THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

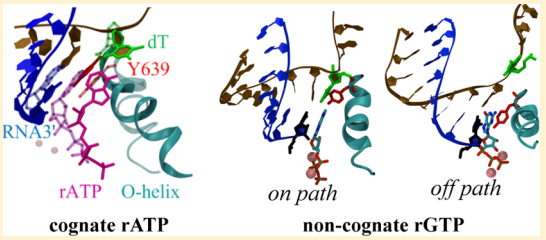
Nucleotide Selectivity at a Preinsertion Checkpoint of T7 RNA Polymerase Transcription Elongation

Chao E, Baogen Duan, and Jin Yu*

Beijing Computational Science Research Center, Beijing 100193, China

Supporting Information

ABSTRACT: Nucleotide selection is crucial for transcription fidelity control, in particular, for viral T7 RNA polymerase (RNAP) lack of proofreading activity. It has been recognized that multiple kinetic checkpoints exist prior to full nucleotide incorporation. In this work, we implemented intensive atomistic molecular dynamics (MD) simulations to quantify how strong the nucleotide selection is at the initial checkpoint of an elongation cycle of T7 RNAP. The incoming nucleotides bind into a preinsertion site where a critical tyrosine residue locates nearby to assist the nucleotide selection. We calculated the relative binding free energy between a noncognate nucleotide and a cognate one at a preinsertion configuration via alchemical simulations, showing that a small selection free energy or the binding free energy difference (~3 k_BT) exists between the two nucleotides. Indeed, another preinsertion configuration favored by the noncognate nucleotides was identified, which appears to be off path for further nucleotide insertion and additionally assists the nucleotide selection. By chemical master equation (CME) approach, we show that the small selection free energy at the preinsertion site along with the off-path noncognate nucleotide filtering can help substantially to reduce the error rate and to maintain the elongation rate high in the T7 RNAP transcription.



INTRODUCTION

Substrate selection is important for enzymatic reactions with association and/or catalytic specificities. In RNA polymerase (RNAP) transcription elongation, the selection of nucleotide substrate is based on the Watson−Crick (WC) base pairing between a template DNA nucleotide (nt) and the incoming RNA nucleoside triphosphate (or rNTP), and is further promoted by the RNAP acitivities. In viral RNAP species, such as that from bacteriophage T7, proofreading has not been detected, hence, the transcription accuracy or fidelity control relies fully on the nucleotide selectivity prior to the nucleotide incorporation.[1−3] Experimentally, multiple kinetic or conformational rearrangement steps have been detected preceding the catalysis of the RNA or DNA polymerase enzymes.[4−6] Accordingly, the nucleotide selection against the noncognate/wrong nucleotide is allowed to happen stepwise at each kinetic checkpoint,[4] i.e., either through an accelerated backward transition to the previous kinetic state, or via a deaccelerated forward transition toward the next kinetic state, in comparison to the cognate/right nucleotide.[7] The variations of the kinetic rates of the noncognate nucleotide relative to that of the cognate species are equivalent to the variations of binding free energies or activation barriers on the reaction paths for the different nucleotide species.

In a previous mathematical framework of the stepwise nucleotide selection of a nonproofreading polymerase, we characterized the selection strength $\exp(\Delta_i/k_BT)$ (with k_B the Boltzmann constant and T the temperature) at the i-th kinetic checkpoint by obtaining a selection free energy Δ_i that defines the difference of the binding free energies or activation barriers between the right and wrong nucleotide species. We then demonstrated that significant stepwise selection free energies Δ_i would accumulate along the reaction path to collectively impact on the elongation rate v (or speed) and error rate ε.[7] In particular, the error rate can scale approximately as $\exp(-\sum\Delta_i)$ with the accumulated or total selection free energy as $\sum\Delta_i$. In T7 RNAP with a transcription error rate measured at ~10^{-4},[3] $\sum\Delta_i$ is then estimated as ~10 k_BT.

The elongation cycle of T7 RNAP consists of five steps: the NTP binding/preinsertion, insertion/tight binding, catalysis, the pyrophosphate ion (PPi) product release, and the polymerase translocation (see Figure 1).[6,8,9] Correspondingly, the nucleotide selection can happen first by energetically destabilizing the noncognate NTP at the preinsertion site, so that there exists a positive relative binding free energy between the noncognate NTP and the cognate one (i.e., the initial selection free energy $\Delta_1 \equiv \Delta\Delta G_b$, or the relative binding free energy) and the unbinding rate of the noncognate NTP is enhanced by $\exp(\Delta_1/k_BT)$ upon that of the cognate NTP. Next, the selection proceeds through the nucleotide insertion, either by kinetically slowing down the insertion rate of the noncognate NTP comparing to that of the cognate one, or by destabilizing energetically the noncognate NTP at the binding/insertion site, so that the reversal rate of the noncognate NTP insertion is enhanced upon that of the cognate NTP. Finally,

Special Issue: Klaus Schulten Memorial Issue

Received: November 19, 2016
Revised: February 15, 2017
Published: February 15, 2017

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX

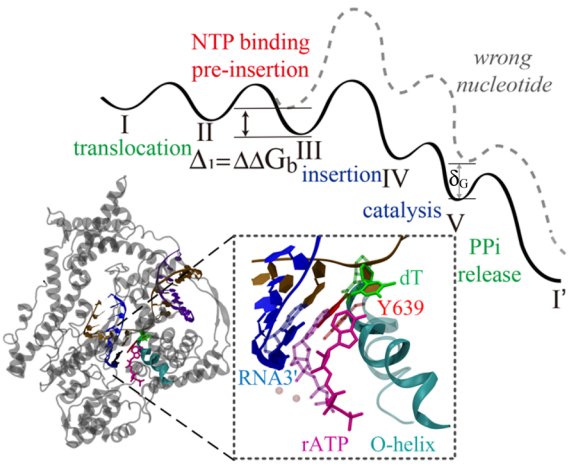


Figure 1. Elongation free energy profiles and structural views of T7 RNAP. Top: The free energy profiles of the cognate/right nucleotide (solid line) and the noncognate/wrong nucleotide (dashed line). Individual transitions including the polymerase translocation (state I → II), the nucleotide preinsertion (II → III), insertion (III → IV), catalysis (IV → V), and the PPi product release (V → I′) are indicated. Current study focuses on the nucleotide selection activity at the preinsertion checkpoint, i.e., calculating the selection or relative binding free energy between the noncognate and cognate NTP as $\Delta_1 = \Delta\Delta G_b$. The in-solution constraint of the free energy difference between the noncognate and cognate NTP at the 3′-end of the RNA is δ_G. A structural view (graphics produced by VMD[10]) of the T7 RNAP preinsertion complex bound with cognate rATP is shown (PDB: 1S0V),[8] with a zoomed in view of the preinsertion site (RNA, blue; template DNA, brown; the template nt for the incoming nucleotide to pair with, green; the O-helix, cyan; Tyr639, red; rATP, magenta; magnesium ions in pink). The active site of the insertion complex (rATP, template nt, Tyr639, and the O-helix) is also shown transparently for an easy comparison.

the selection can also go through the catalysis reaction, kinetically, as long as the cognate NTP catalyzes faster than the noncognate species. According to the respective free energy profiles for the cognate and noncognate NTP incorporation (Figure 1 top), we see that the free energy differences between the two profiles, or the selection free energies, accumulate gradually along the reaction path. Note that a constant or intrinsic free energy difference δ_G exits between the noncognate and the cognate NTP incorporation at the 3′-end of the synthesizing RNA strand as that in solution condition, which is an equilibrium constraint independent of the polymerase enzyme activity.

