Supplementary Information

Figure 6: Pol β's Reaction Pathway

General pathway for nucleotide insertion by DNA polymerases (a), and corresponding crystal closed (b, red) and open (c, green) conformations of a pol β/DNA complex. E: DNA polymerase; dNTP: 2′-deoxyribonucleoside 5′-triphosphate; PP_i: pyrophosphate; DNA_n/DNA_{n+1}: DNA before/after nucleotide incorporation to DNA primer. Binding of correct nucleotides (step 1) induces a conformational change from an open to a closed pol β state (step 2); incorrect units may inhibit or alter this change. The 'thumb' in left-handed pol β corresponds to the 'fingers' in right-handed DNA polymerases. After nucleotide incorporation (step 3), a closed-to-open rearrangement (step 4) occurs prior to the release of product pyrophosphate (step 5). (Figure kindly provided by Linjing Yang).

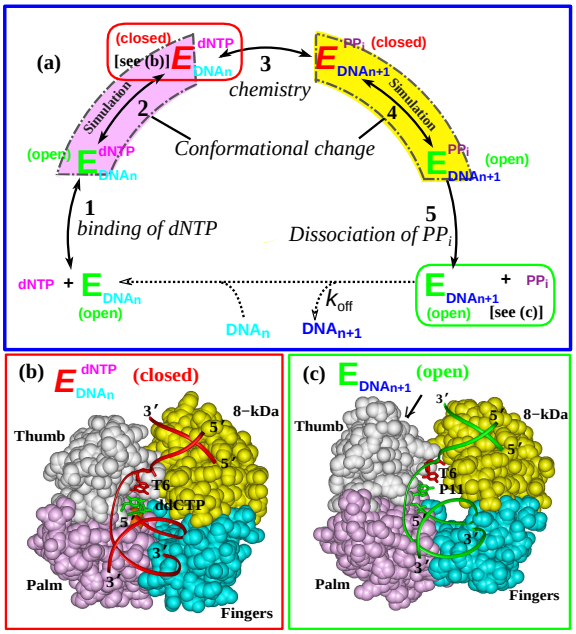

Figure 6: Pol β's reaction pathway.

Figure 7: Convergence Analysis

Left: For four transition state (TS) regions, 5 sample trajectories out of 150–200 harvested in path sampling are shown. The dihedral-angle interval between the dashed lines in TS 3 defines the TS ensemble for Arg258's partial-rotation. **Right (top):** Order parameter autocorrelation functions $\langle \chi_i(0)\chi_i(t)\rangle$ (in units of Å^2 for TS 1, and rad^2 for TS 2–4), where $\langle\cdot\rangle$ denotes the average over the ensemble of generated trajectories. Autocorrelation functions are plotted with initial point $\langle \chi_i(0)\chi_i(0)\rangle \approx \langle \chi_A\rangle^2$ and end point $\langle \chi_i(0)\chi_i(\tau)\rangle \approx \langle \chi_A\rangle\langle \chi_B\rangle$, to indicate crossing the barrier region between A and B over time τ (see Fig. 8 in Appendix A); χ_4 was shifted by 180° before computing $\langle \chi_4(0)\chi_4(t)\rangle$ to include in the same plot. **Right (bottom):** commitment probability distributions (CPDs) for TS regions 1–4 for initial (left) and converged TPS (right) trajectories, each calculated from 150–200 trajectories at a resolution of 0.25.

