was calculated as the cumulative sum of average differences

between mapping and perturbation potential energies using Zwanzig's FEP formula.[52] See eqs S1−S5 and S20 for a detailed description of the FEP method.

**Calibration of the Free-Energy Profile of the RS → IS1 Reaction in Solution.** The EVB reaction free energy of the GB proton transfer in the reference water system was set to 10.8 kcal mol$^{-1}$, according to the difference in p$K_a$ between the hydrogen donor (dRib) and acceptor (acetate)

$$\Delta g_{0,\text{EVB,water}} = 2.303RT(pK_{a,\text{dRib}} - pK_{a,\text{acetate}}) \quad (1)$$

where $R = 1.9872041 \times 10^{-3}$ kcal mol$^{-1}$ K$^{-1}$ is the gas constant, $T$ = 298 K is the simulation temperature, p$K_{a,\text{dRib}}$ = 12.67,[70] and p$K_{a,\text{acetate}}$ = 4.74. The EVB activation free energy of the GB proton transfer was set to 12.6 kcal mol$^{-1}$, that is, 1.8 kcal mol$^{-1}$ above the reaction free energy,[57] which is in accordance with the assumption of a low activation barrier for proton transfer between two oxygen atoms.[71]

The FEP reaction free energy of the SB proton transfer in the reference water system was set to 7.7 kcal mol$^{-1}$, according to the difference between p$K_a$ of dRib and pH of 7

$$\Delta g_{0,\text{FEP,water}} = 2.303RT(pK_{a,\text{dRib}} - pH) \quad (2)$$

The differences in p$K_a$ between the enzyme and water systems were calculated from FEP free energies (see eq S20)

$$\Delta pK_{a,\text{enzyme}} = \frac{1}{2.303RT}(\Delta g_{\text{FEP,enzyme}} - \Delta g_{\text{FEP,water}}) \quad (3)$$

and used for the calculation of the FEP reaction free energy of the RS → IS1 step in the enzyme systems

$$\Delta g_{0,\text{FEP,enzyme}} = 2.303RT(pK_{a,\text{dRib}} + \Delta pK_{a,\text{enzyme}} - pH) \quad (4)$$

The value of $\Delta g_{\text{FEP,water}}$ in eq 3 is a weighted average, based on Mg$^{2+}$ occupancy, of the FEP free energies obtained from FEP simulations of the water system with/without Mg$^{2+}$ ions. The Mg$^{2+}$ occupancy was set to 0.7, which yielded $\Delta pK_a$ for Pol$\beta$3 and Pol$\beta$4 in accordance with the difference between the approximate inflection point in the $k_{\text{pol}}$ versus pH plot of 8.3 for dGTP·dC nascent base pair in Pol$\beta$[41] and the p$K_a$ of dRib.

**Calibration of the Free-Energy Profile of the IS1 → IS2 Reaction in Solution.** The EVB activation free energy of the IS1 → IS2 step in the reference water system was set to 23.1 kcal mol$^{-1}$,[57] which was calculated as the sum of three contributions

$$\Delta g_{\text{EVB,water}}^\ddagger = \Delta g_{\text{cpmp}}^\ddagger + \Delta g_{\text{cage}}^\ddagger + \Delta g_{\text{mg}}^\ddagger \quad (5)$$

where $\Delta g_{\text{cpmp}}^\ddagger$ = 27.5 kcal mol$^{-1}$ is the activation free energy for the hydrolysis of (3-chlorophenyl) methyl phosphate, whose p$K_a$ = 9.02[72] is similar to p$K_{a4}$ = 8.93 of PP$_i$,[73] by hydroxide anion in solution;[74,75] $\Delta g_{\text{cage}}^\ddagger$ = −2.4 kcal mol$^{-1}$ is the solvent cage effect;[57] and $\Delta g_{\text{mg}}^\ddagger$ = −2.0 kcal mol$^{-1}$ is the correction for the presence of Mg$^{2+}$ ions in the reference water system.[76] The EVB reaction free energy of 19.5 kcal mol$^{-1}$ for the IS1 → IS2 step was obtained as follows

$$\Delta g_{0,\text{EVB,water}} = \Delta g_{\text{EVB,water}}^\ddagger + \Delta\Delta g_0^\ddagger + c \quad (6)$$

where $\Delta g_{\text{EVB,water}}^\ddagger$ = 23.1 kcal mol$^{-1}$ is the activation free energy of the IS1 → IS2 step (eq 5), $\Delta\Delta g_0^\ddagger$ = −2.6 kcal mol$^{-1}$ is the difference between the reaction and activation free energies used by Florián et al.,[57] and $c$ = −1.0 kcal mol$^{-1}$ is the correction introduced in this study to make the intermediate observable on the EVB free-energy surface and to change the rate-limiting step

in aqueous solution from the IS2 → PS to the IS1 → IS2 step, in accordance with ref 69.

**Calibration of the Free-Energy Profile of the IS2 → PS Reaction in Solution.** The EVB activation free energy of the IS2 → PS step in the reference water system was set to 2.6 kcal mol$^{-1}$, the value used by Florián et al.[57] The EVB reaction free energy of the IS2 → PS step in the reference water system was set to −32.8 kcal mol$^{-1}$ in the GB pathway and −29.7 kcal mol$^{-1}$ in the SB pathway (−32.8 + 10.8 − 7.7 kcal mol$^{-1}$; cf. eqs 1 and 2), which yielded the intended cumulative RS → PS reaction free energy of −2.5 kcal mol$^{-1}$ in both pathways, a value in accordance with the equilibrium constant of ~10$^2$ observed for the extension of a single-stranded DNA by one nucleotide, dA$_n$ + dATP ↔ dA$_{n+1}$ + PP$_i$, catalyzed by terminal deoxynucleotidyl-transferase.[77]

**Calculation of the Catalytic Effect.** The free-energy profiles of the entire nucleotidyl-transfer reaction were created by combining the free-energy profiles of individual steps: RS ↔ IS1, IS1 ↔ IS2, and IS2 ↔ PS, where IS2 was either IS2a or IS2b. The value of IS1 was obtained from EVB (in the GB pathway) or FEP (in the SB pathway). The IS1 → PS part in the SB pathway was calculated as the average of two versions, SB1 and SB2, which differed in the composition of the simulated system (see Tables S1, S3, and S5). The catalytic effect was then calculated as the difference in activation free energies between reactions in the enzyme and water systems. To facilitate comparison of the calculated and observed reaction energetics, the calculated activation free energies (kcal mol$^{-1}$) were converted into the catalytic rate constants (s$^{-1}$; eq 7) and the observed rate constants were converted into activation free energies (eq 8)[55]

$$k_{pol} = \frac{k_B T}{h} e^{-\Delta g^{\ddagger}/RT} \quad (7)$$

$$\Delta g^{\ddagger} = -RT \ln\left(k_{pol} \frac{h}{k_B T}\right) \quad (8)$$

where $T$ (K) is the simulation (eq 7) or experimental temperature (eq 8), $k_B = 1.3806488 \times 10^{-23}$ m$^2$ kg s$^{-2}$ K$^{-1}$ is the Boltzmann constant, and $h = 6.62606957 \times 10^{-34}$ m$^2$ kg s$^{-1}$ is the Planck constant.

**Structural Analysis of MD Trajectories.** VMD[78] and PyMOL molecular graphics programs were used for the visualization of MD trajectories and individual structures, respectively. The geometrical parameters, including RMSD, interatomic distances, and angles, were extracted from MD trajectories using VMD. PyMOL was used for creating all molecular structure images for this article. O$_{nuc}$−P$_{\alpha}$ and P$_{\alpha}$−O$_{lg}$ distances in TSs were calculated as weighted averages of all FEP $k$ steps that contributed to the TS bin. (Note that configurations from different $k$ steps may contribute to one EVB bin, and, vice versa, configurations from the same $k$ step may belong to different EVB bins; see eq S16.) Animations of the nucleotidyl-transfer reaction were generated from the final snapshots of the FEP steps that corresponded to each state (FEP simulations) or contributed by the highest number of energy points to each state (EVB simulations).

## ■ RESULTS

**Free-Energy Profiles of the Proton Transfer Reaction.** Two mechanisms for the deprotonation of the 3′-OH group of the primer DNA strand were considered: (1) GB proton transfer to the proximal and distal carboxylic oxygens of Asp256/490/

485 (labeled O$_{\delta2}$ and O$_{\delta1}$, respectively, in Figure 2) and to the proximal carboxylic oxygen (O$_{\delta2}$) of Asp192/429/429 in the Polβ/tPolλ/tPolλΔL1 systems and (2) SB proton transfer to bulk water.

*GB Proton Transfer (RS → IS1).* The proton transfer to the active site aspartates in enzyme systems and to the corresponding acetate ions in the reference water system (Figure S2; Table S29) was modeled using the EVB approach: The reactants and products were defined using RS and IS1 VB states, and the RS → IS1 transformation was carried out using the FEP method. The initial set of six enzyme systems comprised Polβ1, Polβ3, Polλ1, Polλ2, Polλ3, and Polλ4. O$_{\delta2}$ of Asp192/429/429 was the least feasible proton acceptor with minimum and maximum activation free energy (Δ$g^{\ddagger}$) values of 24.5 and 32.1 kcal mol$^{-1}$, respectively, among the six enzyme systems. Of the two carboxylic oxygens of Asp256/490/485, O$_{\delta1}$, which accepted the proton with the minimum, maximum, and average Δ$g^{\ddagger}$ of 16.7, 23.4, and 19.1 kcal mol$^{-1}$, respectively, was surpassed in feasibility by O$_{\delta2}$, which accepted the proton with the minimum, maximum, and average Δ$g^{\ddagger}$ values of 10.6, 13.7, and 12.0 kcal mol$^{-1}$, respectively. The proton transfer to O$_{\delta2}$ of Asp256/490/485 also showed the lowest reaction free energies (Δ$g_0$) with an average of 6.5 kcal mol$^{-1}$ among the six enzyme systems, which was 10.5 and 7.5 kcal mol$^{-1}$ lower than the average Δ$g_0$ for the two less favorable proton acceptors, O$_{\delta2}$ of Asp192/429/429 and O$_{\delta1}$ of Asp256/490/485, respectively.

To increase the robustness of the calculated Δ$g^{\ddagger}$ and Δ$g_0$ for the preferred proton transfer to O$_{\delta2}$ of Asp256/490/485, EVB simulations were also performed in the reverse direction (RS ← IS1), and Polβ2 and Polβ4 were added to the set of enzyme systems. The reverse reaction failed to proceed in Polβ4, because the Morse potential for the distance between O$_{nuc}$ and the proton did not provide the sufficient driving force needed for overcoming the energy barrier caused by a strong electrostatic attraction between the proton and O$_{\delta2}$ of Asp190. In other enzyme systems, the values of both Δ$g^{\ddagger}$ and Δ$g_0$ were on average 1.2 and 1.8 kcal mol$^{-1}$ lower than the corresponding free energies in the forward direction. The final values of Δ$g^{\ddagger}$ and Δ$g_0$ were obtained by averaging the free energies from the forward and reverse proton transfer (RS ↔ IS1). The GB proton transfer in the Polβ and Polλ systems showed average catalytic effects of 2.7 and 0.2 kcal mol$^{-1}$, respectively, when compared to that of the reference water system calibrated to Δ$g^{\ddagger} = 12.6$ kcal mol$^{-1}$. The proton transfer product state (IS1) in the Polβ and Polλ systems was on average 6.0 and 4.4 kcal mol$^{-1}$ more stable than the corresponding IS1 in the water system calibrated to Δ$g_0 = 10.8$ kcal mol$^{-1}$. Thus, the GB proton transfer reaction was endergonic with average free-energy cost of 4.8 and 6.4 kcal mol$^{-1}$ in the Polβ and Polλ systems, respectively.