The existence of the preinsertion site in T7 RNAP and similar single-subunit RNA or DNA polymerases is a prominent structural feature discovered in previous high-resolution structural studies.[8,11] The preinsertion site locates next to the active site for the nucleotide incorporation. In our previous study of T7 RNAP, a critical tyrosine residue Tyr639 was recognized to be able to detect and grab on the noncognate NTPs at the preinsertion site.[12,13] It was found that the noncognate rNTP associates closely with Tyr639 but keeps far from the template nt, while the cognate rNTP forms WC base pairing with the template nt but stays away from Tyr639. On the other hand, for multisubunit RNAPs in higher organisms, a preinsertion elongation structure had also been identified.[14] Additionally, in multisubunit RNAPs, there exists an entry (E) site for the nucleotide binding prior to the binding into the preinsertion site.[15,16] It would actually be interesting to probe how the nucleotide selection proceeds at an early or preinsertion stage for both types of RNAPs. Nevertheless, detailed computational studies focused on the preinsertion nucleotide selection were still lack of, though substantial amount of work had been devoted to study overall conformational dynamics throughout the RNAP elongation cycle.[17−19] Current work would fill the gap then by implementing atomistic molecular dynamics (MD) simulations to the comparatively simple single-subunit T7 RNAP to provide structural and quantitative basis for the preinsertion nucleotide selection.

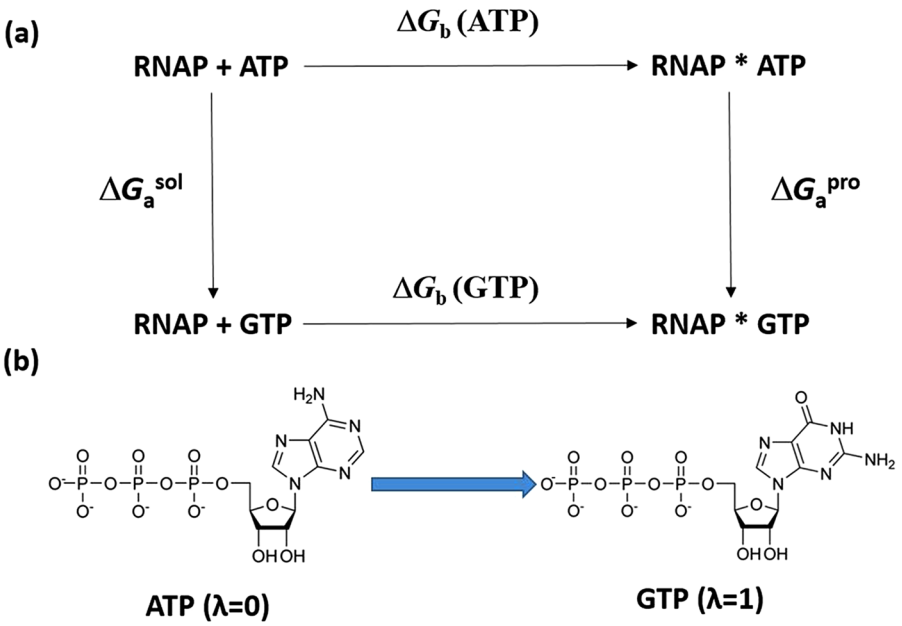


Figure 2. Computational scheme to calculate the relative binding free energy between the cognate ATP and noncognate GTP. (a) A thermodynamic cycle is used to obtain the relative binding free energy $\Delta\Delta G_b$. (b) The alchemical transition from ATP to GTP.

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX

The aim of this work is to essentially quantify how strong the nucleotide selection is at the initial checkpoint of T7 RNAP, i.e., to determine the nucleotide selection free energy Δ_1 as the relative binding free energy $\Delta\Delta G_b$ between the noncognate rNTP and the cognate one at the preinsertion site. In a traditional way of calculating the relative binding free energy between two types of ligands to a receptor protein, one performs an alchemical transition to allow gradual transformation of one type of ligand to the other type at the binding site.[20−22] An accurate calculation relies on whether the transformation is achieved sufficiently slow or quasi-statically. In current system, we have two types of the ligands, the cognate and noncognate rNTP, binding to the preinsertion site of the polymerase receptor protein. The nucleotide contains a negatively charged triphosphate group, while the base is comparatively flexible and can sample a variety of conformations in the absence of the base pairing. In addition, the cognate rNTP and the noncognate ones may also bind into different preinsertion configurations.[12] All these characters make the conformational relaxation slow and the sampling challenging. In this work, we conducted intensive MD simulations to the full T7 RNAP-DNA-RNA elongation complex for over four microseconds accumulatively to allow necessary local conformational relaxations during the alchemical transitions.

Furthermore, we analyzed our results in the kinetic framework of the RNAP elongation via the chemical master equation (CME) approach.[7,18] Through the CME approach one can see how stepwise selection free energies (Δ_i) at individual kinetic checkpoints impact on the overall elongation rate and error rate. In particular, the selection free energy Δ_1 was obtained from our submicrosecond atomistic simulations, while the CME links transitions between kinetic states in an elongation cycle that spans to tens of milliseconds.

Below we demonstrate first how we performed the MD simulations and the alchemical free energy calculations to obtain the relative binding free energy Δ_1. We then used the CME approach to obtain the elongation rate and error rate in the presence of the nucleotide selection. In particular, we show how much the initial nucleotide selection impacts on the elongation rate and error rate, i.e., the measures of speed and fidelity control, respectively. Last, we discuss how an off-path preinsertion configuration favored by the noncognate rNTP would additionally assist the speed-fidelity control.

METHODS

Calculating the Relative Binding Free Energy between the Noncognate and Cognate NTP. The relative binding free energy between ligands GTP and ATP at the preinsertion site can be evaluated through the thermodynamic cycle below (see Figure 2a): $\Delta\Delta G_b \equiv \Delta G_b$ (GTP) $- \Delta G_b$ (ATP) $= \Delta G_a^{pro} - \Delta G_a^{sol}$, where ΔG_a^{pro} and ΔG_a^{sol} are the free energies of transforming ATP alchemically into GTP in the protein complex and in the free solution, respectively.

The alchemical free energy was calculated using the free energy perturbation (FEP) method,[23] in which the free energy difference between the two reference states a and b is defined as

$$\Delta G(a \to b) = -\frac{1}{\beta}\mathrm{In}\langle \exp\{-\beta[H_b(x, p_x) - H_a(x, p_x)]\}\rangle_a \quad (1)$$

