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

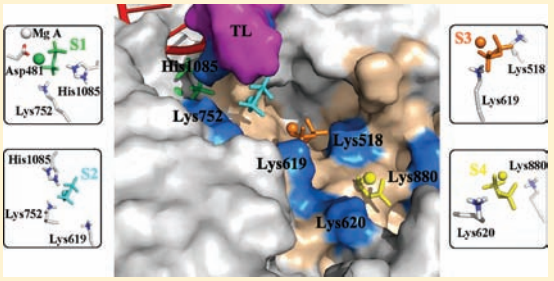
Dynamics of Pyrophosphate Ion Release and Its Coupled Trigger Loop Motion from Closed to Open State in RNA Polymerase II

Lin-Tai Da,[†] Dong Wang,*[,‡] and Xuhui Huang*[,†]

[†]Department of Chemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
[‡]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California 92093-0625, United States

Supporting Information

ABSTRACT: Pyrophosphate ion (PP_i) release after nucleotide incorporation is a necessary step for RNA polymerase II (pol II) to enter the next nucleotide addition cycle during transcription elongation. However, the role of pol II residues in PP_i release and the mechanistic relationship between PP_i release and the conformational change of the trigger loop remain unclear. In this study, we constructed a Markov state model (MSM) from extensive all-atom molecular dynamics (MD) simulations in the explicit solvent to simulate the PP_i release process along the pol II secondary channel. Our results show that the trigger loop has significantly larger intrinsic motion after catalysis and formation of PP_i, which in turn aids PP_i release mainly through the hydrogen bonding between the trigger loop residue H1085 and the $(Mg{-}PP_i)^{2-}$ group. Once PP_i leaves the active site, it adopts a hopping model through several highly conserved positively charged residues such as K752 and K619 to release from the pol II pore region of the secondary channel. These positive hopping sites form favorable interactions with PP_i and generate four kinetically metastable states as identified by our MSM. Furthermore, our single-mutant simulations suggest that H1085 and K752 aid PP_i exit from the active site after catalysis, whereas K619 facilitates its passage through the secondary channel. Finally, we suggest that PP_i release could help the opening motion of the trigger loop, even though PP_i release precedes full opening of the trigger loop due to faster PP_i dynamics. Our simulations provide predictions to guide future experimental tests.



INTRODUCTION

DNA-dependent RNA polymerases (RNAPs), central cellular machinery in the gene expression process, are responsible for the transcription of DNA into RNA in both prokaryotic and eukaryotic cells.[1,2] Multisubunit RNA polymerase II (pol II) in eukaryotic cells plays a unique role in the synthesis of mRNA (mRNA), which carries the genetic information from template DNA to protein.[3−5] In the past decade, an atomic-level view of the transcription mechanism of pol II[6−23] and its counterparts in bacteria[24−27] has been established.

The nucleotide addition cycle of pol II is a multistep process.[3−5,28] In the first step, a nucleotide triphosphate (NTP) is proposed to bind in the addition site (A site) via either an entry site (E site) in the secondary channel[10] or a template site in the downstream DNA channel.[29] Next, a highly conserved domain in pol II, named trigger loop (TL), undergoes dramatic structural transition from an open state to a closed state, which can position the incoming NTP and promote the subsequent catalytic reactions.[13,26,30] Substitutions in or deletion of TL in eukaryotic pol II[31] or prokaryotic RNAP[26,32,33] can substantially slow the elongation rate. A two-metal ion mechanism has been proposed to catalyze the formation of the phosphodiester bond.[34,35] In this mechanism, the first Mg^{2+} ion, termed as metal A, binds independently of NTP and facilitates the nucleophilic attack of the 3′-oxygen from the RNA terminal nucleotide on the α-phosphate of the substrate. The second Mg^{2+} ion, metal B, is likely to accompany the substrate NTP and product pyrophosphate (PP_i) during binding or release. Upon addition of incoming nucleotide to the 3′-RNA chain, PP_i is released and the trigger loop switches to the open state. Pol II then translocates one base pair downstream from the pre- to post-translocation register and is ready for a new cycle of nucleotide addition.[36,37]

Structural[13,26] and genetic studies[31,38−42] have shown that several key residues on the TL domain (residues Q1078 to T1095 from Rpb1) of pol II participate in regulating the substrate recognition and catalytic activity. In the recent crystal structure of pol II,[13] Q1078 is found to interact directly with the 3′−OH group of guanosine 5′-triphosphate (GTP). L1081 also makes direct contact with the nucleotide base, and mutation of the counterpart residue in *Escherichia coli* (M932A) can decrease the nucleotide addition rate.[41] Another highly conserved residue among different species, H1085, is found to interact directly with the phosphate group of NTP.[13,26] Substitutions of H1085 or its counterparts in both eukaryotic and prokaryotic RNAPs can significantly decrease the rate of the polymerization[31,41] as well as the reverse reaction,

Received: November 12, 2011
Published: December 28, 2011

dx.doi.org/10.1021/ja210656k | J. Am. Chem.Soc. 2012, 134, 2399−2406

pyrophosphorolysis.[42,43] These results suggest that this histidine may make stabilizing contacts in the transition state[33,41] or may be directly involved in the catalytic reaction as a proton donor.[44,45]