*Specific Base (SB) Proton Transfer (RS → IS1).* The proton transfer to bulk water was modeled using the FEP approach: The reactants and products were defined by the RS and IS1 FEP states; the RS ↔ IS1 transformation was carried out in the forward (proton annihilation) and reverse (proton creation) directions. The obtained FEP free energies were averaged and converted into the SB proton transfer reaction free energies on the basis of the differences in p$K_a$ of the 3′-OH group between the enzyme and reference water systems (Table S30). The SB proton transfer reaction was endergonic with average reaction free energies of 1.3, 6.6, and 10.2 kcal mol$^{-1}$ in Polβ, tPolλ, and tPolλΔL1 systems, respectively. Compared to the reference water system, Δ$g_0$ values were 6.4 kcal mol$^{-1}$ lower in Polβ, 1.1 kcal mol$^{-1}$ lower in tPolλ, and 2.5 kcal mol$^{-1}$ higher in tPolλΔL1.

Compared to the GB proton transfer, $\Delta g_0$ values were 3.5 kcal mol$^{-1}$ lower in Pol$\beta$, 0.2 kcal mol$^{-1}$ higher in tPol$\lambda$, and 3.9 kcal mol$^{-1}$ higher in tPol$\lambda\Delta$L1.

**Free-Energy Profiles of the Concerted Nucleophilic Attack and PP$_i$ Formation Reactions.** The concerted $O_{nuc}$–$P_\alpha$ formation and $P_\alpha$–$O_{lg}$ cleavage reactions were split into two successive steps by introducing an artificial VB state (IS2) between the proton transfer product (IS1) and the final product of the nucleotidyl-transfer reaction (PS): IS1 → IS2 → PS. Two variants of IS2 were employed for exploring the tightness of the TS of the overall reaction: IS2a for the compact TS and IS2b for the loose TS. IS2a and IS2b differed in the atomic charges in the reactive region ($-0.65e$, $1.16e$, and $-0.55e$ for $O_{nuc}$, $P_\alpha$, and $O_{lg}$ in IS2a vs $-0.75e$, $1.22e$, and $-0.65e$ for the corresponding atoms in IS2b; see Tables S6 and S7) and in the equilibrium P–O distances in the Morse potential energy function (1.90 Å for $O_{nuc}$–$P_\alpha$ and $P_\alpha$–$O_{lg}$ distances in IS2a vs 2.28 Å for $O_{nuc}$–$P_\alpha$ and 2.25 Å for $P_\alpha$–$O_{lg}$ distances in IS2b; see Tables S13–S15).

*Part I of the P–O Bond Formation and Cleavage Reactions (IS1 → IS2).* In the GB pathway, the IS1 → IS2a reaction proceeded with an average $\Delta g^{\ddagger}$ of 17.6 kcal mol$^{-1}$ among Pol$\beta$1, Pol$\beta$2, and Pol$\beta$4, and 15.7 kcal mol$^{-1}$ among Pol$\lambda$. Pol$\beta$3 was an outlier with an average $\Delta g^{\ddagger}$ of 21.9 kcal mol$^{-1}$ (squares and circles in Figure S3; Table S31). The IS1 → IS2b reaction proceeded with average $\Delta g^{\ddagger}$ values of 18.5 and 16.2 kcal mol$^{-1}$ in Pol$\beta$ and Pol$\lambda$ systems, respectively (triangles and pentagons in Figure S3; Table S32). IS2b was less stable (by 1.5 and 2.2 kcal mol$^{-1}$ in Pol$\beta$ and Pol$\lambda$ systems, respectively) than IS2a. However, considering both $\Delta g^{\ddagger}$ and $\Delta g_0$, the energetics of the IS1 → IS2a and IS1 → IS2b reactions were too similar to allow for a confident designation of a preferred IS2 geometry.

The IS1 → IS2 → PS reaction steps following the SB proton transfer were modeled using two different approaches to compensate for the total charge of the systems for the lost proton: (1) setting the net charge of the remote Arg152/389/389, located >14 Å from the $P_\alpha$ atom of dNTP in Pol$\beta$/tPol$\lambda$/tPol$\lambda\Delta$L1, to $+1e$ (SB1; Figure S4A; Tables S33 and S34) or (2) adding one unrestrained Na$^+$ ion into the active site, about 4 Å from the $P_\alpha$ atom of dNTP (SB2; Figure S4B; Tables S35 and S36). The energetics of the IS1 → IS2 → PS reaction steps obtained with the SB1 and SB2 structures could not be confidently distinguished (Tables S33–S36 and S41–S44); therefore, the $\Delta g^{\ddagger}$ and $\Delta g_0$ values were averaged (Tables S37, S38, S45, and S46).

The IS1 → IS2a reaction in the SB pathway (squares and circles in Figure S4; Table S37) proceeded with an average $\Delta g^{\ddagger}$ of 12.8 kcal mol$^{-1}$ in Pol$\beta$ systems and 11.8 kcal mol$^{-1}$ in Pol$\lambda$ systems. Among the eight enzyme systems, the maximum difference in $\Delta g^{\ddagger}$ was only 2.2 kcal mol$^{-1}$, and the standard deviation of the average $\Delta g^{\ddagger}$ of 12.3 kcal mol$^{-1}$ was only 0.3 kcal mol$^{-1}$. The average catalytic effect among the enzyme systems of 10.8 kcal mol$^{-1}$ was 4.3 kcal mol$^{-1}$ larger than that for the corresponding reaction in the GB pathway. Similarly to $\Delta g^{\ddagger}$, the values of $\Delta g_0$ showed very little variation among the enzyme systems, averaging 8.4 kcal mol$^{-1}$ in Pol$\beta$ systems and 7.5 kcal mol$^{-1}$ in Pol$\lambda$ systems. Thus, the free-energy difference between $\Delta g_0$ and $\Delta g^{\ddagger}$ was, on average, nearly identical in the water ($-4.6$ kcal mol$^{-1}$), Pol$\beta$ ($-4.4$ kcal mol$^{-1}$), and Pol$\lambda$ ($-4.3$ kcal mol$^{-1}$) systems.

The IS1 → IS2b reaction in the SB pathway (triangles and pentagons in Figure S4; Table S38) proceeded with an average $\Delta g^{\ddagger}$ of 11.8 kcal mol$^{-1}$ in Pol$\beta$ systems and 11.6 kcal mol$^{-1}$ in Pol$\lambda$ systems. Among the eight enzyme systems, the maximum

difference in $\Delta g^{\ddagger}$ was only 2.3 kcal mol$^{-1}$, and the standard deviation of the average $\Delta g^{\ddagger}$ of 11.7 kcal mol$^{-1}$ was only 0.3 kcal mol$^{-1}$; these parameters are nearly identical to those of the reaction involving IS2a. The average catalytic effect among the enzyme systems of 11.4 kcal mol$^{-1}$ was 0.6 kcal mol$^{-1}$ larger compared to that of the IS1 → IS2a reaction and 4.9 kcal mol$^{-1}$ larger compared to that of the IS1 → IS2b reaction in the GB pathway. Similarly to $\Delta g^{\ddagger}$, the values of $\Delta g_0$ also showed very little variation among the enzyme systems, averaging 5.7 kcal mol$^{-1}$ in Pol$\beta$ systems and 5.6 kcal mol$^{-1}$ in Pol$\lambda$ systems. However, the free-energy difference between $\Delta g_0$ and $\Delta g^{\ddagger}$ of $-6.0$ kcal mol$^{-1}$ in the enzyme systems was 1.4 kcal mol$^{-1}$ larger compared to that of the reference water system and 1.6 kcal mol$^{-1}$ larger compared to that of the reaction involving IS2a. Thus, the IS1 → IS2b reaction was more favorable than the corresponding reaction involving IS2a in both $\Delta g^{\ddagger}$ and $\Delta g_0$.

*Part II of the P–O Bond Formation and Cleavage Reactions (IS2 → PS).* In the GB pathway, the IS2a → PS and IS2b → PS reactions proceeded with average $\Delta g^{\ddagger}$ values of 3.3 and 3.0 kcal mol$^{-1}$ among Pol$\beta$ and Pol$\lambda$ systems, that is, 0.7 and 0.4 kcal mol$^{-1}$ above $\Delta g^{\ddagger}$ of the reference water system (Figure S5; Tables S39 and S40). The reactions were exergonic with average $\Delta g_0$ values in Pol$\beta$ and Pol$\lambda$ systems of $-34.0$ and $-29.7$ kcal mol$^{-1}$ in the reactions involving IS2a and IS2b, respectively, that is, 1.2 kcal mol$^{-1}$ below and 3.1 kcal mol$^{-1}$ above $\Delta g_0$ of the reference water system ($-32.8$ kcal mol$^{-1}$). In the SB pathway, the IS2a → PS and IS2b → PS reactions proceeded with average $\Delta g^{\ddagger}$ values of 3.3 and 2.3 kcal mol$^{-1}$ among Pol$\beta$ and Pol$\lambda$ systems (Figure S6; Tables S45 and S46); these $\Delta g^{\ddagger}$ values were very similar to those of the GB pathway. The IS2a → PS and IS2b → PS reactions in the SB pathway were exergonic with an average $\Delta g_0$ in Pol$\beta$ and Pol$\lambda$ systems of $-38.7$ kcal mol$^{-1}$ in the reactions involving IS2a or IS2b, that is, 9.0 kcal mol$^{-1}$ below $\Delta g_0$ of the reference water system ($-29.7$ kcal mol$^{-1}$).

**Cumulative Free-Energy Profiles of the Nucleotidyl-Transfer Reaction.** The average activation and reaction free energies of the three successive steps, RS → IS1, IS1 → IS2, and IS2 → PS, were adjoined, and the cumulative $\Delta g^{\ddagger}$ and $\Delta g_0$, relative to RS, were obtained for each step. The artificial IS2 in the water systems was removed by shifting the free energy of IS2 in both GB and SB pathways (IS2a by $+3.17$ and $+3.25$ kcal mol$^{-1}$; IS2b by $+3.16$ and $+3.27$ kcal mol$^{-1}$); these free-energy shifts were then applied to the corresponding IS2 in the enzyme systems (Tables 3 and 4). The reaction coordinate of each state (Tables S47–S50) was obtained as the cumulative sum of energy gaps of the individual reaction steps (RS to PS in the GB pathway and IS1 to PS in the SB pathway) or by adoption from the GB pathway (RS → IS1 in the SB pathway).