Here $\beta^{-1} \equiv k_B T$, with k_B the Boltzmann constant and T the temperature. $H_a(x,p_x)$ and $H_b(x,p_x)$ are the Hamiltonians at the states a and b, respectively. $\langle ... \rangle_a$ denotes the ensemble average over configurations at the initial state a. Both forward and backward alchemical transformations were performed in the simulation to evaluate the free energy change through the bidirectional sampling,[24] utilizing the Bennett acceptance ratio (BAR) method,[25] as implemented in the GROMACS package.[26,27] In this method, the Gibbs free energy from state i to j is given by

$$\Delta G(i \to j) = \beta^{-1}\left(\mathrm{In}\frac{\langle f(H_i - H_j + C)\rangle_j}{\langle f(H_j - H_i + C)\rangle_i}\right) + C \quad (2)$$

where f is the Fermi function $f(x) = [1 + \exp(\beta x)]^{-1}$, H_i and H_j are the Hamiltonians at the states i and j, respectively. The value for C is determined iteratively to satisfy $N_j \langle f(H_i - H_j + C)\rangle_j = N_i \langle f(H_j - H_i + C)\rangle_i$, and the free energy difference is then obtained as

$$\Delta G(i \to j) = -\beta^{-1}\mathrm{In}\frac{N_j}{N_i} + C \quad (3)$$

$$\Delta G(a \to b) = \sum_{i=1}^{n-1} \Delta G(i \to i+1) \quad (4)$$

where N_i and N_j represent the number of coordinate frames at λ_i and λ_j, respectively. The magnitude of the minimum square error σ^2 can be then found by taking the variance of eq 2 (see eqs 10(a), 10(b), and 11 of Bennett's paper).[25]

$$\sigma^2 = \frac{2}{\bar{N}}\left[\left(\int \frac{2\rho_i\rho_j}{\rho_i + \rho_j} dq^n\right)^{-1} - 1\right] \quad (5)$$

where $\bar{N}$ is some value lying between N_i and N_j,[25] ρ_i and ρ_j are the configuration space density functions, and q is the coordinate of the configuration space with n degrees of freedom. The integral in eq 5 indeed shows the "overlap" between the two configuration densities.[25]

Set Up of the MD Simulation. All MD simulations were performed using GROMACS-4.6.5 package.[26,27] The Amber99sb force field with ParmBSC0 nucleic acid parameters was used.[28−31] The ATP/GTP parameters were obtained from Carlson et al.[32]

The preinsertion complex of T7 RNAP was built from the crystal structure[8] (with PDB ID: 1S0V). The crystal waters within 10 Å of ATP molecule and two Mg^{2+} ions in the active site were retained. The RNAP complex was solvated with explicit TIP3P water[33] in a cubic box and the minimum distance from the protein to the wall was 10 Å. To neutralize the system and make the salt concentration 0.1 M, 115 Na^+ ions and 81 Cl^- ions were added. The full simulation system contained ~134 000 atoms. For all simulations, the cutoff of van der Waals (vdW) and the short-range electrostatic interactions was set to 9 and 10 Å, respectively. Particle mesh Ewald (PME) method[34,35] was used to evaluate the long-range electrostatic interactions. All the MD simulations were run at 1 bar and 310 K using the Parrinello−Rahman Barostat[36,37] and the velocity rescaling thermostat,[38] respectively. The LINCS algorithm was used to constrain all the chemical bonds.[39] The time-step was 2 fs and the neighbor list was updated every 5 steps. The solvated system was minimized with the steepest descent algorithm, followed by 100 ps of MD simulation within the canonical ensemble and 200 ps of simulation within the NPT ensemble. Position restraints on the heavy atoms of the

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX

protein and nucleic acids (NA) chains were imposed at the beginning of the simulation. Following the constrained simulation, unconstrained MD simulations were carried out for 100 ns.[12]

In the alchemical free energy simulation, Mg-ATP (GTP) was converted to Mg-GTP (ATP) in the forward (backward) direction (Figure 2b). To do that, we first added 3 dummy atoms to ATP. Then a topology file including ATP and GTP force field parameters was constructed by PYMACS package.[40] In the forward simulation, the mutation of ATP into GTP was controlled through the parameter λ from 0 to 1 with an increment of 0.05, and *vice versa* in the backward direction. The electrostatic and vdW interaction were simultaneously changed during the simulation. Twenty one simulation windows were included in each direction, and these simulations were conducted in the free solution and in the protein complex, respectively.

In the free-solution simulation, Mg-NTP was solvated in a cubic box with 4105 TIP3P waters, the minimum distance from Mg-NTP to the wall was 10 Å. To neutralize the system and keep the ionic concentration at 0.10 M, 10 Na^+ ions and 8 Cl^- ions were added. There were totally 12 380 atoms in the final system. The simulations in the protein complex were conducted as summarized early. For both the free solution and protein bound simulations, we followed the same procedure for each window: (i) 50 000 steps of energy minimization with a steepest descent algorithm; (ii) 100 ps of MD simulation within the canonical ensemble; (iii) 100 ps NPT equilibration with atomic position restrains on the heavy atoms of the protein and NA chains; (iv) the productive runs with the restraints removed and within the NPT ensemble, carried out for 100 and 5 ns in the protein complex and free solution, respectively. In total, 210 ns free-solution simulation data and 4.2 μs protein-bound simulation data were collected to evaluate the relative binding free energy between ATP and GTP.

Master Equation Approach To Solve the Five-State Elongation Kinetics. Here we consider a five-state polymerase elongation cycle (see Figure 1) including the pretranslocation state (I), the post-translocation state (II), the preinsertion state (III), the insertion state (IV), and the product state (V). Correspondingly, one defines a population vector $\Pi = (P_I\ P_{II}\ P^r_{III}\ P^w_{III}\ P^r_{IV}\ P^w_{IV}\ P^r_V\ P^w_V)^T$ to represent the probability distributions of the five states I to V (for both the right and wrong species from states III to V, labeled by r and w, respectively, for state III, IV, and V that can differentiate the right and wrong nucleotide species). The master equation for the distributions is

$$\frac{d}{dt}\Pi = M\Pi \tag{6}$$

where M is a 8 × 8 transition matrix as

$$\begin{pmatrix}
-k_{II+} & k_{II-} & 0 & 0 & 0 & 0 & k_{I+} & k_{I+} \\
k_{II+} & -k_{II-}-k_{III+} & k_{III-} & \eta_{III-}k_{III-} & 0 & 0 & 0 & 0 \\
0 & i_r k_{III+} & -k_{III-}-k_{IV+} & 0 & k_{IV-} & 0 & 0 & 0 \\
0 & (1-i_r)k_{III+} & 0 & -\eta_{III-}k_{III-}-\dfrac{k_{IV+}}{\eta_{IV+}} & 0 & \eta_{IV-}k_{IV-} & 0 & 0 \\
0 & 0 & k_{IV+} & 0 & -k_{IV-}-k_{V+} & 0 & k_{V-} & 0 \\
0 & 0 & 0 & \dfrac{k_{IV+}}{\eta_{IV+}} & 0 & -\eta_{IV-}k_{IV-}-\dfrac{k_{V+}}{\eta_{V+}} & 0 & \dfrac{k_{V-}\eta_G}{\eta_{III-}\eta_{IV+}\eta_{IV-}\eta_{V+}} \\
0 & 0 & 0 & 0 & k_{V+} & 0 & -k_{I+}-k_{V-} & 0 \\
0 & 0 & 0 & 0 & 0 & \dfrac{k_{V+}}{\eta_{V+}} & 0 & -k_{I+}-\dfrac{k_{V-}\eta_G}{\eta_{III-}\eta_{IV+}\eta_{IV-}\eta_{V+}}
\end{pmatrix}$$

with k_{II+} and k_{II-} the forward and backward translocation rate, k_{III+} and k_{III-} the NTP binding ($\propto$[NTP]) and unbinding rate, k_{IV+} and k_{IV-} the catalytic and its reverse rate, k_{V+} and k_{V-} the the catalytic and its reverse rate, and k_{I+} the product release rate. At a very low product (PPi) concentration, the last step (V → I′) is almost irreversible ($k_{I-} \to 0$). i_r is the portion of right/cognate nucleotides from solution at "input" (i_r = 25% by default for four equally mixed nucleotides). $\eta_G \equiv \exp(\delta_G)$ is a constant set at 10 by default (for δ_G about 2−3 k_BT). Note that δ_G is introduced early (see Introduction Section) as the in-solution free energy difference between the wrong/noncognate and right/cognate nucleotide incorporation to the 3′-end of the RNA strand.