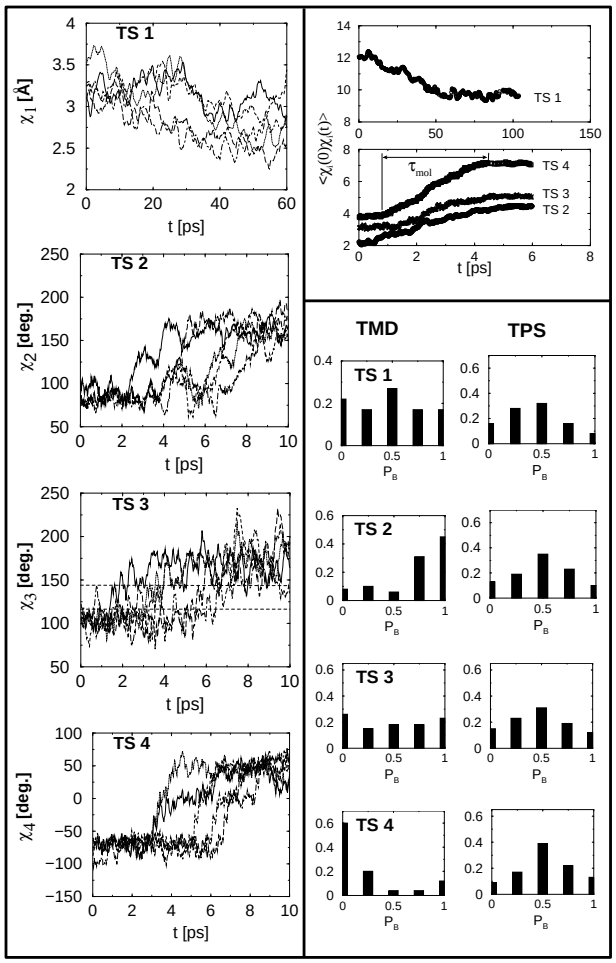

Figure 7: Convergence analysis

Appendix 1: Transition Path Sampling – Background

Transition path sampling (TPS) of Chandler and coworkers [1] aims to capture rare events (excursions or jumps between metastable basins in the free energy landscape) in molecular processes by essentially performing Monte Carlo sampling of symplectic dynamics trajectories, for which acceptance or rejection is determined by selected statistical criteria that characterize the ensemble of trajectories. Method details, beautiful illustrations, and applications to relatively small systems are available [2–5]. Here we only recapitulate essential features.

TPS exploits the separation of timescales in rare-event processes (e.g., long timescale $\tau_{\rm long}$ to bring a system to the top of the free-energy barrier compared to the short timescale $\tau_{\rm short}$ for dynamics within the barrier region) and saddle-like character of the free-energy landscape at transition-state (TS) regions to connect the various free energy basins. Namely, starting from an initial trajectory that captures a barrier crossing — which can be generated by algorithms that use guiding fields — TPS employs Metropolis Monte Carlo (MC) sampling of segments of (reversible and symplectic) molecular dynamics (MD) trajectories longer than $\tau_{\rm short}$ but shorter than $\tau_{\rm long}$. Despite the unphysical nature of the initial sampling trajectories, the protocol converges to yield physically meaningful trajectories passing through the saddle region. TPS samples different molecular dynamics trajectories using the shooting algorithm [6] (which perturbs initial momenta of atoms in a randomly chosen time interval, subject to the conservation of Maxwellian distribution of velocities, total linear and angular momentum, and detailed balance) to perform random walk steps in the space of trajectories, accepted or rejected according to selected statistical criteria (given by a path action as described below, see Eq. **3**) that characterize the ensemble of trajectories [1].

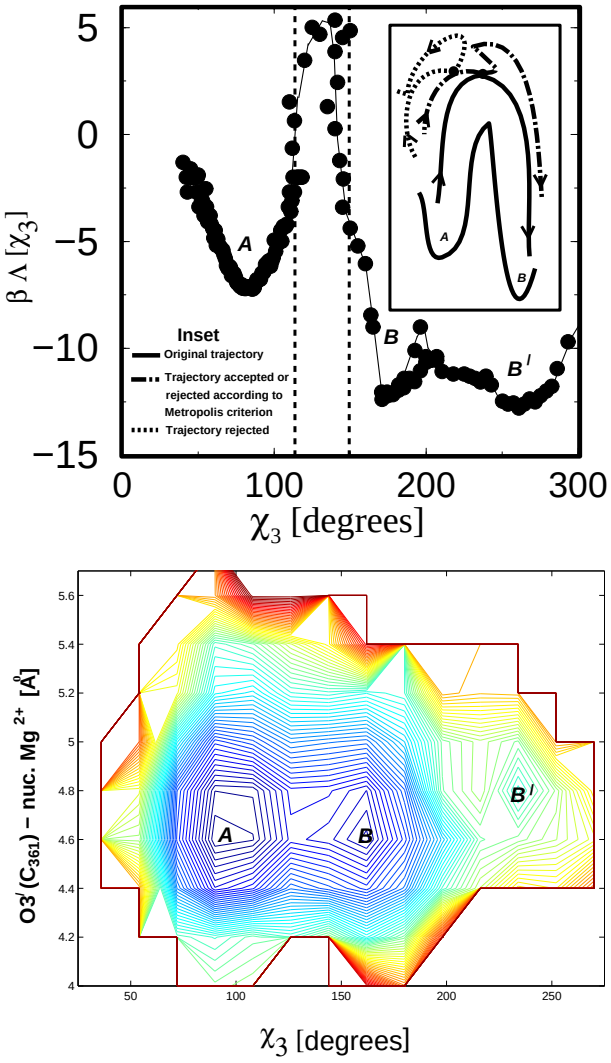

Figure 8: Free energy landscape with basins for Arg-258 rotation: A (unrotated); B (partially rotated); B' (fully rotated). The potential of mean force $\Lambda(\chi_3)$ (Upper) is obtained using umbrella sampling, and the two-dimensional free energy landscape (Lower) is obtained from TPS simulations. Ordinate (Lower) corresponds to distance between the nucleotide binding Mg^{2+} ion and the oxygen O3′ of the last primer (cytosine) residue. The dihedral angle space between the dashed lines (Upper) defines the TS ensemble for Arg-258's partial rotation. *Inset* illustrates how TPS trajectories are harvested (see text).