*Activation Free Energies of the Nucleotidyl-Transfer Reaction.* The highest point in the free-energy profiles of the enzyme-catalyzed reactions ($\Delta G^{\ddagger}$) was TS2 (as in the water system), IS2a, or TS3 (Figure S7). TS2 was the overall TS in 24 cases, including all Pol$\beta$ and Pol$\lambda$ systems in the GB and SB pathways involving IS2b. IS2a in the GB pathway was the overall TS in Pol$\beta$3, Pol$\lambda$1, Pol$\lambda$2, and Pol$\lambda$3. TS3 in the GB pathway involving IS2a was the overall TS in Pol$\beta$1, Pol$\beta$2, and Pol$\lambda$4; TS3 in the SB pathway involving IS2a was the overall TS in Pol$\lambda$1. The maximum difference in the free energy of TS2, IS2, and TS3 ranged from 0.5 to 4.4 kcal mol$^{-1}$ in the GB pathway, and from 0.5 to 4.6 kcal mol$^{-1}$ in the SB pathway.

No significant differences were observed in the values of $\Delta G^{\ddagger}$ in the reactions involving IS2a and IS2b. In the GB pathway, the average $\Delta G^{\ddagger}$ was 23.9, 23.2, and 21.7 kcal mol$^{-1}$ in Pol$\beta$, tPol$\lambda$,

**Table 3. Cumulative Activation and Reaction Free Energies in the GB Pathway of the Nucleotidyl-Transfer Reaction in the Water, Polβ, and Polλ Systems**[a]

| system | TS1 | IS1 | TS2 | IS2 | TS3 | PS |
|---|---|---|---|---|---|---|
| | | | kcal mol$^{-1}$ | | | |
| water | 12.6 | 10.8 | 33.9 | 33.8 | 32.9 | −2.5 |
| Polβ1 | 10.2 | 5.2 | 23.1 | 23.8 | 24.1 | −13.7 |
| | | | 22.3 | 19.7 | 19.5 | −13.7 |
| Polβ2 | 9.7 | 4.4 | 21.7 | 22.2 | 22.5 | −14.0 |
| | | | 20.9 | 18.1 | 18.2 | −13.3 |
| Polβ3 | 8.5 | 3.2 | 25.2 | 28.0 | 27.4 | −11.8 |
| | | | 23.3 | 23.1 | 22.4 | −8.9 |
| Polβ4 | 11.2 | 6.2 | 23.8 | 23.3 | 23.2 | −12.2 |
| | | | 26.5 | 23.2 | 23.4 | −9.7 |
| Polλ1 | 12.4 | 6.8 | 23.2 | 23.7 | 23.5 | −14.5 |
| | | | 23.7 | 21.1 | 20.6 | −12.0 |
| Polλ2 | 13.4 | 6.0 | 22.3 | 23.0 | 22.5 | −15.1 |
| | | | 22.5 | 21.1 | 20.8 | −12.9 |
| Polλ3 | 11.5 | 5.9 | 20.9 | **21.0** | 19.8 | −17.3 |
| | | | 22.0 | 19.3 | 18.1 | −14.5 |
| Polλ4 | 12.4 | 6.8 | 21.6 | 21.4 | 22.0 | −15.1 |
| | | | 21.9 | 18.7 | 17.5 | −16.6 |

[a]TS1, transition state 1; IS1, intermediate state 1; TS2, transition state 2; IS2, intermediate state 2a or 2b (see Table 2); TS3, transition state 3; PS, product state. Upper rows for the enzyme systems: IS1 → IS2a → PS; lower rows for the enzyme systems: IS1 → IS2b → PS. Numbers in bold denote the lowest rate-determining activation free energies among the GB (this table) and SB pathways (Table 4).

**Table 4. Cumulative Activation and Reaction Free Energies in the SB Pathway of the Nucleotidyl-Transfer Reaction in the Water, Polβ, and Polλ Systems**[a]

| system | IS1 | TS2 | IS2 | TS3 | PS |
|---|---|---|---|---|---|
| | | kcal mol$^{-1}$ | | | |
| water | 7.7 | **30.8** | 30.7 | 29.8 | −2.5 |
| Polβ1 | 1.6 | 14.6 | 13.7 | 13.5 | −28.3 |
| | | **13.6** | 11.4 | 10.2 | −30.6 |
| Polβ2 | 0.0 | 11.8 | 11.1 | 11.5 | −30.3 |
| | | **11.5** | 8.1 | 6.9 | −33.3 |
| Polβ3 | 2.2 | 15.2 | 13.9 | 13.8 | −29.4 |
| | | **13.8** | 11.4 | 9.9 | −32.4 |
| Polβ4 | 1.2 | 14.7 | 13.9 | 13.0 | −29.0 |
| | | **13.2** | 10.9 | 9.6 | −31.3 |
| Polλ1 | 7.4 | 19.3 | 19.4 | 19.8 | −23.9 |
| | | **18.5** | 16.3 | 16.2 | −24.4 |
| Polλ2 | 5.8 | 17.9 | 16.5 | 16.4 | −24.0 |
| | | **17.2** | 14.1 | 12.6 | −27.8 |
| Polλ3 | 11.6 | 23.1 | 22.8 | 21.9 | −20.0 |
| | | 22.4 | 20.5 | 19.0 | −22.4 |
| Polλ4 | 8.8 | **20.9** | 19.1 | 18.7 | −22.3 |
| | | 21.9 | 19.6 | 17.9 | −22.7 |

[a]IS1, intermediate state 1; TS2, transition state 2; IS2, intermediate state 2a or 2b (see Table 2); TS3, transition state 3; PS, product state. Upper rows for the enzyme systems: IS1 → IS2a → PS; lower rows for the enzyme systems: IS1 → IS2b → PS. Numbers in bold denote the lowest rate-determining activation free energies among the SB (this table) and GB pathways (Table 3).

and tPolλΔL1 systems. In the SB pathway, the average $\Delta G^{\ddagger}$ was 13.6, 18.4, and 22.1 kcal mol$^{-1}$ in Polβ, tPolλ, and tPolλΔL1 systems. The average catalytic effect in the GB pathway was 10.0, 10.7, and 12.2 kcal mol$^{-1}$ in Polβ, tPolλ, and tPolλΔL1 systems. The average catalytic effect in the SB pathway was 17.2, 12.4, and 8.7 kcal mol$^{-1}$ in Polβ, tPolλ, and tPolλΔL1 systems.

*LFERs.* LFERs with correlation coefficients >0.80 were detected for the free energies of TS3 versus PS in the GB pathway involving IS2b ($R^2$ of 0.85) and for the free energies of IS1 versus PS, TS2 versus PS, and TS3 versus PS in the SB pathway involving IS2a or IS2b with $R^2$ ranging from 0.94 to 0.98 (Figure 3). In the SB pathway, the slope of the linear regression lines in IS1 versus PS, TS2 versus PS, and TS3 versus PS relationships for the reactions involving IS2a and IS2b ranged from 0.92 to 1.08. Averaging over the compact (via IS2a) and loose (via IS2b) variants of the SB pathway yielded the following LFER equations for relating the free-energy changes in IS1, TS2, TS3, and PS

$$\Delta\Delta g_{IS1} = 1.00\Delta\Delta g_{PS} \tag{9}$$

$$\Delta\Delta g_{TS2} = 0.94\Delta\Delta g_{PS} \tag{10}$$
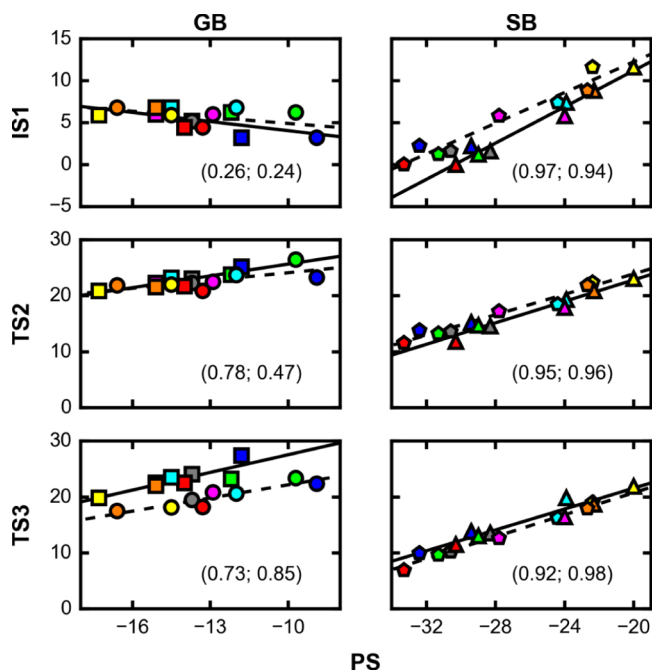
$$\Delta\Delta g_{TS3} = 0.95\Delta\Delta g_{PS} \tag{11}$$

*Comparison of the Calculated Energetics with the Available Experimental and Theoretical Data.* The lowest $\Delta G^{\ddagger}$ values for Polβ were obtained in the SB pathway: 12.5 kcal mol$^{-1}$ for dCTP·dG and 13.5 kcal mol$^{-1}$ for dGTP·dC nascent base pairs, which was 2.9−4.3 kcal mol$^{-1}$ below the experimental and 1.5−5.6 kcal mol$^{-1}$ below the theoretical values obtained from the literature (Table S51). The lowest $\Delta G^{\ddagger}$ values for tPolλ were obtained in the SB pathway: 17.9 kcal mol$^{-1}$ for the dCTP· dG nascent base pair, which was 0.4 kcal mol$^{-1}$ above the



**Figure 3.** Linear free energy relationships (LFERs, kcal/mol) among IS1, TS2, TS3, and PS in the GB and SB pathways of the nucleotidyl-transfer reaction in Polβ and Polλ systems. Gray, Polβ1; red, Polβ2; blue, Polβ3; green, Polβ4; cyan, Polλ1; magenta, Polλ2; yellow, Polλ3; orange, Polλ4. Squares and triangles, the reactions involving IS2a; circles and pentagons, the reactions involving IS2b. Solid lines, linear regressions for the reactions involving IS2a; dashed lines, linear regressions for the reactions involving IS2b; correlation coefficients ($R^2$) of the linear regressions are given in parentheses (solid line; dashed line).

experimental and 0.3−0.5 kcal mol$^{-1}$ above the theoretical (albeit for a different nascent base pair) values obtained from the literature (Table S52). The lowest $\Delta G^{\ddagger}$ values for tPol$\lambda\Delta$L1 were 21.0 kcal mol$^{-1}$ for dGTP·dC in the GB pathway and 20.9 kcal mol$^{-1}$ for dGTP·dT in the SB pathway; the latter was 1.7 kcal mol$^{-1}$ above the experimental value for tPol$\lambda$ obtained from the literature and represented the first reported prediction for tPol$\lambda\Delta$L1 for which no experimental pre-steady-state kinetic data exist. The values of $\Delta G_0$ for Pol$\beta$ and Pol$\lambda$ were 12.8−30.9 kcal mol$^{-1}$ below the corresponding values that had been reported in other theoretical studies.

**Geometries of the RS, IS, TS, and PS of the Nucleotidyl-Transfer Reaction.** Animations of the nucleotidyl-transfer reaction generated from the representative structures of each state showed similar atomic motions in the active site for all enzyme systems (Supporting Files S3). The geometric determinants of $\Delta G^{\ddagger}$ and $\Delta G_0$ were searched for in the structures of IS1, TS2, TS3, and PS.