The first selection strategy, denoted S_1, rejects wrong nucleotides by accelerated unbinding transitions from the preinsertion state III (at the selection strength $\eta_{III-} \equiv k^w_{III-}/k^r_{III-}$, with w and r for the wrong and right nucleotides, respectively). The second selection strategy, denoted S_2, inhibits the wrong nucleotide insertion by slowing down the insertion comparing to that of the cognate one (from III to IV, $\eta_{IV+} \equiv k^r_{IV+}/k^w_{IV+}$). The third selection strategy S_3 has an enhanced reversal rate for the wrong nucleotide insertion (from IV to III, $\eta_{IV-} \equiv k^w_{IV-}/k^r_{IV-}$). The fourth selection strategy S_4 inhibits the catalytic rate of the wrong nucleotide comparing to that of the right one $\eta_{V+} \equiv k^r_{V+}/k^w_{V+}$. We can then solve for the steady state solution of eq 6 to obtain the overall elongation rate v as $v \equiv (P^r_V + P^w_V)k_{I+}$ and the error rate as the elongation rate of the wrong NTPs over that of all NTPs (both right and wrong) $\varepsilon \equiv P^w_V/(P^r_V + P^w_V)$.

When there is an additional off-path preinsertion configuration or site (as addressed later in the text), we accordingly lowered i_r, the portion of right/cognate nucleotides at the first binding checkpoint.

RESULTS

We first show the alchemical free energy calculations between noncognate rGTP and the cognate rATP at the preinsertion site. Then we analyze the essential structural features in the alchemical transitions and compare them with that of a previously identified noncognate rNTP preinsertion configuration. Last, we show how much the initial nucleotide differentiation can impact on the elongation rate and error rate, respectively.

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX

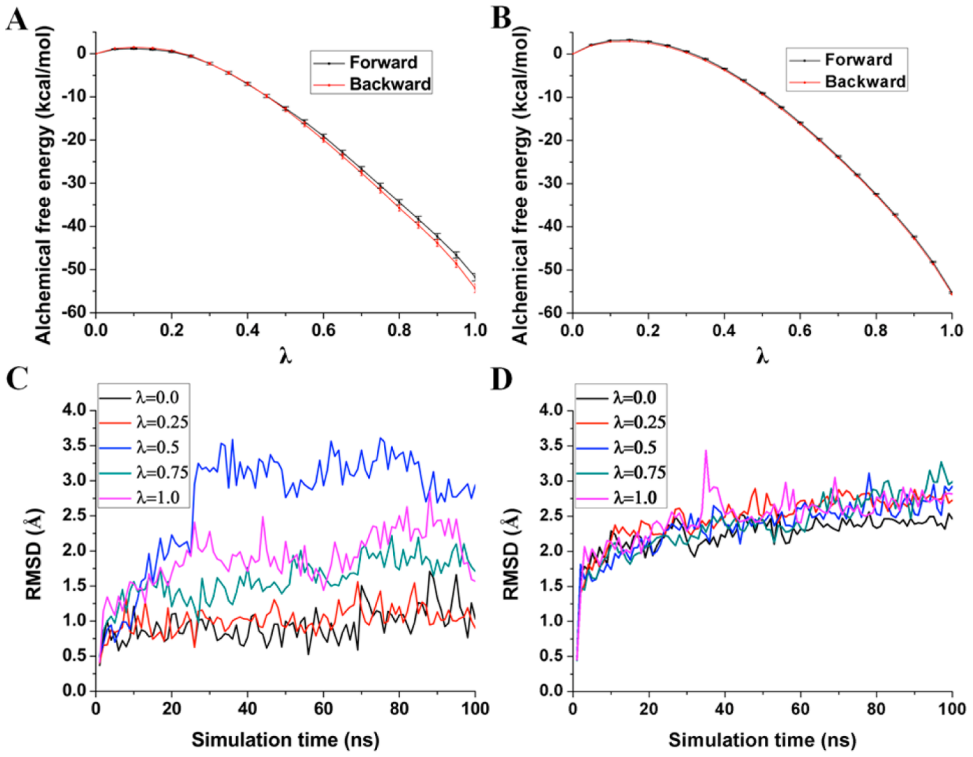


Figure 3. Alchemical free energy calculation and local conformation equilibration. The alchemical free energies for the rATP (λ = 0) to rGTP (λ = 1) transition forward and backward, conducted at the preinsertion site of the T7 RNAP protein complex and in the free solution are shown in (A) and (B), respectively. The RMSD values for local residues (heavy atoms of Y639, NTP, template dT, the 3′-end RNA base pair, and Cα atoms of the O-helix 627−640) around the preinsertion site and that for the full protein are shown in (C) and (D), respectively. The RMSD values were calculated according to structural alignments using the palm domain, and were monitored for five MD simulation windows (λ = 0, 0.25, 0.5, 0.75, 1) along the forward direction.

Relative Binding Free Energy Is ∼3 k_BT on Average between the Noncognate and Cognate rNTP at the Preinsertion. In the crystal structure of the preinsertion complex of T7 RNAP,[8] the DNA template nt is dT and the cognate nucleotide is rATP. Accordingly, rGTP was chosen as a noncognate rNTP here. Based on a thermal dynamical cycle of the two rNTP ligands binding to the receptor polymerase protein, the relative binding free energy between the two equals to their alchemical free energy changes between the protein and water environment (see Methods Section). Accordingly, we conducted MD simulations of the alchemical transitions between rATP and rGTP in the preinsertion site of T7 RNAP, and in water, respectively.

The alchemical transition was conducted in both forward (λ increase from 0 to 1) and backward (λ decreases from 1 to 0) directions. Twenty-one simulation windows were used for each direction. In the protein environment, 100 ns MD simulation was conducted in each window, and data between 20 to 100 ns were collected for the free energy integration. In the water environment, 5 ns simulation was conducted in each window. In both cases, the overall forward and backward free energy calculations agree relatively well (see Figure 3A and B), showing that the quasi-stability was kept reasonably during the alchemical transition. Nevertheless, there were still detectable free energy differences between the forward and backward calculations, so that we used the average values of the two. Accordingly, we obtained $\Delta G_a^{pro} \sim -53.6 \pm 0.3$ kcal/mol in the protein complex and $\Delta G_a^{sol} \sim -55.4 \pm 0.1$ kcal/mol in the water solution. As a result, the relative binding free energy between rGTP and rATP is $\Delta\Delta G_b = \Delta G_a^{pro} - \Delta G_a^{sol} \sim 1.8 \pm 0.2$ kcal/mol, or $\sim 3 \pm 0.4\ k_BT$. Note that the error estimation here along with the free energy calculation was conducted using the BAR method[25] (see eqs 2−5 in Methods Section).

We then checked the convergence in the free energy calculation. The convergence in water is fast, as bulk water relaxes at the picosecond time scale. In the protein environment, the convergence takes a much longer time. Anyway, the forward and backward alchemical free energy curves showed more or less convergence starting after ∼20 ns (see Supporting Information SI Figure S1). The deviations between the forward and backward free energies calculated in the protein complex would also lead to fluctuations of their average values, but the average fluctuation amplitudes were no more than ∼1.6 kcal/mol ($< \sim 3\ k_BT$) throughout the simulation. Hence, the average value of $\Delta\Delta G_b$ would not fluctuate significantly away from ∼3 k_BT, and $\Delta\Delta G_b > 0$ is also guaranteed.