After formation of the phosphodiester bond, the product PP_i group is thought to release through the secondary channel (or pore) of the pol II[46] (see Figure 1). This channel consists of a narrow pore and a funnel-shaped region. The pore region is surrounded by several discontinuous parts of the Rpb1 and Rpb2 chains of pol II and also contains the E site.[10] In addition, several positively charged residues, such as K752, K619, K518, and K620, are located in the pore region. The relationship between the PP_i release and translocation is among the least well-understood steps in pol II nucleotide addition cycle. No structure of this state has been obtained. Previous studies[47] led to a suggestion that, for the single subunit T7 RNA polymerase, disassociation of the product PP_i can induce a protein conformational change that is directly coupled with translocation (called the power-stroke mechanism). In contrast, for multisubunit RNA polymerases, a Brownian ratchet model[48] has been proposed in which the DNA/RNA hybrid can oscillate between the pre- and post-translocation states driven by thermal fluctuations. For multisubunit RNAPs, both single molecular[49] and biochemical experiments[50] suggest that the PP_i release is uncoupled with the translocation process.

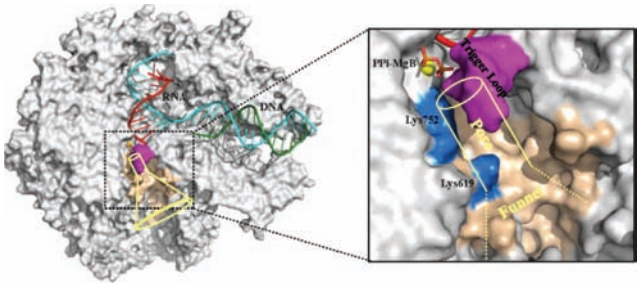


Figure 1. Pyrophosphate ion (PP_i) releases through the secondary channel of the pol II. This channel, colored in wheat, consists of a narrow pore region and a funnel-shaped region. Trigger loop (TL), RNA, template DNA, and nontemplate DNA are shown in magenta, red, cyan, and green, respectively. The PP_i is shown as orange sticks, and its bound $Mg^{2+}B$ ion is shown as a yellow sphere. Two critical lysine residues (K752 and K619) along the secondary channel are highlighted in blue.

Although experimental and theoretical studies[47−53] have provided a wealth of information for the relationship between the PP_i release and structural changes of the polymerase, the atomic-level mechanism of PP_i release remains a mystery, and no structure of a multisubunit RNAP with bound PP_i has been reported. Several important questions remain unanswered. How does PP_i release along the secondary channel? What is the role of pol II residues in modulating PP_i release? Is PP_i release coupled with the opening motion of the TL? To address these questions, we constructed a Markov state model (MSM) from 122 unbiased 6 ns all-atom molecular dynamics (MD) simulations in the explicit solvent with the system size of ∼370 000 atoms. The aggregated simulation time is more than 700 ns. Our results indicate that a few positively charged residues along the secondary channel can facilitate the process of the PP_i release by forming "hopping" sites. Within each hopping site, PP_i can form favorable interactions with pol II and greatly increase its residence time. Furthermore, we observed a strong coupling between the process of PP_i release and the opening motion of the TL.

METHODS

1. Setup of the Pol II Elongation Complex. The structure of pol II was modeled from the crystal structure of GTP-bound pol II complex (PDB ID, 2E2H). Missing parts of pol II were added by SWISS pdb software.[54] More specifically, missing residues 156−160, 186−198, 315−318, 1177−1186, 1232−1235, and 1235−1253 in Rpb1, residues 249−250, 336−344, 438−445, 503−508, 669−677, 715−721, 733−734, and 920−934 in Rpb2, residues 1−2, 49−51, 83−94, and 112−116 in Rpb5, and residues 64−75 in Rpb8 were rebuilt. For the longer missing residues, 71−89 and 133−163 in Rpb2 located on the surface of the protein, three residues GAG were applied to replace these two motifs. Other parts in the complex, including the downstream DNA, DNA−RNA hybrid, two Mg^{2+} ions, eight Zn^{2+} ions, and one GTP, were retained.

1.1. GTP-Bound Complex. We have made a small modification on the GTP-bound pol II complex structure (PDB ID, 2E2H) because the O3′ atom of the primer terminal RNA chain is missing in this crystal structure. This leads to the improper coordination modes between the $Mg^{2+}A$ and the oxygen atoms bonded to P_α of the GTP. To add the missing oxygen atom and fix the coordination modes of the $Mg^{2+}A$, we used the pretranslocation state structure[7] (PDB ID, 1I6H) as the template to refine the structure of 2E2H for the RNA nucleotide at the *i* position, the base, sugar, and monophosphate of GTP, and the metal $Mg^{2+}A$. After energy minimization, our new model exhibits good binding modes between the GTP and the active site atoms.