*TS Structures.* $P_{\alpha}-O_{lg}$ and $O_{nuc}-P_{\alpha}$ distances in TS2 and TS3 differed in the pathways involving IS2a and IS2b (Figure 4;
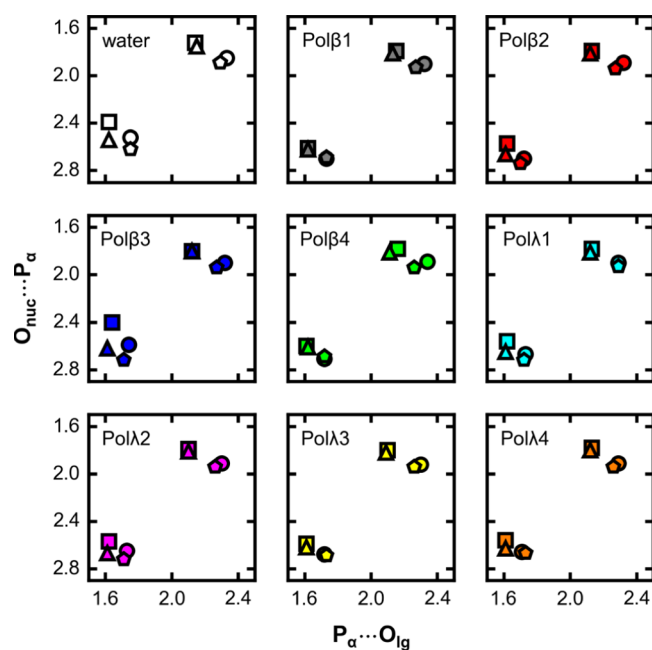


**Figure 4.** P−O distances (Å) in TS2 (lower left) and TS3 (upper right) in the GB (squares and circles) and SB (triangles and pentagons) pathways of the nucleotidyl-transfer reaction in the water, Pol$\beta$, and Pol$\lambda$ systems. Squares and triangles, TSs in the reactions involving IS2a; circles and pentagons, TSs in the reactions involving IS2b.

Tables S53−56). TS2a and TS3a featured shorter $P_{\alpha}-O_{lg}$ (by 0.11 and 0.17 Å on average) and $O_{nuc}-P_{\alpha}$ (by 0.09 and 0.12 Å on average) distances than TS2b and TS3b among all enzyme systems in the GB and SB pathways. $P_{\alpha}-O_{lg}$ distances in TS2 and TS3 were only marginally shorter (by 0.01 and 0.03 Å on average) and $O_{nuc}-P_{\alpha}$ distances were longer (by 0.06 and 0.03 Å on average) in the SB than the GB pathway. The water systems showed no difference in $P_{\alpha}-O_{lg}$ distances in TS2 between the GB and SB pathways but showed longer $O_{nuc}-P_{\alpha}$ in TS2 of the SB (than the GB) pathway (by 0.13 Å). Compared to water systems, enzymes showed: (1) similar $P_{\alpha}-O_{lg}$ distance (shorter by 0.02 Å on average); (2) longer $O_{nuc}-P_{\alpha}$ distance in TS2 (by 0.12 Å on average) and TS3 (by 0.06 Å on average) in the GB and SB pathways, with TS2a of Pol$\beta$3 in the GB pathway being an

outlier, having water-like $O_{nuc}-P_{\alpha}$ distance and also the highest $\Delta G^{\ddagger}$ (28.0 kcal mol$^{-1}$) among all enzyme systems.

TS2 structures in all enzyme systems were stabilized by a network of hydrogen-bonding interactions (Figures 5 and S8−S15). The Watson−Crick hydrogen-bonding interactions between complementary nucleobases were always present, as were the interactions between the $O_{3\beta}$ atom of dNTP and an
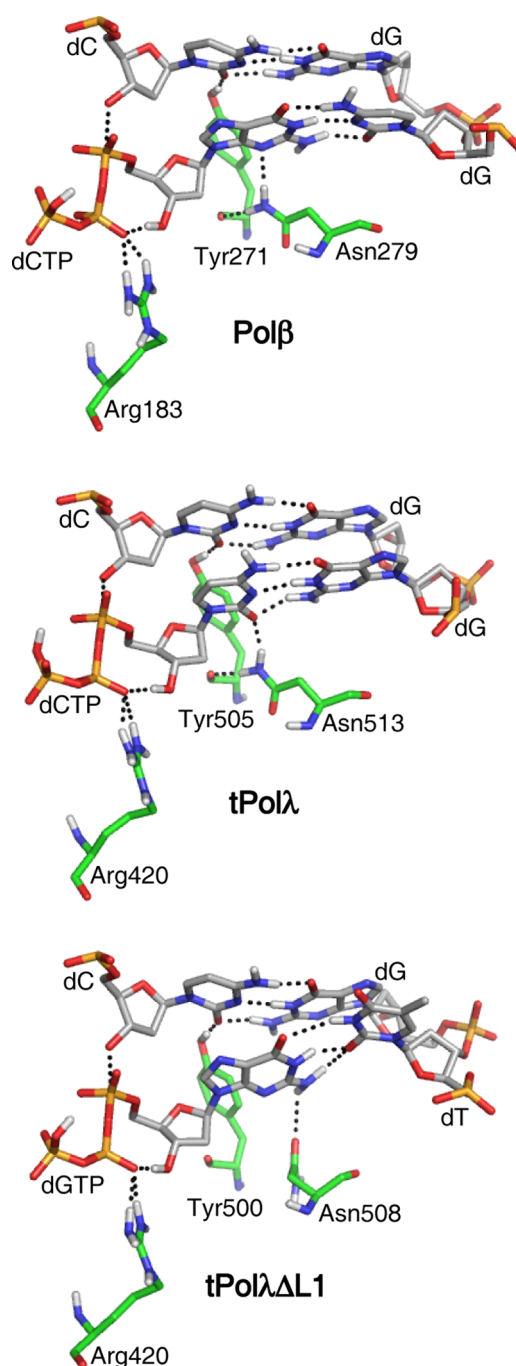


**Figure 5.** Representative TS2 structures in Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1 systems. Pol$\beta$, Pol$\beta$4 in SB2b pathway; tPol$\lambda$, Pol$\lambda$2 in SB2b pathway; tPol$\lambda\Delta$L1, Pol$\lambda$4 in SB2b pathway. Atoms are colored by element: green, carbon in amino acid residues; gray, carbon in nucleotides; red, oxygen; blue, nitrogen; orange, phosphorus; white, hydrogen. Hydrogens bound to carbon atoms are omitted for clarity. Dotted lines indicate hydrogen-bonding interactions and the partial $O_{nuc}-P_{\alpha}$ bond. See Figures S4−S11 for the full set of representative TS2 structures.

amino group of Arg183/420/420 in Polβ/tPolλ/tPolλΔL1 and between the hydroxyl group of Tyr271/505/500 in Polβ/tPolλ/ tPolλΔL1 and the carbonyl-group oxygen of the 3′-terminal dC of the primer DNA strand. The arginine could also interact with the hydroxyl group of dNTP, which could also form alternative interactions with the $O_{1\beta}$ atom of dNTP and with the carbonyl-group oxygen of the tyrosine which could simultaneously form a hydrogen bond with the amino group of Asn279/513/508 in Polβ/tPolλ/tPolλΔL1. The asparagine could also be involved in alternative hydrogen-bonding interactions with the nucleobases of the dNTP·dN nascent base pair.
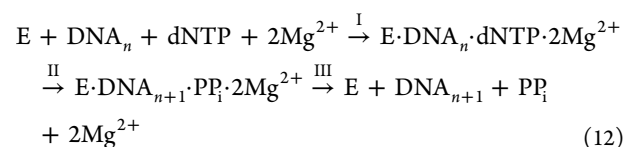
*PS Structures.* PS structures of all enzyme systems showed both similarities to and differences from the 4BU3 X-ray crystal structure of the Polβ product complex (Figures S16−S23). The similarities in the atomic positions included: (1) the nucleotides; (2) the $PP_i$ product; (3) the $Mg_B^{2+}$ ion coordinating Asp190/427/ 427 and Asp192/429/429 in Polβ/tPolλ/tPolλΔL1 with $PP_i$; and (4) Arg183/420/420 in Polβ/tPolλ/tPolλΔL1 interacting with the $O_{1\beta}$ atom of the inserted nucleotide. The differences in atomic positions included: (1) closer interaction between the amino groups of Arg149/386/386 and oxygen atoms of $PP_i$ in the PS structures; and (2) the $Mg_A^{2+}$ ion coordinating the $O_{nuc}$, Asp256/490/485, Asp190/427/427, and Asp192/429/429 in the PS structures of Polβ/tPolλ/tPolλΔL1 versus coordinating $O_{lg}$ of $PP_i$ and a phosphate-group oxygen of the inserted dNMP in the X-ray crystal structure; in the PS structures, the new crystallographic position of the $Mg_A^{2+}$ ion was occupied by a sodium ion (Polλ1 in the SB2 pathway involving IS2b) or (in all other cases) by water molecules.

*Cumulative Charge Profiles.* The simple geometrical analysis failed to explain the observed LFERs. Thus, this phenomenon had to be due to a complex combination of short- and long-range electrostatic interactions. But an attempt to quantify these interactions, as the charge distribution around $O_{nuc}$ in IS1 and around the $O_{3\beta}$ atom of $PP_i$ in PS, did not provide any hint about the source of the LFER (Figure S24): No obvious link between the cumulative charge profiles and $\Delta g_0$ could be established.

## ■ DISCUSSION

X-family DNA polymerases β and λ are involved in DNA damage repair and DNA recombination. In vitro and in vivo knockout experiments have demonstrated their vital roles: Polβ and Polλ double null mouse embryonic fibroblasts are hypersensitive to alkylating and oxidizing DNA-damaging agents;[15] Polβ knockout mice exhibit apoptosis in post-mitotic neuronal cells and die at birth;[79] heterozygous Polβ[+/−] mice develop normally, but have increased incidence of lymphomas;[80] and Polλ knockout mice are viable and fertile but display reduced immunoglobulin heavy-chain junction variability due to impaired nonhomologous end-joining.[81] Polβ and Polλ protect the integrity of the genome by being the principal nucleotidyltransferases in single-nucleotide gap-filling DNA synthesis during BER;[15] Polλ also contributes to generating the diverse repertoire of immunoglobulins and T-cell receptors via the gap-filling activity during V(D)J recombination.[81] When overexpressed in tumors, Polβ and Polλ can act as oncogenes by increasing the mutation rate and enhancing the capability of DNA repair.[82] Polβ and Polλ play their biological roles by dRP lyase and nucleotidyltransferase activities; we explore the latter.