The shooting algorithm [6] generates an ensemble of molecular dynamics trajectories connecting two local minima (metastable states) A and B (see Fig. 8) in a free energy landscape via Monte Carlo sampling. For a given dynamics trajectory, the state of the system (i.e., basin A or B) is characterized by defining a set of order parameters $\chi = \{\chi_1, \chi_2, \cdots\}$. These order parameters are geometric quantities such as dihedral angles, bond distances, rms deviations of selected residues with respect to a reference structure, and so on. For biomolecules, as we show later, the key to a successful TPS application is identifying these key variables. Here, the groundwork simulations were important [7–9]. To formally identify a basin, *the population operator* h_A indicates if a particular molecular configuration associated with a time t of a molecular dynamics trajectory belongs to basin A:

$$h_A(\chi(t)) = \begin{cases} 1 & \text{if } \chi(t) \in A, \\ 0 & \text{otherwise.} \end{cases} \quad (\mathbf{1})$$

The *trajectory operator* H_B identifies a visit to basin B in a trajectory of length τ:

$$H_B\{\chi\}_\tau = \begin{cases} 1 & \text{if there exists } 0 < t < \tau \text{ such that } h_B(t) = 1, \\ 0 & \text{otherwise} \end{cases} \quad (\mathbf{2})$$

The idea in TPS is to generate many trajectories that connect A to B from one such existing pathway (see Figs. 7 and 8 in text). This is accomplished by a Metropolis algorithm that generates an ensemble of trajectories $\{\chi\}_\tau$ of length τ according to a path action $S\{\chi\}_\tau$ given by:

$$S\{\chi\}_\tau = \rho(0)\, h_A(\chi_0)\, H_B\{\chi\}_\tau, \quad (\mathbf{3})$$

where $\rho(0)$ is the probability of observing the configuration at $t = 0$ ($\rho(0) \propto \exp(-\beta E(0))$, in the canonical ensemble). A new trajectory $\{\chi^*\}_\tau$ is generated from an existing trajectory $\{\chi\}_\tau$ using the shooting algorithm [6], by perturbing the momenta of atoms at a randomly chosen time. The perturbation scheme is symmetric, i.e., the probability of generating a new set of momenta from the old set is the same as the reverse probability of generating the old set from the new set. Moreover, the scheme conserves the equilibrium distribution of momenta and the total linear momentum (and, if desired, total angular momentum). The acceptance probability implied by the above procedure is given by

$$P_{\text{acc.}} = \min\left[1, S\{\chi^*\}/S\{\chi\}\right]. \quad (\mathbf{4})$$

Together, these criteria ensure preservation of detailed balance, and thus according to the Metropolis algorithm [10], generate an ensemble of trajectories consistent with the path action S.

The ergodicity and convergence of each TPS run is monitored by calculating the autocorrelation function of the order parameter $\langle \chi_i(0)\chi_i(t)\rangle$ (Fig. 7) associated with each transition state i, where $\langle\cdot\rangle$ denotes the average over the ensemble of generated trajectories. In each

case, the autocorrelation function is plotted from $\langle \chi_i(0)\chi_i(0)\rangle \approx \langle \chi_A \rangle^2$ to the time τ where $\langle \chi_i(0)\chi_i(\tau)\rangle \approx \langle \chi_A \rangle \langle \chi_B \rangle$; this range is spanned during our sampling time τ, indicating that the transition state regions between A and B are crossed during this interval (see Fig. 8). The gradual transition of the autocorrelation functions between these values indicates decorrelation of the generated trajectories in each TPS run; strongly correlated trajectories would lead to an abrupt change in the correlation function for the chosen values of τ. The characteristic relaxation time τ_{mol} associated with the crossing of each TS region is given by the time taken for the gradual transition of the autocorrelation function $\langle \chi_i(0)\chi_i(t)\rangle$, as shown in Fig. 7 (see also Table 1 of main text). The value of τ_{mol} provides an estimate for the timescale associated with barrier crossing at the transition state region (see Appendix 4); the length τ of the MD trajectories in TPS should thus be greater than the transient time to commit to a basin, i.e., $\tau > \tau_{\mathrm{mol}}$.