From the RMSD values of the local residues around the NTP preinsertion site and that of the full protein complex, one can see that the binding site reached a local equilibrium after ∼20 to 30 ns (Figure 3C), while the full complex only gradually approached to the equilibrium after tens of nanoseconds yet kept evolving toward the end of the 100 ns simulation as λ increases (Figure 3D). Note that a jump of the local RMSD at λ = 0.5 in between 20 and 30 ns appeared as the rNTP molecule adjusted a dihedral angle between the base and the sugar during the transition, and the phosphate group deviated as well from the original position. In brief, the local conformational relaxation around the preinsertion site achieved fairly well during the alchemical transition, while the overall conformation relaxation of the protein complex had not necessarily yet.

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX

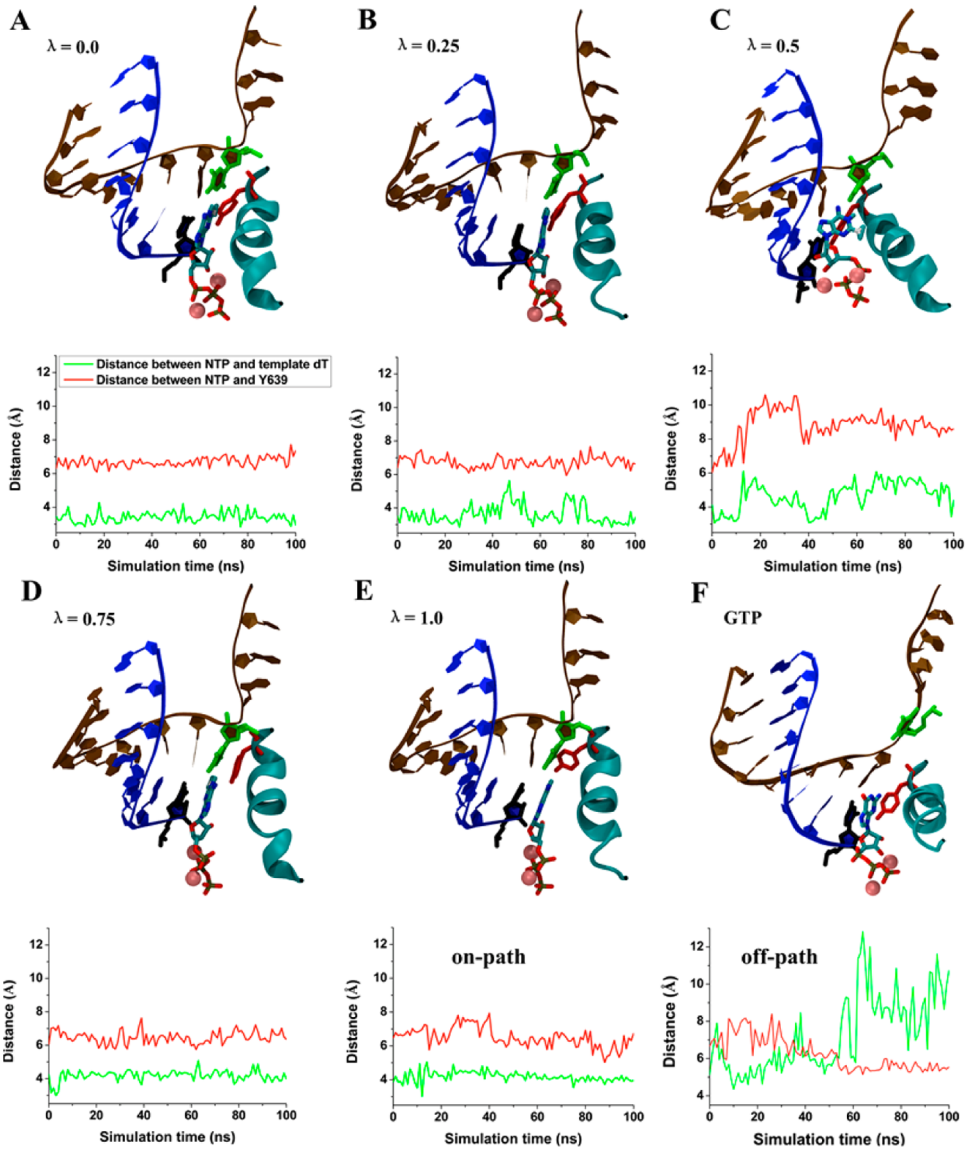


Figure 4. Preinsertion configurations during the alchemical transition from rATP to rGTP. The five structural snapshots along with measurements were taken from five simulation windows (A–E) for λ = 0, 0.25, 0.5, 0.75, and 1, respectively. The two measured distances were between the centers of the mass of the rNTP base and Y639 side ring (red line), and between the N1 atom of rNTP and the N3 atom of the template dT (the two involved in hydrogen bonding; green line). In particular, (E) is quite similar to (A) and is regarded as on path for further NTP insertion. The previously identified rGTP preinsertion configuration is shown in (F),[12] and it is suggested to be off path for further NTP insertion.

Noncognate Preinsertion Configuration Obtained Alchemically Here Is Structurally Similar to the Cognate Preinsertion Configuration. In our previous study,[12] we found that Tyr639 differentiates cognate rNTP from the noncognate ones at the preinsertion site by interacting closely with the noncognate species but weakly with the cognate one.[12] At the same time, the template nt stays closely with the cognate NTP to form the WC base pairing in the equilibrated preinsertion state, but deviates from the noncognate species. The study thus demonstrated notable structural variations and different positioning between the respective preinsertion configurations of the cognate and noncognate rNTPs.

To make a comparison, we also monitored how local configurations of the preinsertion NTP, Tyr639, and the template dT evolved as λ increased from 0 to 1. The configurations at five typical simulation windows $\lambda \sim$ 0, 0.25, 0.5, 0.75, and 1 are shown in Figure 4 (A–E). One can see that prior to the alchemical change at $\lambda \sim 0$ (Figure 4A) for the rATP preinsertion, the distance from rATP base to the Tyr639 side ring was comparatively large (~7 Å), while the WC base pairing or hydrogen bonding was formed well between rATP and the template dT. At the central transition intermediate $\lambda \sim$ 0.5, large fluctuations and variations of the rATP-Tyr639 distance were shown, while hydrogen bonding broke between the intermediate rNTP and the template dT. At the end of the transition $\lambda \sim 1$ (Figure 4E), the rGTP-Tyr639 distance dropped slightly to ~6 Å; however, the association distance between the converted rGTP and the template dT remained similar to that between rATP and dT all the way during the transition, even though the base pairing or hydrogen bonding was broken.