Polβ and Polλ share the common mechanism of the nucleotidyl-transfer reaction:[19]

$$E + DNA_n + dNTP + 2Mg^{2+} \xrightarrow{I} E{\cdot}DNA_n{\cdot}dNTP{\cdot}2Mg^{2+}$$
$$\xrightarrow{II} E{\cdot}DNA_{n+1}{\cdot}PP_i{\cdot}2Mg^{2+} \xrightarrow{III} E + DNA_{n+1} + PP_i$$
$$+ 2Mg^{2+} \tag{12}$$

In this study, we modeled the chemical transformation of RS to PS (reaction II in eq 12) in Polβ, truncated Polλ (tPolλ), and Loop1 mutant of tPolλ (tPolλΔL1) using FEP and EVB simulations in three steps

$$RS \xrightarrow{TS1} IS1 \xrightarrow{TS2} IS2 \xrightarrow{TS3} PS \tag{13}$$

where IS1 and IS2 are the intermediate states, RS → IS1 is the proton transfer step (DNA-3′-OH → DNA-$O_{nuc}^-$ + $H^+$), and IS1 → IS2 → PS comprises the $O_{nuc}$−$P_\alpha$ bond formation and the $P_\alpha$−$O_{lg}$ bond cleavage reactions. Recent ab initio QM calculations of a model reaction between tetrahydrofuranol and methyl triphosphate have suggested that IS2 does not exist,[69] which means that the reaction scheme of the uncatalyzed reaction between 2′-deoxyribose (dRib) and dNTP can be simplified into

$$RS \rightarrow IS1 \rightarrow PS \tag{14}$$

We removed IS2 in the FEP/EVB free-energy profiles by increasing the reaction free energy $(\Delta g_0)$ of the IS1 → IS2 step by about 3 kcal mol$^{-1}$, which resulted in the free energy of IS2 being almost equal to TS2 and TS3. When we applied the same free-energy shift to IS2 in the free-energy profiles of the reactions catalyzed by Polβ and Polλ, we also obtained a flat free-energy surface from TS2 to TS3.

$$RS \xrightarrow{TS1} IS1 \xrightarrow{TS2\sim IS2\sim TS3} PS \tag{15}$$

The absence of IS2 in the water, Polβ, and Polλ systems means that, using the terminology offered by Lassila et al.,[83] the $O_{nuc}$−$P_\alpha$ bond formation and the $P_\alpha$−$O_{lg}$ bond cleavage reactions are concerted,[30,44,48,49,75,84] and, unlike T7 DNA polymerase,[57] the studied enzymes do not change the concertedness of the reactions. The broad, flat TS region of the IS1 → PS reaction in the water and enzyme systems accommodates a range of associative, pentavalent-intermediate-like geometries: from the "early" TS (with less $O_{nuc}$−$P_\alpha$ than $P_\alpha$−$O_{lg}$ bonding) to the "late" TS (with more $O_{nuc}$−$P_\alpha$ than $P_\alpha$−$O_{lg}$ bonding), and with the total bonding at the $P_\alpha$ atom between 1.2 (compact TS) and 0.9 (loose TS) relative to RS.

IS1, the proton transfer intermediate, might also be artificial as it is absent in QM/MM calculations of the nucleotidyl-transfer reaction catalyzed by tPolλ,[49] which (if true) simplifies the reaction mechanism into the most simple form possible

$$RS \xrightarrow{TS} PS \tag{16}$$

But if IS1 does not exist, then the rate-limiting step includes all chemical transformations; can the proton transfer to Asp490 be distinguished as the rate-limiting step in such a scenario?[49] (No, it cannot.) We assumed that IS1 does exist (eq 15),[43−48] which allowed us to view the proton transfer as a distinct step. We could thus explore various proton transfer mechanisms and assess their relative feasibility from the calculated activation $(\Delta g^{\ddagger})$ and reaction free energies $(\Delta g_0)$.

In the RS → IS1 reaction, the 3′-OH group of the 3′-terminal nucleotide of the primer DNA strand loses its proton; the oxygen becomes $O_{nuc}$, poised for the nucleophilic attack on the $P_\alpha$ atom of dNTP. The proton transfer is facilitated by the catalytic magnesium ion $(Mg_A^{2+})$ and by the availability of a suitable proton

acceptor. What are the requirements for the initial proton acceptor? It must (1) be thermodynamically "willing" to accept the proton, (2) be located in the vicinity of the 3′-OH group (or be able to migrate there), and (3) its protonation should not impede the subsequent reaction steps. In Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, there are six candidates that might satisfy these conditions (see Figure 2): (1) oxygen of a water molecule migrating to the active site from the bulk solution,[30,46,48,51] (2) the distal carboxylic oxygen of Asp256/490/485 ($O_{\delta 1}$), (3) the proximal carboxylic oxygen of Asp256/490/485 ($O_{\delta 2}$),[43,45,47,49,51] (4) the proximal carboxylic oxygen of Asp192/429/429 ($O_{\delta 2}$),[45,49] (5) oxygen of the water molecule that coordinates the $Mg_A^{2+}$,[44,45,49] and (6) a nonbridging oxygen of the $P_\alpha$-group of dNTP ($O_{\alpha 2}$).[42,45,49]

Two QM/MM studies compared multiple proton transfer pathways: In Pol$\beta$, Alberts et al.[45] found the $e \rightarrow c$ proton transfer to be more feasible than the direct proton transfer to acceptors $c$, $d$, or $f$; in tPol$\lambda$, Cisneros et al.[49] found the proton transfer to acceptor $c$ to be more feasible than the proton transfer to acceptors $d$, $e$, or $f$. Neither of these two studies, however, assessed the possibility of the proton transfer to the bulk solution (acceptor $a$), which requires calculating p$K_a$ of the 3′-OH group. The proton transfer to acceptor $a$ has so far been reported only for Pol$\beta$, using FEP or protein-dipoles-Langevin-dipoles linear-response approximation (PDLD/S-LRA) simulations, by Sucato et al.,[30] Xiang et al.,[46] Ram Prasad and Warshel,[48] and Matute et al.:[51] acceptor $a$ ($\Delta g_0$ of 4 kcal mol$^{-1}$) seems to be at least as good as $c$ ($\Delta g_0$ of 6 kcal mol$^{-1}$).[51] How about the remaining proton acceptors? We did not find any report about the feasibility of acceptor $b$. Acceptor $e$ was an intermediate for the transfer to $d$[44] or $c$[45] in QM/MM calculations with Pol$\beta$, or scored badly ($\Delta E^\ddagger \sim$ 35 kcal mol$^{-1}$) in QM/MM calculations with tPol$\lambda$.[49] And acceptor $f$ was first suggested for Pol$\beta$, based on pioneering QM calculations, by Abashkin et al.,[42] but the proton transfer to acceptor $f$ failed to provide reasonable activation potential energies in two subsequent QM/MM studies on Pol$\beta$[45] and tPol$\lambda$.[49] We therefore excluded $e$ and $f$ from our consideration and modeled the proton transfer to the four remaining acceptors: the SB proton transfer to the bulk solution ($a$) using FEP simulations, and the GB proton transfer to $b$, $c$, and $d$ using EVB simulations.

We found the acceptor $d$ unlikely ($\Delta g^\ddagger > 24$ kcal mol$^{-1}$) because of the long distance between $O_{nuc}$ and $O_{\delta 2}$ of Asp192/429/429 in Pol$\beta$/tPol$\lambda$/tPol$\lambda\Delta$L1 and because of steric hindrance and electrostatic repulsion by $Mg_A^{2+}$. The proton transfer to acceptor $b$ proceeded with better energetics ($\Delta g^\ddagger > 16$ kcal mol$^{-1}$) but required a conformational change of Asp256/490/485 in Pol$\beta$/tPol$\lambda$/tPol$\lambda\Delta$L1 for reducing the distance between the $O_{nuc}$ and the $O_{\delta 1}$ atom of the aspartate; the conformational change perturbed the interaction of the aspartate with $Mg_A^{2+}$. The proton transfer to acceptor $c$ proceeded smoothly, over the shortest distance, and without disturbing the interaction between $O_{\delta 2}$ of the aspartate and $Mg_A^{2+}$, which contributed to a better energetics: $\Delta g^\ddagger$ ($\Delta g_0$) was 10 (5), 13 (6), and 12 (6) kcal mol$^{-1}$ in Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, respectively. But we obtained similar, or even more favorable, energetics for the proton transfer to the bulk solution. This mechanism is structurally supported by the existence of the half-open active site, replete with water molecules, which form a continuous chain all the way to the bulk solution in crystal structures of ternary complexes in the so-called "closed conformation". Proton transfer reactions along such proton-acceptor chains are very fast,[71] which means that the corresponding activation free-energy

barriers are primarily determined by their reaction free energy. If we assume a small <2 kcal mol$^{-1}$ free-energy difference between $\Delta g^\ddagger$ and $\Delta g_0$ for the proton transfer to the bulk solution at pH 7, this pathway yields $\Delta g^\ddagger$ ($\Delta g_0$) of 3 (1), 9 (7), and 12 (10) kcal mol$^{-1}$ in Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, respectively. Thus, the proton transfer to the bulk solution (via Asp256/490/485, an active site water molecule, or a hydroxide anion) seems to be the proton transfer pathway of the least resistance.[51] Moreover, the subsequent IS1 $\rightarrow$ IS2 reaction step, which also contributes into the overall free-energy barrier, proceeds with lower $\Delta g^\ddagger$ when preceded by the SB proton transfer. Two questions remain unresolved: (1) What makes the proton transfer to the bulk solution superior over the proton transfer to the active site aspartate in Pol$\beta$ and tPol$\lambda$, but not in tPol$\lambda\Delta$L1? (2) What is the source of the large differences in $\Delta g_0$ of the proton transfer to the bulk solution among the three enzyme systems?

Although the seemingly simple proton transfer to the bulk water yielded large differences between the studied systems, the subsequent concerted $O_{nuc}$–$P_\alpha$ bond formation and $P_\alpha$–$O_{lg}$ bond cleavage reactions (IS1 $\rightarrow$ PS), modeled using three VB states and involving large changes in atomic positions and charge distributions, proceeded uniformly in Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, regardless of the initial coordinates (X-ray crystal structures: 2FMP, 2PXI, 2PFO, 2PFP, or 3PML) and the nascent base pair (dCTP·dG, dGTP·dC, or dGTP·dT): $\Delta g^\ddagger$ of 12 kcal mol$^{-1}$ and $\Delta g_0$ of −32 kcal mol$^{-1}$ for the reactions involving compact or loose TS geometries. The activation free-energy barrier of this step is 9, 6, and 0 kcal mol$^{-1}$ larger than the assumed $\Delta g^\ddagger$ of the proton transfer in Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1. Thus, the P−O bond formation/cleavage reaction is the rate-limiting step in Pol$\beta$ (in agreement with refs 30, 43, 46, 48, 51 and in disagreement with ref 44) and tPol$\lambda$ (in disagreement with ref 49), and rate-colimiting (with the proton transfer) in tPol$\lambda\Delta$L1

$$\text{rate} = \frac{k_1 k_2}{k_{-1}}[\text{RS}] \tag{17}$$

where $k_1$ is the rate constant of the RS $\rightarrow$ IS1 reaction, $k_{-1}$ is the rate constant of the RS $\leftarrow$ IS1 reaction, and $k_2$ is the rate constant of the IS1 $\rightarrow$ PS reaction. Because $k_2$ is a constant in Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, the relative reaction rates among these enzymes depend solely on the p$K_a$ of the $O_{nuc}$

$$\Delta G^\ddagger = \Delta g_{0,\text{RS}\rightarrow\text{IS1}} + \Delta g^\ddagger_{\text{IS1}\rightarrow\text{PS}} = \Delta g_{0,\text{RS}\rightarrow\text{IS1}} + 12$$

$$\text{kcal mol}^{-1} \tag{18}$$

Equation 18 describes an intermolecular LFER (not to be confused with the intramolecular Brønsted LFER) between $\Delta g_0$ of the proton transfer step and $\Delta G^\ddagger$ of the overall reaction with a slope of ~1. This idea can be tested experimentally using nonhydrolyzable dNTP analogues with the $O_{lg}$ atom replaced by a $CH_2$, $CF_2$, or NH group, which will trap the system in the RS $\rightleftharpoons$ IS1 reaction and thus allow the evaluation of the $K_{eq} = k_1/k_{-1}$ term in eq 17. If the values of $K_{eq}$ for the dNTP analogues correlate with the values of catalytic rate constants $k_{pol}$, measured for the corresponding native dNTPs, the proposed intermolecular LFER will be validated. The values of $\Delta\Delta G^\ddagger$ could then be calculated using linear-response simulations on the reactant and either TS[63,85] or IS1 structures.