Conserving the path action as described above both conserves the equilibrium distributions of the individual (metastable) states and ensures that the accepted molecular dynamics trajectories connect the two metastable states in question (Fig. 7). The shooting algorithm based on the Metropolis scheme (e.g., ref. [10]) conserves microscopic reversibility. Taken together, starting from a initial trajectory consistent with the path action, TPS generates an ensemble of MD trajectories guaranteed to converge to the correct ensemble as defined by the path action. The ensemble of trajectories represents configurations that constitute the correct pathway for hopping between the metastable states. Based on work by Crooks [11] and Jarzynski [12], we have shown (unpublished work) that the TPS of symplectic MD trajectories obeying microscopic reversibility satisfies the fluctuation-dissipation theorem [13].

Appendix 2: Transition Path Sampling – Biomolecular Applications

In general, harvesting mechanistic pathways by TPS for large biomolecular systems requires: (i) a robust protocol to generate initial trajectories, (ii) a careful identification of the different transition state regions, (iii) an implementation of the TPS sampling code to work with standard and well established integrators and force fields such as CHARMM [14], (iv) development and application of tests to assess convergence of TPS, and (v) reliable procedures for computing the reaction free energy pathway. We use a divide-and-conquer approach to implement the above steps for biomolecular systems, as we describe in turn.

(i) Generating initial trajectories that capture rare events is accomplished via targeted molecular dynamics (TMD) [15,16], high-temperature MD, and umbrella sampling [17–20] that use guiding (altered) fields. For example, we generate trajectories starting from the crystal open structure that capture pol β's closing using TMD on a pol β/DNA/dCTP complex with explicit water, salt, and magnesium ions [9] and similarly from the closed structure capturing the opening [7]. The TMD simulations were performed by introducing an additional restraint force as implemented in CHARMM based on the rms distance with respect to the closed polymerase conformation (1BPY). The functional form of the rms restraint energy is as follows:

$$E_{\rm rms} \;=\; K\;\left[D_{\rm rms}(X(t), X^{\rm target}) - d_0\right]^2, \qquad (5)$$

where K is a force constant, $D_{\rm rms}$ represents the relative rms distance for a selected set of atoms between the instantaneous conformation $X(t)$ and the reference $X^{\rm target}$, and d_0 is an offset constant (in Å). The restraint force is applied only to the heavy atoms in pol β. With a decrease in d_0 as a function of simulation interval, the conformational change is driven from the initial (open) to final (closed) conformation, and many configurations are generated for inspection.

(ii) Identifying different transition state (TS) regions is based on analysis of histograms of variables characterizing the motions of key residues (i.e., dihedral angles, distances, etc.) in the initial trajectories [free and TMD trajectories of the opening [7] and closing [9] as described above]. Although the initial trajectory is likely far from the physical pathway, a slice of the free energy landscape far from the actual path often displays similar characteristics (e.g., specific slow local motions) as that along the reaction coordinate. As a first approximation, the number of independent variables χ_i that display a bimodal distribution is used to characterize the various TS regions (see Table 1).

The next step is to confirm the existence of the TS regions by initiating a series of short (of order 10–100 ps, see Table 1), unrestrained MD trajectories from the different TS regions and calculating the "commitment probability distribution" (CPD) functions (i.e., probability that a particular trajectory will commit to a particular proximal free energy basin) for each TS region [2]. The CPDs are calculated by first choosing an ensemble of molecular configurations corresponding to a TS region (also called the TS ensemble). For example, in our application, the TS ensemble for the partial rotation of Arg-258 was defined as a window of dihedral-angle space corresponding to the barrier region as shown in Figs. 8 and 7. For each configuration in such a TS ensemble, four trajectories are initiated with a randomly

chosen set of momenta from a Maxwell distribution, and the commitment probability P_B is determined as the fraction of trajectories committing to basin B. The frequency distribution of P_B for the configurations in the TS ensemble yields the CPD. In general, the CPD for each TS region of the initial (unphysical) trajectory is bimodal with a minimum at 1/2 and maxima at 0 and 1 [1] (Fig. 7).

(iii) TPS is implemented using CHARMM version C28a2 by using shooting and shifting [1,21] algorithms and configurational bias to enhance the Metropolis sampling while preserving the detailed balance criteria. The shooting and shifting moves are coded in a PERL script that calls CHARMM for trajectory generation. The code is designed to handle multiple TS regions that are characterized by sets of arbitrary dihedral angles, distances, and any configurational quantity calculable within CHARMM. The Verlet integrator in CHARMM with a time step of 1 fs is used for generating the individual molecular dynamics trajectories in TPS. Electrostatic and van der Waals interactions are smoothed to zero at 12 Å, and periodic boundary conditions are used in three dimensions during the dynamics simulations.