We note that the current alchemical transition started from an equilibrated structure of the cognate rATP preinsertion complex, in which the WC base pairing already formed. During

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX–XXX

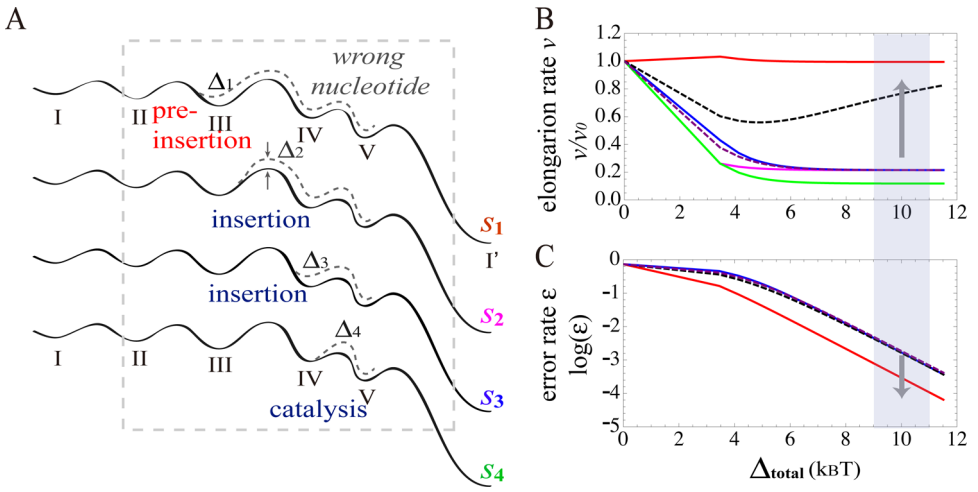


Figure 5. Impacts of the nucleotide selections on the elongation rate and error rate in the T7 RNAP transcription elongation. (A) A schematic diagram showing the four kinetic checkpoints or selection methods (S_1, S_2, S_3, and S_4) during the elongation.[7] The elongation rate v, normalized by v_0, the elongation rate without any selection, is shown in (B), and the error rate ε is shown in logarithmic value in (C), and both v/v_0 and $\log(\varepsilon)$ change with the total selection free energy $\Delta \equiv \sum\Delta_i$ ($i = 1\cdots, 4$). The red and magenta solid curves show the first two selections (S_1 and S_2), and the blue and green solid curves show the last two selections (S_3 and S_4). Note that the error rates are the same for S_1 and S_2, and for S_3 and S_4, so only S_1 and S_3 curves are visible in (C). The dark and purple dashed curves show two combined selection methods: (dark) with an equal selection free energy for every checkpoint $\Delta_1 = \Delta_2 = \Delta_3 = \Delta_4 = \Delta/4$, and a similar strategy can be used by T7 RNAP; (purple) no initial selection for $\Delta_2 = \Delta_3 = \Delta_4 = \Delta/3$. The region at $\Delta \sim 10\ k_BT$ is highlighted by a bar in (B) and (C). With an off-path binding site considered for the noncognate rNTP at preinsertion, the portion of the cognate nucleotide on-path increases effectively, so that the overall elongation rate increases and the error rate decreases for all the selection methods (indicated by the two gray arrows in B and C. See SI Figure S2 as well).

the alchemical change, with sufficiently well relaxation in the local region of the preinsertion site, the noncognate rGTP remained largely in the same preinsertion configuration as the cognate rATP. This preinsertion configuration thus appears quite differently from that obtained from our previous study (see Figure 4F), which was made by alchemically converting rATP to rGTP in the crystal structure of the rATP preinsertion complex. The WC base pairing between rATP and the template dT was not yet formed in the rATP bound crystal structure.[8] Accordingly, the converted rGTP molecule was able to drift to its favored preinsertion configuration immediately. It is interesting to notice that the template dT deviated far from rGTP in the previously identified preinsertion configuration, and dT indeed fluctuated around a translocation intermediate position in between the post-translocation and pretranslocation states of T7 RNAP.[9,41]

In brief, via the alchemical simulation, we obtained a preinsertion configuration of the noncognate rNTP, which is structurally similar to the equilibrated preinsertion configuration of the cognate rNTP. In this configuration, rNTP associates closely with template nt while stays away from Tyr639. On the other hand, in our previously identified preinsertion configuration for the noncognate rNTP, Tyr639 stays close by while the template nt is put to a distance away. We suggest that the previously identified preinsertion configuration is an off-path binding state, not suitable for further nucleotide insertion. In contrast, the currently obtained preinsertion configuration of the noncognate rNTP is similar enough to that of the cognate rNTP, and is an on-path binding state for further nucleotide insertion.

Impact of the Initial Nucleotide Selection on the Elongation Rate and the Error Rate. Previously we have studied stepwise nucleotide selection in a kinetic framework.[7] We analyzed the five-state model of the T7 RNAP elongation and identified four kinetic checkpoints for the nucleotide selection (see Figure 5A): At the first checkpoint (as selection S_1), the noncognate nucleotide is destabilized relative to the cognate one at the preinsertion site by a relative binding free energy Δ_1 (or the noncognate NTP with an enhanced unbinding rate by a factor of $\exp(\Delta_1)$ above that of the cognate unbinding rate) ; at the second checkpoint (as selection S_2), the insertion of the noncognate nucleotide is hindered by an elevated activation barrier of Δ_2 above that of the cognate one (or the noncognate NTP with a decreased insertion rate by $\exp(-\Delta_2)$ below that of the cognate insertion rate); at the third checkpoint (as selection S_3), the noncognate nucleotide bound to the insertion site, after leaking from the first two checkpoints, is less stabilized than the cognate one by a relative binding free energy Δ_3; and finally (as selection S_4), the noncognate nucleotide in the active or insertion site faces with an elevated activation barrier for catalysis above that of the cognate one, by Δ_4. As a good measure of the overall selection capacity of the RNAP,[7] the total selection free energy, the summation of the individual terms $\Delta \equiv \sum\Delta_i$ ($i = 1\cdots, 4$), is monitored here. Note that we estimated Δ to be $\sim 10\ k_BT$,[7] for the measured $\sim 10^{-4}$ error rate.[3]

In general, selection against the noncognate nucleotides by the RNAP would slow down the elongation rate or speed. It has been noticed that in an elongation cycle with translocation being fast, NTP concentration sufficiently high, and the rate-limiting step being the NTP insertion transition (III →IV), the elongation rate can be maintained high even with a weak initial selection S_1.[7] In Figure 5B, we show how the total selection free energy Δ affects the elongation rate v (normalized by v_0, the elongation rate without the RNAP selection), for individual selections from S_1 to S_4, respectively, and from two combined selection strategies as well. One sees that under S_1 individually, a very high elongation rate v is maintained. In the combined selection strategy without the initial selection (with an equal amount of selection free energy $\Delta/3$ applied through S_2 to S_4),

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX

the elongation rate drops to ~0.2 v_0 at $\Delta > \sim 6\ k_BT$. By including the initial selection in the other combined selection (with an equal amount of $\Delta/4$ applied through S_1 to S_4), the elongation rate significantly improves to ~0.8 v_0, e.g., for $\Delta \sim 10\ k_BT$. In particular, at $\Delta \sim 10\ k_BT$, one has $\Delta_1 \sim 2.5\ k_BT$, which is close to that obtained in our alchemical calculation above. A likely combined strategy with small but nontrivial initial nucleotide selection may actually work for T7 RNAP.

On the other hand, including only a small amount of initial selection free energy into the combined strategy cannot lower down the error rate significantly. A lowest error rate can only be achieved when almost all the selection free energy is devoted to the initial checkpoint, i.e., $\Delta \sim \Delta_1$ or as for the selection strategy S_1 (see Figure 5C), which is unlikely as a single initial checkpoint can hardly support the complete nucleotide selection.