The uniformly depressed $\Delta g^\ddagger$ of the rate-limiting P−O bond formation/cleavage step is also the main contribution (−11 kcal mol$^{-1}$) to the average catalytic effect of −17, −12, and −9 kcal mol$^{-1}$ in Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, which we obtained by comparing $\Delta G^\ddagger$ for the enzymes and for the reference reaction

between dRib and dNTP in solution. This catalytic contribution appears to be very robust with respect to small perturbations of the protein structure and even with respect to mispairing between the templating and incoming nucleobases.[86] When the closed, "catalytically competent", ternary complexes of DNA polymerases are used for the simulations, the fidelity control originates from the preceding steps: thermodynamics of dNTP binding[44,86−88] or the deprotonation kinetics of the 3′-OH group of the primer DNA strand[44] because these steps are more sensitive to "small" structural variations in the ternary complexes. When the enzyme carries out the reaction while in the (partially-)open, "catalytically incompetent", conformation,[89,90] the catalysis of the P−O bond formation/breakage reactions is greatly diminished.[46,85]

In search for the main source of the catalytic effect of Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, we explored the representative structures of the RS, IS, TS, and PS and asked what interactions are responsible for the higher stability of the TS in the enzyme compared to that in the solution. The characteristic feature of the active sites of Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1 is a network of favorable electrostatic (including hydrogen-bonding) interactions, which interconnects the three apparent binding sites for the nucleobase, dRib, and triphosphate groups of dNTP.[40,63,90,91] These interactions, present already in the RS and preserved during the entire nucleotidyl-transfer reaction, involve the sidechains of Arg149/386/386, Arg183/420/420, and Asn279/513/508, the backbone and sidechain of Tyr271/505/500, and the nucleobase of the templating nucleotide in Pol$\beta$/tPol$\lambda$/tPol$\lambda\Delta$L1. In the solution reaction, the role of the amino acid residues and the templating nucleotide is fulfilled only suboptimally by water molecules, which are incapable of establishing an electrostatically preorganized reaction cage.[32,87,92]

Assuming the validity of the TS theory (eq 7), we can make a direct comparison between the calculated and available experimental $\Delta G^{\ddagger}$ of the nucleotidyl-transfer reaction catalyzed by Pol$\beta$ and tPol$\lambda$: Our FEP/EVB simulations overestimate the catalytic effect of Pol$\beta$ by 3−4 kcal mol$^{-1}$[30,31,41,93−95] but give an accurate prediction of the catalytic effect of tPol$\lambda$.[29] Note that all our calculated $\Delta G^{\ddagger}$ values are genuine predictions because the only experimental inputs were the X-ray crystal structures of ternary complexes of Pol$\beta$, tPol$\lambda$, and tPol$\lambda\Delta$L1, p$K_a$ of dRib, and pH of maximum $k_{pol}$ in Pol$\beta$. In this light, the low single digit difference in $\Delta G^{\ddagger}$ might seem small except that $k_{pol}$ scales exponentially with $\Delta G^{\ddagger}$: the difference of 4 kcal mol$^{-1}$ in $\Delta G^{\ddagger}$ translates into >300-fold difference in $k_{pol}$. In addition, our $\Delta G^{\ddagger}$ values carry a cumulative uncertainty of 6 kcal mol$^{-1}$. The standard deviations of individual reaction steps that add up to this total uncertainty have been assessed by simulating the same reaction in the forward and reverse directions. Thus, the activation and reaction free energies for each reaction step, each reaction mechanism, and each enzyme were determined by at least four (and up to 16) separate 510 ps FEP or FEP/EVB simulations. Of course, larger number of longer simulations (at least by 1 order of magnitude) might be a way to improve the conformational sampling along the reaction coordinate and thus, a way to increase the accuracy and reduce the uncertainty of $\Delta G^{\ddagger}$ values. However, in studies aimed at exploring, quantitatively, complex enzymatic reaction mechanisms, the strategy of running multiple short simulations and averaging over very different simulation setups[96] is more cost-effective than running fewer but longer simulations for the same total simulation time. Previous QM/MM and FEP/EVB predictions for Pol$\beta$ were more accurate than ours, for unknown reasons, deviating only by 1

to 2 kcal mol$^{-1}$ from the experimental $\Delta G^{\ddagger}$.[30,44,46] The only computational prediction for tPol$\lambda$ was reported by Cisneros et al.,[49] who obtained an activation potential energy of 17.6 kcal mol$^{-1}$ using QM/MM calculations, that is, in perfect agreement with our FEP/EVB activation free energy of 17.9 kcal mol$^{-1}$.

The FEP/EVB approach, which was developed more than 35 years ago, remains underappreciated and underused by the molecular modeling community at large, despite multiple case studies demonstrating its robustness and versatility in modeling of enzyme catalysis. No molecular modeling tool can stand on its own and FEP/EVB is not an exception. But when complemented by QM and QM/MM calculations, linear-response simulations and, most importantly, by experiments (X-ray crystallography, NMR, pre-steady-state kinetics, kinetic isotope effect, Brønsted LFER, and our proposed intermolecular LFER), the FEP/EVB approach can provide a unique insight into the structure−dynamics−function−evolution relationships of DNA polymerases $\beta$ and $\lambda$ and into their (ab)normal roles in DNA repair and recombination.

## CONCLUSIONS

Assuming our computational model is a correct imitation of reality, the three introductory questions can be answered as follows: (1) What is the preferred mechanism for the deprotonation of the 3′-OH group of the primer DNA strand? The preferred mechanism for the deprotonation of the 3′-OH group of the primer DNA strand is the transfer of the proton to the bulk aqueous solution via a multitude of possible pathways, involving active site amino acid residues (including, but not limited to, Asp256/490/485 in Pol$\beta$/tPol$\lambda$/tPol$\lambda\Delta$L1), hydroxide anions accessing the active site, and/or chains of water molecules connecting the active site to the bulk solution. (2) What is the degree of concertedness of the P−O bond formation and cleavage reactions? The $O_{nuc}$−$P_\alpha$ bond formation and $P_\alpha$−$O_{lg}$ bond cleavage reactions are concerted, and proceed via an early, loose TS. (3) Which of the three reaction steps limits the overall reaction rate? Only two steps can be distinguished: proton transfer followed by $O_{nuc}$−$P_\alpha$ bond formation and $P_\alpha$−$O_{lg}$ bond cleavage; and whereas both these steps contribute to the overall activation barrier of the nucleotidyl-transfer reaction, the P−O bond formation/cleavage represents the rate-limiting step. The overall activation barrier (and hence $k_{pol}$), which can be perturbed by small alterations in the structures of the catalytically competent ternary complexes, is linearly proportional to the stability of the deprotonated nucleophilic oxygen, that is, to the p$K_a$ of the 3′-OH group of the primer DNA strand in the enzyme active site. The observed intermolecular LFER can be exploited, computationally or experimentally, for predicting the catalytic effects of nondeleterious mutations in DNA polymerases.

## ASSOCIATED CONTENT

**Ⓢ Supporting Information**

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b08581.

> Structural models in PDB format (Supporting File S1)
> Structural models in Q topology (TOP) format (Supporting File S2)
> Animations of the nucleotidyl-transfer reaction (Supporting File S3)
> EVB method (eqs S1−S19); FEP method (eqs S1−S5 and S20); structural models (Tables S1−S5; Figure S1); EVB parameters (Tables S6−S24); FEP parameters (Tables

S25–S28); free-energy profiles (Tables S29–S52; Figures S2–S7); TS geometry (Tables S53–S56; Figures S8–S15); PS geometry (Figures S16–S23); cumulative charge profiles (Figure S24) (PDF)

## ■ AUTHOR INFORMATION

**Corresponding Author**

*E-mail: jfloria@luc.edu. Phone: +1 773 508 3785. Fax: +1 773 508 3086.

**ORCID** 🔟

Martin Klvaňa: 0000-0002-5076-4936

**Author Contributions**

M.K. and J.F. designed the research; M.K. performed the calculations; M.K., U.B., and J.F. interpreted the data and wrote the article.

**Notes**

The authors declare no competing financial interest.

## ■ ABBREVIATIONS

BER, base excision repair; EVB, empirical valence bond; FEP, free-energy perturbation; GB, general base; LFER, linear free-energy relationship; MD, molecular dynamics; Pol$\beta$, human DNA polymerase $\beta$; Pol$\lambda$, human DNA polymerase $\lambda$; QM, quantum mechanics; SB, specific base; tPol$\lambda$, truncated Pol$\lambda$; tPol$\lambda\Delta$L1, the Loop1 mutant of tPol$\lambda$; TS, transition state

## ■ REFERENCES

(1) Koshland, D. E., Jr. Special essay. The Seven Pillars of Life. *Science* **2002**, *295*, 2215–2216.

(2) Moon, A. F.; García-Díaz, M.; Batra, V. K.; Beard, W. A.; Bebenek, K.; Kunkel, T. A.; Wilson, S. H.; Pedersen, L. C. The X Family Portrait: Structural Insights into Biological Functions of X Family Polymerases. *DNA Repair* **2007**, *6*, 1709–1725.

(3) Singhal, R. K.; Prasad, R.; Wilson, S. H. DNA Polymerase $\beta$ Conducts the Gap-Filling Step in Uracil-Initiated Base Excision Repair in a Bovine Testis Nuclear Extract. *J. Biol. Chem.* **1995**, *270*, 949–957.

(4) Nicholl, I. D.; Nealon, K.; Kenny, M. K. Reconstitution of Human Base Excision Repair with Purified Proteins. *Biochemistry* **1997**, *36*, 7557–7566.

(5) Fortini, P.; Pascucci, B.; Parlanti, E.; Sobol, R. W.; Wilson, S. H.; Dogliotti, E. Different DNA Polymerases Are Involved in the Short- and Long-Patch Base Excision Repair in Mammalian Cells. *Biochemistry* **1998**, *37*, 3575–3580.

(6) Wilson, S. H. Mammalian Base Excision Repair and DNA Polymerase Beta. *Mutat. Res.* **1998**, *407*, 203–215.

(7) Iwanaga, A.; Ouchida, M.; Miyazaki, K.; Hori, K.; Mukai, T. Functional Mutation of DNA Polymerase $\beta$ Found in Human Gastric Cancer—Inability of the Base Excision Repair In Vitro. *Mutat. Res.* **1999**, *435*, 121–128.

(8) García-Díaz, M.; Bebenek, K.; Kunkel, T. A.; Blanco, L. Identification of an Intrinsic 5′-Deoxyribose-5-Phosphate Lyase Activity in Human DNA Polymerase $\lambda$: A Possible Role in Base Excision Repair. *J. Biol. Chem.* **2001**, *276*, 34659–34663.