(iv) Testing for convergence of TPS involves (1) calculating the autocorrelation functions associated with order parameters to check for the decorrelation of paths. Paths can be considered decorrelated if the autocorrelation function shows a gradual transition (increase or decrease of the value) between approximately $\langle\chi_A\rangle^2$ and $\langle\chi_A\rangle\langle\chi_B\rangle$.

The sampling quality can also be assessed by calculating order-parameter correlation functions in path space, $\langle\chi_i^*(0)\chi_i^*(n)\rangle_{\rm NS}$, where n represents the harvested trajectories, χ_i^* is the value of the order parameter evaluated at a particular time-slice at the bottleneck of the transition, and NS denotes no shifting with respect to the first trajectory (to remove the trivial decorrelation because of the shifting moves). Fig. 9 shows such a correlation function for the last 30 harvested trajectories for TS 1–4. It is evident from the figure that on an average, every 10th to 20th trajectory is statistically decorrelated; therefore the 150 to 200 trajectories that we generate for each TS ensure sufficiently good sampling.

(2) Recalculating CPD functions for each TS region to ensure that CPDs are unimodal with a peak at 1/2 (Fig. 7). The saddle region associated with each transition state is manifested by a unimodal CPD peaked around 1/2 [2]. (3) Extending a few harvested trajectories and confirming that the open and closed conformations of pol β are sampled. In our case, a global criterion such as the rms deviation of the heavy atoms in the enzyme with respect to the open (1BPX) and the closed (1BPY) crystal structures is used to quantify the proximity to the reference states. A failure to observe the correct bounds of the order parameter autocorrelation functions likely indicates inadequate sampling — MD segments are shorter than $\tau_{\rm short}$ — or use of incorrect variables to characterize the metastable basins. Pathological behavior in 2 above indicates that convergence of TPS has not been achieved, and/or a particular TS region in the initial trajectory is not a saddle in the actual pathway.

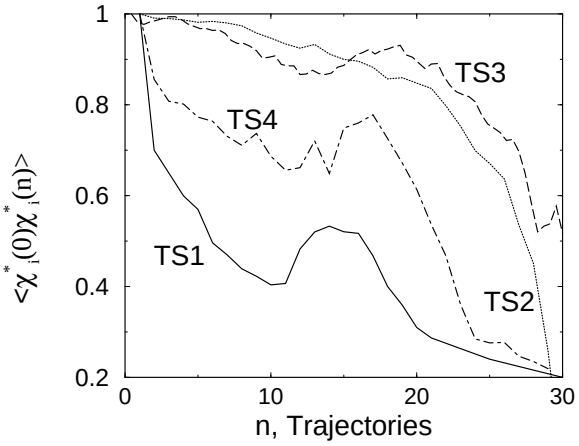

Figure 9: Decorrelation of transition paths for TS 1–4. The plots are normalized to begin at unity by normalizing with $\langle\chi_i^*(0)\chi_i^*(0)\rangle_{\rm NS}$.

Failure to sample the reference states in 3 suggests the presence of additional TS regions and requires additional path sampling near TS regions closest to open or closed conformations of pol β, after which steps **i**–**iv** above are repeated.

(v) Computing free energies along the reaction pathway is the most time consuming and important phase. We have reduced the time and statistical error by applying a novel procedure on the shooting and shifting algorithms rather than a biasing force as described separately (unpublished work). Essentially, the free energy barrier is determined by calculating the potential of mean force using histogram methods and a modified umbrella sampling scheme based on TPS. The order parameter variable χ_i characterizing each transition state i is divided into 6–10 windows. In each window, the probability distribution $P(\chi_i)$ is calculated from a series of modified path sampling runs performed using an appropriate guiding function (action) (unpublished work). The potential of mean force $\Lambda(\chi_i)$ in each window is given by ($\beta = 1/k_BT$):

$$\beta\Lambda(\chi_i) = -\ln[P(\chi_i)] + \text{constant}. \qquad (\mathbf{6})$$

The arbitrary constant associated with each window is adjusted by the method prescribed by Lynden-Bell and coworkers [22] to make the Λ function continuous. The overall free energy is calculated using the equation:

$$\exp(-\beta F) = \int_{\chi_{\min}}^{\chi_{\max}} \exp(-\beta\Lambda(\chi_i))\, d\chi_i, \qquad (\mathbf{7})$$

where the integration domain characterizes the metastable state. The free energies of all metastable states and barriers were calculated by numerically integrating Eq. **7** in the appropriate region.