However, we found in our simulation studies that the noncognate rNTP would adopt two different preinsertion configurations: One configuration similar to that of the cognate rNTP and is ready for nucleotide insertion, while the other configuration has the template nt located far away and appears to be an off-path binding configuration nonsuitable for nucleotide insertion. As long as the off-path preinsertion configuration is favored by the noncognate rNTP, it can be quite important for the fidelity control. At the high NTP concentration with sufficiently fast binding activities, if there is a slight free energy bias, say ~2 k_BT, toward the off-path preinsertion configuration, then ~90% of the noncognate rNTPs would bind into the off-path configuration. Accordingly, the portion of the cognate rNTPs recruited to the elongation cycle effectively increases, i.e., from 25% (with four equally populated rNTPs in solution) to ~80%. Accordingly, the elongation rate increases to some extent, while the error rate consistently drops for all the selection methods (see SI Figure S2). For example, for the combined selection with equally distributed selection free energy ($\Delta_i \sim \Delta/4$, i = 1···4) and $\Delta \sim 10\ k_BT$, the elongation rate improves from 0.8 v_0 to about v_0, while the error rate ε decreases from $\sim 10^{-3}$ to 10^{-4}, when the off-path noncognate rNTP preinsertion site is additionally considered.

DISCUSSION

Nucleotide selection is essential for transcription fidelity control, in particular, for nonproofreading polymerase, such as the single subunit viral T7 RNAP. We show that there are four kinetic checkpoints stepwise prior to nucleotide incorporation in a five-state T7 RNAP elongation cycle,[7] with an accumulated total selection free energy ~10 k_BT for the transcription error rate $\sim 10^{-4}$.[3]

In this work, we quantified the T7 RNAP selectivity at the nucleotide preinsertion, the first checkpoint of the RNAP elongation cycle. We implemented intensive all-atom MD simulations to calculate the relative binding free energy between a noncognate rGTP and a cognate rATP at the preinsertion site. The calculation was conducted via the alchemical simulations between rATP and rGTP, and the relative binding free energy turns out to be ~3 k_BT on average to select against the noncognate rGTP, right above thermal fluctuation level. Detailed inspections show that the binding configurations of the cognate rATP and noncognate rGTP appear similar in the alchemical simulations. In this configuration, the template nt associates closely with rNTP to potentially form base pairing interactions, while Tyr639 side chain keeps away from rNTP. We consider it an on-path preinsertion configuration ready for further nucleotide insertion.

However, in our previous MD study, it was found that rGTP at preinsertion could associate closely with Tyr639 but stay far from the template nt.[12] In that configuration, the noncognate rGTP is caught by Tyr639 so that it is unlikely to be further inserted. We thus propose it an off-path preinsertion configuration prohibited from nucleotide insertion. The appearance of the two preinsertion configurations for the noncognate rNTP can indeed bring advantages to the fidelity control, as illustrated below.

First we consider only the on-path preinsertion configuration, which is available to both the cognate rNTP and likely, a small portion of noncognate rNTPs. Further stepwise selection kinetics would proceed just as that shown in Figure 5A. Accordingly, we found that by including only a small amount of the selection free energy at the initial checkpoint, i.e., the preinsertion site in T7 RNAP, it would help to maintain the elongation rate or speed high, though the small amount of selection free energy does not contribute much to lower the error rate.

Interestingly, we recognize that there is an additional preinsertion configuration available to the noncognate rNTP, while the configuration seems off path or unable to proceed to the insertion step. Likely, this off-path preinsertion configuration is favored by the noncognate rNTP more than the on-path preinsertion configuration. Consequently, the off-path binding by the noncognate rNTPs at the preinsertion site would then filter out a large portion of the noncognate species, so that to prevent these nucleotides from being recruited to the elongation cycle. Effectively, it is equivalent to immediately reduce the error rate at the first kinetic checkpoint upon the NTP binding. Hence, the noncognate nucleotide filtering by this additional off-path preinsertion configuration would significantly improve the transcription fidelity control, without impacting negatively on the overall speed.

Notably, one sees that Tyr639 residue plays an important role in assisting differentiating the noncognate NTP and the cognate one,[12] so that the noncognate NTP filtering is made possible upon the preinsertion. Tyr639 is a highly conserved residue among single subunit RNA polymerases[42] and DNA polymerases.[43,44] However, there seems no corresponding residue playing a same role for nucleotide screening at preinsertion of multisubunit RNAP II.[14] Indeed, Thr831 residue located in the middle of the bridge helix of RNAP II points to the active site similarly as Tyr639 (see structural reviews and comparison in SI Figure S3), but it is lacking a side ring and is mainly to facilitate backtracking of RNAP II.[45] Interestingly, in the off-path binding configuration of noncognate rNTP, we noticed that the template nt drifted back from a post-translocation position to an intermediate position in between the post-translocation and pretranslocation states of T7 RNAP. It indicates that the noncognate rNTP does not likely serve as a paw as the cognate rNTP does in the Brownian ratchet model of RNAP.[13,46] Accordingly, the template nt was able to move back to resume the Brownian motions. In multisubunit RNAP, however, due to the presence of an E-site, the noncognate NTP binds to the preinsertion site can slip back by itself to the E-site via base rotation away from the template nt.[15] Hence, the initial screening or filtering of noncognate NTP in multisubunit RNAP appears to proceed quite differently from that in the single subunit T7 RNAP. A

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX

comparative investigation on the structural and energetic detail of the nucleotide selection in the multisubunit RNAP II preinsertion would be further expected.

CONCLUSIONS

In order to examine the nucleotide selectivity and quantify the selection strength at the initial checkpoint of T7 RNAP transcription elongation, we implemented intensive all-atom MD simulations and conducted alchemical free energy calculations. We obtained the initial selection free energy as the relative binding free energy between the noncognate and cognate rNTP at the preinsertion site as ~3 k_BT, which indicates that the noncognate NTP is less stabilized and more likely to unbind than the cognate one. In addition, we found that the currently obtained preinsertion configuration of the noncognate rNTP (on path) is notably different from the previously identified one (off path). We thus suggest that an on-path preinsertion instability along with the off-path filtering against the noncognate NTPs together promotes the fidelity control of T7 RNAP transcription elongation. Determining how much the off-path preinsertion configuration is energetically favored by the noncognate NTPs would help to reveal how much the overall elongation rate and error rate are modulated by the initial screening. Further studies on the nucleotide selectivity of this prototypical nonproofreading RNAP would focus on the selections during the insertion process, upon the insertion, and into the catalytic stage.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b11668.

Convergence in the alchemical free energy calculation, speed and fidelity control of T7 RNAP by an off-path filtering of the noncognate rNTPs at preinsertion, and structural comparison of the preinsertion configuration of T7 RNAP and that of yeast RNAP II (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: jinyu@csrc.ac.cn; Phone: +86 10 5698 1807.

ORCID

Jin Yu: 0000-0001-8224-1374

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is supported by National Science Foundation of China (NSFC) grant 11275022 and NSAF U153040. We acknowledge the computational support from the Beijing Computational Science Research Center (CSRC) and Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase). Thanks for comments from Dr Lin-Tai Da. All graphics were made with Visual Molecular Dynamics ("VMD" http://www.ks.uiuc.edu/Research/vmd/).

REFERENCES

(1) Sydow, J. F.; Cramer, P. RNA Polymerase Fidelity and Transcriptional Proofreading. *Curr. Opin. Struct. Biol.* **2009**, *19* (6), 732−739.

(2) Joyce, C. M.; Benkovic, S. J. DNA Polymerase Fidelity: Kinetics, Structure, and Checkpoints. *Biochemistry* **2004**, *43*, 14317−14324.