(9) Podlutsky, A. J.; Dianova, I. I.; Wilson, S. H.; Bohr, V. A.; Dianov, G. L. DNA Synthesis and dRPase Activities of Polymerase $\beta$ Are Both Essential for Single-Nucleotide Patch Base Excision Repair in Mammalian Cell Extracts. *Biochemistry* **2001**, *40*, 809–813.

(10) Sobol, R. W.; Wilson, S. H. Mammalian DNA Beta-Polymerase in Base Excision Repair of Alkylation Damage. *Prog. Nucleic Acid Res. Mol. Biol.* **2001**, *68*, 57–74.

(11) Mitra, S.; Boldogh, I.; Izumi, T.; Hazra, T. K. Complexities of the DNA Base Excision Repair Pathway for Repair of Oxidative DNA Damage. *Environ. Mol. Mutagen.* **2001**, *38*, 180–190.

(12) Cabelof, D. C.; Raffoul, J. J.; Yanamadala, S.; Guo, Z.; Heydari, A. R. Induction of DNA Polymerase $\beta$-Dependent Base Excision Repair in Response to Oxidative Stress In Vivo. *Carcinogenesis* **2002**, *23*, 1419–1425.

(13) Braithwaite, E. K.; Prasad, R.; Shock, D. D.; Hou, E. W.; Beard, W. A.; Wilson, S. H. DNA Polymerase $\lambda$ Mediates a Back-Up Base Excision Repair Activity in Extracts of Mouse Embryonic Fibroblasts. *J. Biol. Chem.* **2005**, *280*, 18469–18475.

(14) Liu, Y.; Prasad, R.; Beard, W. A.; Kedar, P. S.; Hou, E. W.; Shock, D. D.; Wilson, S. H. Coordination of Steps in Single-Nucleotide Base Excision Repair Mediated by Apurinic/Apyrimidinic Endonuclease 1 and DNA Polymerase $\beta$. *J. Biol. Chem.* **2007**, *282*, 13532–13541.

(15) Braithwaite, E. K.; Kedar, P. S.; Stumpo, D. J.; Bertocci, B.; Freedman, J. H.; Samson, L. D.; Wilson, S. H. DNA Polymerases $\beta$ and $\lambda$ Mediate Overlapping and Independent Roles in Base Excision Repair in Mouse Embryonic Fibroblasts. *PLoS One* **2010**, *5*, No. e12229.

(16) Skosareva, L. V.; Lebedeva, N. A.; Rechkunova, N. I.; Kolbanovskiy, A.; Geacintov, N. E.; Lavrik, O. I. Human DNA Polymerase $\lambda$ Catalyzes Lesion Bypass across Benzo[a]pyrene-Derived DNA Adduct during Base Excision Repair. *DNA Repair* **2012**, *11*, 367–373.

(17) Prasad, R.; Beard, W. A.; Strauss, P. R.; Wilson, S. H. Human DNA Polymerase $\beta$ Deoxyribose Phosphate Lyase. Substrate Specificity and Catalytic Mechanism. *J. Biol. Chem.* **1998**, *273*, 15263–15270.

(18) Allinson, S. L.; Dianova, I. I.; Dianov, G. L. DNA Polymerase $\beta$ Is the Major dRP Lyase Involved in Repair of Oxidative Base Lesions in DNA by Mammalian Cell Extracts. *EMBO J.* **2001**, *20*, 6919–6926.

(19) Sawaya, M. R.; Pelletier, H.; Kumar, A.; Wilson, S. H.; Kraut, J. Crystal Structure of Rat DNA Polymerase $\beta$: Evidence for a Common Polymerase Mechanism. *Science* **1994**, *264*, 1930–1935.

(20) Pelletier, H.; Sawaya, M. R.; Wolfle, W.; Wilson, S. H.; Kraut, J. Crystal Structures of Human DNA Polymerase $\beta$ Complexed with DNA: Implications for Catalytic Mechanism, Processivity, and Fidelity. *Biochemistry* **1996**, *35*, 12742–12761.

(21) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. Crystal Structures of Human DNA Polymerase $\beta$ Complexed with Gapped and Nicked DNA: Evidence for an Induced Fit Mechanism. *Biochemistry* **1997**, *36*, 11205–11215.

(22) DeRose, E. F.; Kirby, T. W.; Mueller, G. A.; Bebenek, K.; García-Díaz, M.; Blanco, L.; Kunkel, T. A.; London, R. E. Solution Structure of the Lyase Domain of Human DNA Polymerase $\lambda$. *Biochemistry* **2003**, *42*, 9564–9574.

(23) Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. Magnesium-Induced Assembly of a Complete DNA Polymerase Catalytic Complex. *Structure* **2006**, *14*, 757–766.

(24) García-Díaz, M.; Bebenek, K.; Krahn, J. M.; Pedersen, L. C.; Kunkel, T. A. Role of the Catalytic Metal During Polymerization by DNA Polymerase Lambda. *DNA Repair* **2007**, *6*, 1333–1340.

(25) McKenna, C. E.; Kashemirov, B. A.; Upton, T. G.; Batra, V. K.; Goodman, M. F.; Pedersen, L. C.; Beard, W. A.; Wilson, S. H. (R)-$\beta$,$\gamma$-Fluoromethylene-dGTP-DNA Ternary Complex with DNA Polymerase $\beta$. *J. Am. Chem. Soc.* **2007**, *129*, 15412–15413.

(26) Ahn, J.; Werneburg, B. G.; Tsai, M. D. DNA Polymerase $\beta$: Structure-Fidelity Relationship from Pre-Steady-State Kinetic Analyses of All Possible Correct and Incorrect Base Pairs for Wild Type and R283A Mutant. *Biochemistry* **1997**, *36*, 1100–1107.

(27) Matsuda, T.; Vande Berg, B. J.; Bebenek, K.; Osheroff, W. P.; Wilson, S. H.; Kunkel, T. A. The Base Substitution Fidelity of DNA Polymerase $\beta$-Dependent Single Nucleotide Base Excision Repair. *J. Biol. Chem.* **2003**, *278*, 25947–25951.

(28) Beard, W. A.; Wilson, S. H. Structural Insights into the Origins of DNA Polymerase Fidelity. *Structure* **2003**, *11*, 489–496.

(29) Fiala, K. A.; Abdel-Gawad, W.; Suo, Z. Pre-Steady-State Kinetic Studies of the Fidelity and Mechanism of Polymerization Catalyzed by Truncated Human DNA Polymerase λ. *Biochemistry* **2004**, *43*, 6751−6762.

(30) Sucato, C. A.; Upton, T. G.; Kashemirov, B. A.; Batra, V. K.; Martínek, V.; Xiang, Y.; Beard, W. A.; Pedersen, L. C.; Wilson, S. H.; McKenna, C. E.; et al. Modifying the β,γ Leaving-Group Bridging Oxygen Alters Nucleotide Incorporation Efficiency, Fidelity, and the Catalytic Mechanism of DNA Polymerase β. *Biochemistry* **2007**, *46*, 461−471.

(31) Brown, J. A.; Pack, L. R.; Sanman, L. E.; Suo, Z. Efficiency and Fidelity of Human DNA Polymerases λ and β During Gap-Filling DNA Synthesis. *DNA Repair* **2011**, *10*, 24−33.

(32) Warshel, A. Computer Simulations of Enzyme Catalysis: Methods, Progress, and Insights. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443.

(33) Loeb, L. A.; Monnat, R. J. DNA Polymerases and Human Disease. *Nat. Rev. Genet.* **2008**, *9*, 594−604.

(34) Lange, S. S.; Takata, K.; Wood, R. D. DNA Polymerases and Cancer. *Nat. Rev. Cancer* **2011**, *11*, 96−110.

(35) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.

(36) García-Díaz, M.; Bebenek, K.; Sabariegos, R.; Domínguez, O.; Rodríguez, J.; Kirchhoff, T.; García-Palomero, E.; Picher, A. J.; Juárez, R.; Ruiz, J. F.; et al. DNA Polymerase λ, a Novel DNA Repair Enzyme in Human Cells. *J. Biol. Chem.* **2002**, *277*, 13184−13191.

(37) Fiala, K. A.; Duym, W. W.; Zhang, J.; Suo, Z. Up-Regulation of the Fidelity of Human DNA Polymerase λ by its Non-Enzymatic Proline-Rich Domain. *J. Biol. Chem.* **2006**, *281*, 19038−19044.

(38) Bebenek, K.; García-Díaz, M.; Zhou, R.-Z.; Povirk, L. F.; Kunkel, T. A. Loop 1 Modulates the Fidelity of DNA Polymerase λ. *Nucleic Acids Res.* **2010**, *38*, 5419−5431.

(39) Bebenek, K.; Pedersen, L. C.; Kunkel, T. A. Replication Infidelity via a Mismatch with Watson-Crick Geometry. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 1862−1867.

(40) Kraynov, V. S.; Showalter, A. K.; Liu, J.; Zhong, X.; Tsai, M. D. DNA Polymerase β: Contributions of Template-Positioning and dNTP Triphosphate-Binding Residues to Catalysis and Fidelity. *Biochemistry* **2000**, *39*, 16008−16015.

(41) Sucato, C. A.; Upton, T. G.; Kashemirov, B. A.; Osuna, J.; Oertell, K.; Beard, W. A.; Wilson, S. H.; Florián, J.; Warshel, A.; McKenna, C. E.; et al. DNA Polymerase β Fidelity: Halomethylene-Modified Leaving Groups in Pre-Steady-State Kinetic Analysis Reveal Differences at the Chemical Transition State. *Biochemistry* **2008**, *47*, 870−879.

(42) Abashkin, Y. G.; Erickson, J. W.; Burt, S. K. Quantum Chemical Investigation of Enzymatic Activity in DNA Polymerase β. A Mechanistic Study. *J. Phys. Chem. B* **2001**, *105*, 287−292.

(43) Lin, P.; Pedersen, L. C.; Batra, V. K.; Beard, W. A.; Wilson, S. H.; Pedersen, L. G. Energy Analysis of Chemistry for Correct Insertion by DNA Polymerase β. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13294−13299.

(44) Radhakrishnan, R.; Schlick, T. Correct and Incorrect Nucleotide Incorporation Pathways in DNA Polymerase β. *Biochem. Biophys. Res. Commun.* **2006**, *350*, 521−529.

(45) Alberts, I. L.; Wang, Y.; Schlick, T. DNA Polymerase β Catalysis: Are Different Mechanisms Possible? *J. Am. Chem. Soc.* **2007**, *129*, 11100−11110.

(46) Xiang, Y.; Goodman, M. F.; Beard, W. A.; Wilson, S. H.; Warshel, A. Exploring the Role of Large Conformational Changes in the Fidelity of DNA Polymerase β. *Proteins* **2008**, *70*, 231−247.

(47) Lin, P.; Batra, V. K.; Pedersen, L. C.; Beard, W. A.; Wilson, S. H.; Pedersen, L. G. Incorrect Nucleotide Insertion at the Active Site of a G:A Mismatch Catalyzed by DNA Polymerase β. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5670−5674.

(48) Ram Prasad, B.; Warshel, A. Prechemistry Versus Preorganization in DNA Replication Fidelity. *Proteins* **2011**, *79*, 2900−2919.