A coarse-grained potential of mean force along each reaction coordinate is computed for each conformational event, as shown in Figure 10. The order parameter variable χ_i characterizing each transition state i is divided into 6–10 nonoverlapping windows. The probability distribution $P(\chi_i)$ is calculated from a series of path sampling runs (unpublished work) in each order parameter window. The histograms for each window were collected by running 75 trajectories per window.

The potential of mean force was calculated in each discrete window according to Eq. **6**. The slopes of the potential of mean force ($d\beta\Lambda(\chi_i)/d\chi_i$) were evaluated by a linear least squares fit to Λ vs. the χ_i data, in each window, as given in Table 2. The coarse-grained potentials of mean force along the reaction coordinates were calculated from the slopes by using a trapezoidal rule for integration along each χ_i (Fig. 10).

The free energy profiles presented in Fig. 10 involve two main approximations/errors: (i) the statistics obtained from our potential of mean force calculations resolve the free energy to within $\pm 3\, k_BT$. A more accurate resolution of the free energy would involve upwards of 300 trajectories per window. (ii) The slopes $d\beta\Lambda(\chi_i)/d\chi_i$ are assumed to be constant throughout each integration step in the trapezoidal rule. This assumption was deemed reasonable based on analyses of three model systems (unpublished work). This coarse-graining approximation can be circumvented by using overlapping windows as prescribed by Lynden-Bell and coworkers [22].

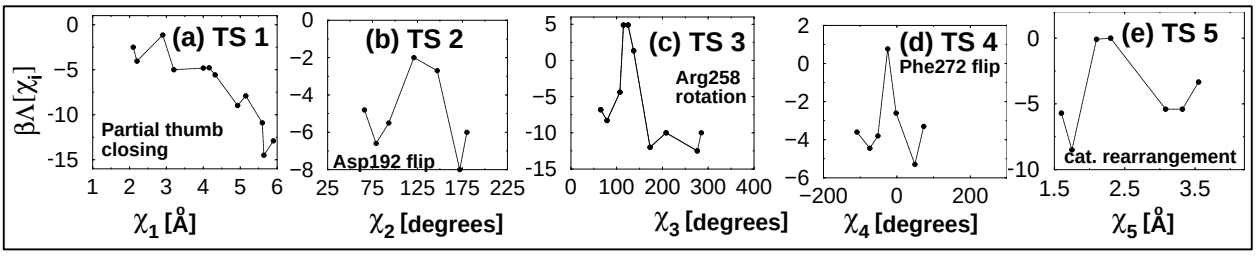

Figure 10: Approximate coarse-grained potential of mean force along the reaction coordinate for different transition state regions. (a) Partial thumb closing (TS 1). The reaction coordinate χ_1 is the rms deviation of the thumb residues (residues 275 to 295) with respect to the closed state. (b) Asp-192 flip (TS 2). The reaction coordinate χ_2 is the dihedral angle characterizing the flip of Asp-192. (c) Arg-258 rotation (TS 3). The reaction coordinate χ_3 is the dihedral angle characterizing the rotation of Arg-258. (d) Phe-272 flip (TS 4). The reaction coordinate χ_4 is the dihedral angle characterizing the flip of Phe-272. (e) Rearrangement of catalytic region and the stabilization of Arg-258 in the fully rotated state (TS 5). The reaction coordinate χ_5 is the distance between the nucleotide binding Mg^{2+} ion and the oxygen atom $O1_\alpha$ of dCTP.

Table 2: Coarse-grained potential of mean force calculations

χ_1 Å	$d\beta\Lambda/d\chi_1$ Å^{-1}	χ_2 deg.	$d\beta\Lambda/d\chi_2$ deg.$^{-1}$	χ_3 deg.	$d\beta\Lambda/d\chi_3$ deg.$^{-1}$	χ_4 deg.	$d\beta\Lambda/d\chi_4$ deg.$^{-1}$	χ_5 Å	$d\beta\Lambda/d\chi_5$ Å^{-1}
2.2	−8.3	72.6	−0.10	59.0	−0.15	−98.0	−0.03	1.7	−20.0
2.8	4.5	85.6	0.08	96.0	0.13	−59.0	−0.03	1.9	22.2
3.1	−15.5	106.0	0.13	112.0	0.92	−45.0	0.18	2.3	0.0
3.6	0.3	134.0	−0.03	119.0	0.00	−15.0	−0.14	2.9	−7.3
4.4	−2.3	159.0	−0.21	142.0	−0.29	28.0	−0.05	3.2	0.0
4.8	−4.2	175.0	0.25	163.0	−0.43	65.0	0.09	3.6	10.1
5.2	5.0			181.0	0.04				
5.5	−4.6			230.0	−0.21				
5.6	−56.2			282.0	0.11				
5.8	4.6								

Steps **i**–**iv** define a self-consistent algorithm for dealing with multiple transition states of the free-energy landscape in complex systems. Together with step **v**, they can yield the overall rate of transition as described in Appendix 4.