(3) Huang, J.; Brieba, L. G.; Sousa, R. Misincorporation by Wild-Type and Mutant T7 RNA Polymerases: Identification of Interactions That Reduce Misincorporation Rates by Stabilizing the Catalytically Incompetent Open Conformation. *Biochemistry* **2000**, *39* (38), 11571−11580.

(4) Yuzenkova, Y.; Bochkareva, A.; Tadigotla, V. R.; Roghanian, M.; Zorov, S.; Severinov, K.; Zenkin, N. Stepwise Mechanism for Transcription Fidelity. *BMC Biol.* **2010**, *8*, 54.

(5) Johnson, K. A. The Kinetic and Chemical Mechanism of High-fidelity DNA Polymerases. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 1041−1048.

(6) Anand, V. S.; Patel, S. S. Transient State Kinetics of Transcription Elongation by T7 RNA Polymerase. *J. Biol. Chem.* **2006**, *281*, 35677−35685.

(7) Yu, J. Efficient Fidelity Control by Stepwise Nucleotide Selection in Polymerase Elongation. *Mol. Based Math. Biol.* **2014**, *2*, 141−160.

(8) Temiakov, D.; Patlan, V.; Anikin, M.; McAllister, W. T.; Yokoyama, S.; Vassylyev, D. G. Structural Basis for Substrate Selection by T7 RNA Polymerase. *Cell* **2004**, *116* (3), 381−391.

(9) Yin, Y. W.; Steitz, T. A. The Structural Mechanism of Translocation and Helicase Activity in T7 RNA Polymerase. *Cell* **2004**, *116* (3), 393−404.

(10) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

(11) Joyce, C. M. Choosing the Right Sugar: How Polymerase Select a Nucleotide Substrate. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 1619−1622.

(12) Duan, B.; Wu, S.; Da, L.-T.; Yu, J. A Critical Residue Selectively Recruits Nucleotides for T7 RNA Polymerase Transcription Fidelity Control. *Biophys. J.* **2014**, *107* (9), 2130−2140.

(13) Yu, J.; Oster, G. A Small Post-Translocation Energy Bias Aids Nucleotide Selection in T7 RNA Polymerase Transcription. *Biophys. J.* **2012**, *102*, 532−541.

(14) Kettenberger, H.; Armache, K.-J.; Cramer, P. Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS. *Mol. Cell* **2004**, *16*, 955−965.

(15) Westover, K. D.; Bushnell, D. A.; Kornberg, R. D. Structural Basis of Transcription: Nucleotide Selection by Rotation in the RNA Polymerase II Active Center. *Cell* **2004**, *119*, 481−489.

(16) Batada, N. N.; Westover, K. D.; Bushnell, D. A.; Levitt, M.; Kornberg, R. D. Diffusion of Nucleoside Triphosphates and Role of the Entry Site to the RNA Polymerase II Active Center. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 17361−17364.

(17) Zhang, L.; Pardo-Avila, F.; Unarta, I.; Cheung, P.; Wang, G.; Huang, X.; Wang, D. Elucidation of the Dynamics of Transcription Elongation by RNA Polymerase II using Kinetic Network Models. *Acc. Chem. Res.* **2016**, *49* (4), 687−694.

(18) Yu, J.; Da, L.-T.; Huang, X. Constructing Kinetic Models to Elucidate Structural Dynamics of a Complete RNA Polymerase II Elongation Cycle. *Phys. Biol.* **2015**, *12*, 016004.

(19) Yu, J. Computational Investigations on Polymerase Actions in Gene Transcription and Replication: Combining Physical Modeling and Atomistic Simulations. *Chin. Phys. B* **2016**, *25*, 018706.

(20) Mobley, D. L.; Klimovich, P. V. Perspective: Alchemical Free Energy Calculations for Drug Discovery. *J. Chem. Phys.* **2012**, *137*, 230901.

(21) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.

(22) Straatsma, T.; McCammon, A. Computational Alchemy. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407−435.

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

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX

(24) Liu, P.; Dehez, F.; Cai, W.; Chipot, C. A Toolkit for the Analysis of Free-Energy Perturbation Calculations. *J. Chem. Theory Comput.* **2012**, *8* (8), 2606−2616.
(25) Bennett, C. H. Efficient Estimation of Free Energy Differences from MC data. *J. Comput. Phys.* **1976**, *22*, 245−268.
(26) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1), 43−56.
(27) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435−447.
(28) Guy, A. T.; Piggot, T. J.; Khalid, S. Single-stranded DNA within Nanopores: Conformational Dynamics and Implications for Sequencing; A Molecular Dynamics Simulation Study. *Biophys. J.* **2012**, *103* (5), 1028−1036.
(29) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65* (3), 712−725.
(30) Joung, I. S.; Cheatham, T. E., III Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112* (30), 9020−9041.
(31) Joung, I. S.; Cheatham, T. E., III Molecular Dynamics Simulations of the Dynamic and Energetic Properties of Alkali and Halide Ions Using Water-model-specific Ion Parameters. *J. Phys. Chem. B* **2009**, *113* (40), 13279−13290.
(32) Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of Polyphosphate Parameters for Use with the AMBER Force Field. *J. Comput. Chem.* **2003**, *24* (9), 1016−1025.
(33) Price, D. J.; Brooks, C. L., III A Modified TIP3P Water Potential for Simulation with Ewald Summation. *J. Chem. Phys.* **2004**, *121* (20), 10096−10103.
(34) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log (N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.
(35) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577−8593.
(36) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182−7190.
(37) Nosé, S.; Klein, M. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50* (5), 1055−1076.
(38) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.
(39) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463−1472.
(40) Seeliger, D.; De Groot, B. L. Protein Thermostability Calculations Using Alchemical Free Energy Simulations. *Biophys. J.* **2010**, *98* (10), 2309−2316.
(41) Yin, Y. W.; Steitz, T. A. Structural Basis for the Transition from Initiation to Elongation Transcription in T7 RNA Polymerase. *Science* **2002**, *298* (5597), 1387−1395.
(42) Cermakian, N.; Ikeda, T. M.; Miramontes, P.; Lang, B. F.; Gray, M. W.; Cedergren, R. On the Evolution of the Single-subunit RNA Polymerases. *J. Mol. Evol.* **1997**, *45* (6), 671−81.
(43) Delarue, M.; Poch, O.; Tordo, N.; Moras, D.; Argos, P. An Attempt to Unify the Structure of Polymerases. *Protein Eng., Des. Sel.* **1990**, *3*, 461−467.
(44) Iyer, L. M.; Abhiman, S.; Aravind, L. A New Family of Polymerases Related to Superfamily A DNA Polymerases and T7-like DNA-dependent RNA Polymerases. *Biol. Direct* **2008**, *3* (1), 39.
(45) Da, L.; Pardo-Avila, F.; Xu, L.; Silva, D.; Zhang, L.; Gao, X.; Wang, D.; Huang, X. Bridge Helix Bending Promotes RNA Polymerase II Backtracking Through a Critical and Conserved Threonine Residue. *Nat. Commun.* **2016**, *7*, 11244.
(46) Wang, H.; Oster, G. Ratchets, Power Strokes, and Molecular Motors. *Appl. Phys. A: Mater. Sci. Process.* **2002**, *75*, 315−323.

DOI: 10.1021/acs.jpcb.6b11668
J. Phys. Chem. B XXXX, XXX, XXX−XXX