(49) Cisneros, G. A.; Perera, L.; García-Díaz, M.; Bebenek, K.; Kunkel, T. A.; Pedersen, L. G. Catalytic Mechanism of Human DNA Polymerase λ with $Mg^{2+}$ and $Mn^{2+}$ from Ab Initio Quantum Mechanical/Molecular Mechanical Studies. *DNA Repair* **2008**, *7*, 1824−1834.

(50) Chaudret, R.; Piquemal, J.-P.; Cisneros, G. A. Correlation between Electron Localization and Metal Ion Mutagenicity in DNA Synthesis from QM/MM Calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 11239−11247.

(51) Matute, R. A.; Yoon, H.; Warshel, A. Exploring the Mechanism of DNA Polymerases by Analyzing the Effect of Mutations of Active Site Acidic Groups in Polymerase β. *Proteins* **2016**, *84*, 1644−1657.

(52) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.

(53) Warshel, A.; Weiss, R. M. An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. *J. Am. Chem. Soc.* **1980**, *102*, 6218−6226.

(54) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, 2nd ed.; Wiley: Hoboken, NJ, 1997.

(55) Eyring, H. The Activated Complex and the Absolute Rate of Chemical Reactions. *Chem. Rev.* **1935**, *17*, 65−77.

(56) Florián, J.; Goodman, M. F.; Warshel, A. Computer Simulation Studies of the Fidelity of DNA Polymerases. *Biopolymers* **2003**, *68*, 286−299.

(57) Florián, J.; Goodman, M. F.; Warshel, A. Computer Simulation of the Chemical Catalysis of DNA Polymerases: Discriminating between Alternative Nucleotide Insertion Mechanisms for T7 DNA Polymerase. *J. Am. Chem. Soc.* **2003**, *125*, 8163−8177.

(58) DeLano, W. *The PyMOL Molecular Graphics System*, Version 0.99rc6 (2006); Schrödinger, LLC.

(59) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486−501.

(60) Marelius, J.; Kolmodin, K.; Feierberg, I.; Åqvist, J. Q: A Molecular Dynamics Program for Free Energy Calculations and Empirical Valence Bond Simulations in Biomolecular Systems. *J. Mol. Graphics Modell.* **1998**, *16*, 213−225.

(61) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(62) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules [J. Am. Chem. Soc. 1995, 117, 51795197]. *J. Am. Chem. Soc.* **1996**, *118*, 2309.

(63) Klvaňa, M.; Jeřábek, P.; Goodman, M. F.; Florián, J. An Abridged Transition State Model to Derive Structure, Dynamics, and Energy Components of DNA Polymerase β Fidelity. *Biochemistry* **2011**, *50*, 7023−7032.

(64) Klvaňa, M.; Murphy, D. L.; Jeřábek, P.; Goodman, M. F.; Warshel, A.; Sweasy, J. B.; Florián, J. Catalytic Effects of Mutations of Distant Protein Residues in Human DNA Polymerase β: Theory and Experiment. *Biochemistry* **2012**, *51*, 8829−8843.

(65) King, G.; Warshel, A. A Surface Constrained All-Atom Solvent Model for Effective Simulations of Polar Solutions. *J. Chem. Phys.* **1989**, *91*, 3647−3661.

(66) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(67) Lee, F. S.; Warshel, A. A Local Reaction Field Method for Fast Evaluation of Long-Range Electrostatic Interactions in Molecular Simulations. *J. Chem. Phys.* **1992**, *97*, 3100−3107.

(68) Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. v; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(69) Zhang, Z.; Eloge, J.; Florián, J. Quantum Mechanical Analysis of Nonenzymatic Nucleotidyl Transfer Reactions: Kinetic and Thermody-

namic Effects of β-γ Bridging Groups of dNTP Substrates. *Biochemistry* **2014**, *53*, 4180−4191.

(70) Izatt, R. M.; Rytting, J. H.; Hansen, L. D.; Christensen, J. J. Thermodynamics of Proton Dissociation in Dilute Aqueous Solution. V. An Entropy Titration Study of Adenosine, Pentoses, Hexoses, and Related Compounds. *J. Am. Chem. Soc.* **1966**, *88*, 2641−2645.

(71) Eigen, M. Proton Transfer, Acid-Base Catalysis, and Enzymatic Hydrolysis. Part I: Elementary Processes. *Angew. Chem., Int. Ed. Engl.* **1964**, *3*, 1−19.

(72) Jencks, W. P.; Rogenstein, J. M. *Handbook of Biochemistry and Molecular Biology, Physical and Chemical Data*, 3rd ed.; CRC Press: Cleveland, United States of America, 1976; Vol. *1*, pp 305−351.

(73) Kerst, A. F. *Environmental Phosphorus Handbook*; John Wiley & Sons: New York, 1973; pp 265−279.

(74) Zalatan, J. G.; Herschlag, D. Alkaline Phosphatase Mono- and Diesterase Reactions: Comparative Transition State Analysis. *J. Am. Chem. Soc.* **2006**, *128*, 1293−1303.

(75) Rosta, E.; Kamerlin, S. C. L.; Warshel, A. On the Interpretation of the Observed Linear Free Energy Relationship in Phosphate Hydrolysis: A Thorough Computational Study of Phosphate Diester Hydrolysis in Solution. *Biochemistry* **2008**, *47*, 3725−3735.

(76) Blaskó, A.; Bruice, T. C. Recent Studies of Nucleophilic, General-Acid, and Metal Ion Catalysis of Phosphate Diester Hydrolysis. *Acc. Chem. Res.* **1999**, *32*, 475−484.

(77) Kato, K.-I.; Gonçalves, J. M.; Houts, G. E.; Bollum, F. J. Deoxynucleotide-polymerizing Enzymes of Calf Thymus Gland: II. Properties of the Terminal Deoxynucleotidyltransferase. *J. Biol. Chem.* **1967**, *242*, 2780−2789.

(78) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.

(79) Sugo, N.; Aratani, Y.; Nagashima, Y.; Kubota, Y.; Koyama, H. Neonatal Lethality with Abnormal Neurogenesis in Mice Deficient in DNA Polymerase β. *EMBO J.* **2000**, *19*, 1397−1404.

(80) Cabelof, D. C.; Ikeno, Y.; Nyska, A.; Busuttil, R. A.; Anyangwe, N.; Vijg, J.; Matherly, L. H.; Tucker, J. D.; Wilson, S. H.; Richardson, A.; et al. Haploinsufficiency in DNA Polymerase β Increases Cancer Risk with Age and Alters Mortality Rate. *Cancer Res.* **2006**, *66*, 7460−7465.

(81) Bertocci, B.; De Smet, A.; Weill, J.-C.; Reynaud, C.-A. Nonoverlapping Functions of DNA Polymerases μ, λ, and Terminal Deoxynucleotidyltransferase during Immunoglobulin V(D)J Recombination In Vivo. *Immunity* **2006**, *25*, 31−41.

(82) Albertella, M. R.; Lau, A.; O'Connor, M. J. The Overexpression of Specialized DNA Polymerases in Cancer. *DNA Repair* **2005**, *4*, 583−593.

(83) Lassila, J. K.; Zalatan, J. G.; Herschlag, D. Biological Phosphoryl-Transfer Reactions: Understanding Mechanism and Catalysis. *Annu. Rev. Biochem.* **2011**, *80*, 669−702.

(84) Kamerlin, S. C. L.; McKenna, C. E.; Goodman, M. F.; Warshel, A. A Computational Study of the Hydrolysis of dGTP Analogues with Halomethylene-Modified Leaving Groups in Solution: Implications for the Mechanism of DNA Polymerases. *Biochemistry* **2009**, *48*, 5963−5971.

(85) Xiang, Y.; Oelschlaeger, P.; Florián, J.; Goodman, M. F.; Warshel, A. Simulating the Effect of DNA Polymerase Mutations on Transition-State Energetics and Fidelity: Evaluating Amino Acid Group Contribution and Allosteric Coupling for Ionized Residues in Human Pol β. *Biochemistry* **2006**, *45*, 7036−7048.

(86) Florián, J.; Goodman, M. F.; Warshel, A. Computer Simulations of Protein Functions: Searching for the Molecular Origin of the Replication Fidelity of DNA Polymerases. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6819−6824.

(87) Florián, J.; Goodman, M. F.; Warshel, A. Theoretical Investigation of the Binding Free Energies and Key Substrate-Recognition Components of the Replication Fidelity of Human DNA Polymerase β. *J. Phys. Chem. B* **2002**, *106*, 5739−5753.

(88) Florián, J.; Warshel, A.; Goodman, M. F. Molecular Dynamics Free-Energy Simulations of the Binding Contribution to the Fidelity of T7 DNA Polymerase. *J. Phys. Chem. B* **2002**, *106*, 5754−5760.

(89) Batra, V. K.; Beard, W. A.; Pedersen, L. C.; Wilson, S. H. Structures of DNA Polymerase Mispaired DNA Termini Transitioning to Pre-catalytic Complexes Support an Induced-Fit Fidelity Mechanism. *Structure* **2016**, *24*, 1863−1875.

(90) Liu, M.-S.; Tsai, H.-Y.; Liu, X.-X.; Ho, M.-C.; Wu, W.-J.; Tsai, M.-D. Structural Mechanism for the Fidelity Modulation of DNA Polymerase λ. *J. Am. Chem. Soc.* **2016**, *138*, 2389−2398.

(91) Martínek, V.; Bren, U.; Goodman, M. F.; Warshel, A.; Florián, J. DNA polymerase β catalytic efficiency mirrors the Asn279−dCTP H-bonding strength. *FEBS Letters* **2007**, *581*, 775−780.

(92) Bren, U.; Martínek, V.; Florián, J. Free Energy Simulations of Uncatalyzed DNA Replication Fidelity: Structure and Stability of T•G and dTTP•G Terminal DNA Mismatches Flanked by a Single Dangling Nucleotide. *J. Phys. Chem. B* **2006**, *110*, 10557−10566.

(93) Ahn, J.; Kraynov, V. S.; Zhong, X.; Werneburg, B. G.; Tsai, M. D. DNA Polymerase β: Effects of Gapped DNA Substrates on dNTP Specificity, Fidelity, Processivity and Conformational Changes. *Biochem. J.* **1998**, *331*, 79−87.

(94) Vande Berg, B. J.; Beard, W. A.; Wilson, S. H. DNA Structure and Aspartate 276 Influence Nucleotide Binding to Human DNA Polymerase β. Implication for the Identity of the Rate-Limiting Conformational Change. *J. Biol. Chem.* **2001**, *276*, 3408−3416.

(95) Oertell, K.; Chamberlain, B. T.; Wu, Y.; Ferri, E.; Kashemirov, B. A.; Beard, W. A.; Wilson, S. H.; McKenna, C. E.; Goodman, M. F. Transition State in DNA Polymerase β Catalysis: Rate-Limiting Chemistry Altered by Base-Pair Configuration. *Biochemistry* **2014**, *53*, 1842−1848.

(96) Lawrenz, M.; Baron, R.; McCammon, J. A. Independent-Trajectories Thermodynamic-Integration Free-Energy Changes for Biomolecular Systems: Determinants of H5N1 Avian Influenza Virus Neuraminidase Inhibition by Peramivir. *J. Chem. Theory Comput.* **2009**, *5*, 1106−1116.