Appendix 3: Protonation States

The protonation states of the titratable side chain groups in the enzyme were chosen based on their individual pKa values and consistent with a solution pH of 7.0 as reported in Table 3. Indeed, in the open crystal complex the three conserved Asp groups are well separated from each other and not closely interacting with the dCTP, and therefore this choice of the protonation state based on pKa of the amino acid group and an overall pH of 7.0 is reasonable.

Table 3: Protonation states of amino acids in pol β

Residue	Charge	pKa
Asp	-1	3.9
Glu	-1	4.3
His	0	6.5
Lys	$+1$	10.8
Arg	$+1$	12.5

Still, a body of recent simulation data suggests that the protonation states are far from clear. In ref. [26], the authors show on the basis of a truncated model of the active site in *ab-initio* calculations, density functional theory (DFT) functionals, and specific basis-set used that the geometry could only be optimized if the assumption that Asp-192 was protonated was made. A report by a different group [27] on the same system, truncated pol β active site claims that geometries can be optimized using high-level DFT without assuming that Asp-192 is protonated. These contrasting observations may reflect artifacts of truncating the active site and ignoring the rest of the protein/DNA/solvent environment.

Note also that the protonation state may change as the conformational change occurs. In classical simulations, it is not possible to allow this change in a physically consistent manner, and that is part of the inherent limitations of classical force fields. Further studies are needed to establish the protonation states. These may include quantum mechanics/ molecular mechanics (QM/MM) simulations in which the titratable residue is included in the quantum mechanical (QM) part and then the free energy change associated with protonation/deprotonation reaction is computed. Alternatively, the relative free energies of the protonated and nonprotonated states could be computed using molecular mechanics/ Poisson-Boltzmann surface area (MM/PBSA) or similar methods to determine the pKa.

Appendix 4: Calculating Reaction Rates

Here, we outline how to estimate the rates [based on transition state theory [23]] associated with the transitions between adjacent metastable states in our overall free energy profile (Fig. 4, upper center). We also outline a procedure for obtaining the correction to the transition state theory approximation.

The free energies of the different metastable states and transition-state regions relative to the open and closed states can be obtained from the potential of mean force calculations, see Appendix 2). Using transition state theory [23], the rate of the transition between adjoining metastable states in Fig. 4 is given by

$$k_{\rm TST}^{A\to B} = \frac{1}{\tau_{\rm mol}} \exp(-\beta \Delta F_{AB}^{\rm barrier}), \qquad (\mathbf{8})$$

where $\tau_{\rm mol}$ is the time to cross the transition-state region and commit to basin B, and $\Delta F_{AB}^{\rm barrier}$ is the free energy of the transition-state region between basins A and B relative to basin A. For example, considering the adjacent states A and B as metastable states 3 and 4 (separated by TS 2) in Fig. 4, $\Delta F_{AB}^{\rm barrier} = F(\text{TS 2}) - F(A)$ and $\Delta F_{BA}^{\rm barrier} = F(\text{TS 2}) - F(B)$; Eq. **8** is then used to compute $k^{A\to B}$ and $k^{B\to A}$ associated with TS 2.

In the ideal gas approximation, the pre-factor $1/\tau_{\rm mol} = k_B T/h$, where h is the Plancks constant. In the reactive flux formalism [24], an estimate for $\tau_{\rm mol}$ is given by $w/\langle|\dot{q}|\rangle^*$, where w is the characteristic width to be crossed along the reaction coordinate q, and $\langle|\dot{q}|\rangle^*$ is the rate of change of the reaction coordinate at the transition state surface.

Because we have relevant data in our application from monitoring convergence, we can use the characteristic time to observe the relaxation of the order parameter autocorrelation function (see Fig. 7) as an estimate for $\tau_{\rm mol}$ for TS 1–4. For TS 5, we can use $\tau_{\rm mol} \approx k_B T/h$ because of the relatively small free energy barrier associated with this TS. The rates of transitions between the adjacent metastable states in Fig. 4 can then be calculated using Eq. **8**[§].

Using the individual rates of transitions between adjoining metastable basins, the overall rate of the conformational change can be determined by modeling the overall process as a network of reactions:

$$\text{Open (MS 1)} \underset{k_{21}}{\overset{k_{12}}{\rightleftarrows}} \text{MS 2} \underset{k_{32}}{\overset{k_{23}}{\rightleftarrows}} \text{MS 3} \underset{k_{43}}{\overset{k_{34}}{\rightleftarrows}} \text{MS 4} \underset{k_{54}}{\overset{k_{45}}{\rightleftarrows}} \text{MS 5} \underset{k_{65}}{\overset{k_{56}}{\rightleftarrows}} \text{MS 6} \underset{k_{76}}{\overset{k_{67}}{\rightleftarrows}} \text{Closed (MS 7)}, \qquad (\mathbf{9})$$

where MS 1–7 correspond to the different metastable states in Fig. 4. The network of reactions can be solved using kinetic Monte Carlo simulations [25] to determine the overall rate of transition between the open and closed states. This is a more involved process, which we may pursue in the future.

Still, an order-of-magnitude estimate for the rate (k) of the overall conformational change is available from

$$k \approx (k_B T/h) \times \exp(-\beta \Delta F_{\rm RL}), \qquad (\mathbf{10})$$

where $\Delta F_{\rm RL}$ is the free-energy of the rate-limiting barrier in the reaction profile relative to the open state. In our case, the rate-limiting barrier is TS 3, corresponding to Arg-258's

[§] A correction to the transition state theory approximation may be obtained by computing the transmission coefficient using the Bennett-Chandler method [24].

rotation, for which $\Delta F_{\mathrm{RL}} = 20.5 \pm 3k_BT$; the corresponding range for the rate is 1.5×10^5 to 4×10^2 s^{-1}.

Literature cited

1. Bolhuis, P. G., Chandler, D., Dellago, C. & Geissler, P. L. (2002) *Annu. Rev. Phys. Chem.* **53**, 291–318.
2. Bolhuis, P. G., Dellago, C. & Chandler, D. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 5883–5888.
3. Geissler, P. L., Dellago, C. & Chandler, D. (1999) *Phys. Chem. Chem. Phys.* **1**, 1317–1322.
4. Radhakrishnan, R. & Trout, B. L. (2003) *Phys. Rev. Lett.* **90**, 158301.
5. ten Wolde, P. R. & Chandler, D. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 6539–6543.
6. Bolhuis, P. G., Dellago, C. & Chandler, D. (1998) *Faraday Discuss.* **110**, 421–436.
7. Yang, L., Beard, W. A., Wilson, S. H., Broyde, S. & Schlick, T. (2002) *J. Mol. Biol.* **317**, 651–671.
8. Yang, L., Beard, W. A., Wilson, S. H., Roux, B., Broyde, S. & Schlick, T. (2002) *J. Mol. Biol.* **321**, 459–478.
9. Arora, K. & Schlick, T. (2003) *Chem. Phys. Lett.* **378**, 1–8.
10. Allen, M. P. & Tildesley, D. J. (1990) *Computer Simulation of Liquids* (Oxford Univ. Press, New York).
11. Crooks, G. E. (1999) Ph.D. Thesis, University of California (Berkeley).
12. Jarzynski, C. (1997) *Phys. Rev. Lett.* **78**, 2690–2693.
13. Chandler, D. (1987) *Introduction to Modern Statistical Mechanics* (Oxford Univ. Press, New York).
14. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comp. Chem.* **4**, 187–217.
15. Schlitter, J., Engels, M. & Krüger, P. (1994) *J. Mol. Graphics* **12**, 84–89.
16. Schlitter, J., Engels, M., Krüger, P., Jacoby, E. & Wollmer, A. (1993) *Mol. Simul.* **10**, 291–309.
17. Paci, E. & Karplus, M. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 6521–6526.
18. Bernèche, S. & Roux, B. (2001) *Nature* **414**, 73–77.
19. Bartels, C. & Karplus, M. (1998) *J. Phys. Chem. B* **102**, 865–880.
20. Harvey, S. C. & Gabb, H. A. (1993) *Biopolymers* **33**, 1167–1172.
21. Dellago, C., Bolhuis, P. G. & Geissler, P. L. (2002) *Adv. Chem. Phys* **123**, 1–81.
22. Lynden-Bell, R. M., Van Duijneveldt, J. S. & Frenkel, D. (1993) *Mol. Phys.* **80**, 801–814.

23. Frost, A. A. & Pearson, R. G. (1961) *Kinetics and Mechanism* (Wiley, New York).

24. Chandler, D. (1978) *J. Chem. Phys.* **68**, 2959–2970.

25. Fichthorn, K. A. & Weinberg, W. H. (1991) *J. Chem. Phys.* **95**, 1090–1096.

26. Abashkin, Y. G., Erickson, J. W. & Burt, S. K. (2001) *J. Phys. Chem. B* **105**, 287–292.

27. Rittenhouse, R. C., Apostoluk, W. K., Miller, J. H. & Straatsma, T. P. (2003) *Proteins: Struct. Funct. Genet.* **53**, 667–682.