1.2. PP_i-Bound Complex. We modeled the PP_i-bound pol II complex by breaking the P_α−O bond of GTP in the crystal structure of pol II (PDB ID, 2E2H) into a PP_i molecule and a newly formed nucleotide. For the residue H1085, both protonated and neutral states are considered. Then the two structures were applied to the minimization. Comparison of the minimized structure of PP_i-bound pol II complex to the crystal structure (2E2H) indicate the minor side chain fluctuations in the active site, suggesting our model is a reasonable starting point (see Supporting Information Figure S1A).

The parameters of the protein residues, DNA, RNA, and ions were taken from the AMBER03 force field.[55] The parameters of GTP were taken from a previous study.[56] Since the binding of the Mg^{2+} to the PP_i may induce significant charge transfer, we have regenerated the partial charges of the $(Mg{-}PP_i)^{2-}$ group and its protonated form $(Mg{-}PP_iH)^{-}$ by RESP fitting to the quantum calculations (HF/6-31G*) using RED III software[57] (see Supporting Information Figure S2).

2. Simulation Details. *2.1. Molecular Dynamics Simulations.* All MD simulations were performed with the GROMACS 4.5.[58] Each protein system was solvated in a cubic box of SPC water molecules[59] with the minimum solute−wall box distance of 7.0 Å. A total of 76 Na^+ ions were added to neutralize the system. The final PP_i-bound pol II complex consists of 372 059 atoms. Long-range electrostatic interactions were treated with the particle-mesh Ewald (PME) summation method.[60,61] van der Waals and short-range electrostatic interactions were cut off at 10 Å. The simulations were run at 310 K and 1 bar using the velocity rescaling thermostat and Berendsen barostat, respectively.[62] All chemical bonds were constrained using LINCS algorithm.[63] We have used a 2 fs time step and updated the neighbor list every 10 steps. The solvated system was minimized with the steepest decent minimization method followed by a 120 ps simulation with the positions of the heavy atoms restrained.

2.2. Steered Molecular Dynamics Simulations. Steered MD simulations[64] were employed to generate the initial release pathways by pulling the PP_i group to three different directions (see Supporting Information Figure S3). One direction is along the wall of the secondary channel choosing the center of mass of the C_α atoms of Rpb1 residues 880−882 and 953−955 as the reference (ref I). The other two directions are along the central region of the channel pointing to the following two points: (1) the middle point of the center of mass of the Rpb1 residues Asp716-Arg720 and the ref I and (2) the middle point of the center of mass of Rpb1 residues Arg726-

Arg731 and the ref I (two gray lines in Supporting Information Figure S3). The external force was applied on the center of mass of the PP_i group with the force constant of 0.5 kJ mol^{-1} $Å^{-2}$ and pulling rate of 0.01 Å/ps. For each pulling direction, six independent steered MD simulations were performed starting from different random structures taken from the MD simulations of the PP_i-bound complex.

Next, we selected the initial conformations from the above steered MD simulations for seeding subsequent unbiased MD simulations. We have clustered the conformations from steered MD simulations into 20 groups using *K*-center clustering algorithm.[65] In our clustering method, the geometric distance between a pair of conformations was described by the root-mean-square deviation (rmsd) value of three PP_i atoms (the bridge oxygen and two phosphate atoms), and the pol II conformation was aligned by the bridge helix domain before each rmsd calculation. We then randomly selected roughly the same number of conformations from each cluster (with a total of 122 conformations) to seed unbiased MD simulations.

3. Constructing the Markov State Model. MSM is a powerful tool for accessing long time scale dynamics with many short simulations at local equilibrium.[66−69] MSM partitions phase space into a number of metastable states, such that intrastate transitions are fast but interstate transitions are slow. This separation of time scales ensures an MSM is Markovian (i.e., that the probability of transitioning from state *i* to state *j* depends only on the identity of *i* and not any previously visited state) and allows MSM built from many short simulations. These dynamics can then be propagated to give global long time scale dynamics:

$$\mathbf{P}(n\Delta t) = [\mathbf{T}(\Delta t)]^n \mathbf{P}(0) \quad (1)$$

where $\mathbf{P}(n\Delta t)$ is a vector of state populations at time $n\Delta t$ and $\mathbf{T}$ is the transition probability matrix.

In order to construct MSM, we first divide 732 000 conformations from the above seeding simulations (122 trajectories, each contains 6000 snapshots) into 274 microstates using the *K*-center clustering algorithm.[65] Next we lumped kinetically related microstates into four macrostates using the robust Perron cluster cluster analysis (PCCA+) algorithm[70] implemented in the MSMBuilder software.[65]

We check if the model is Markovian by examining the behaviors of the implied time scales.[71] Implied time scales (τ_k) can be calculated as

$$\tau_k = -\frac{\tau}{\ln \mu_k(\tau)} \quad (2)$$

where μ_k is an eigenvalue of the transition matrix with the lag time τ. Each implied time scale describes an aggregate transition between two subsets of states. If the model is Markovian, the exponentiation of $\mathbf{T}$ should be equivalent to an MSM constructed with a longer lag time according to eq 1. Under this condition, the implied time scales should be independent of the lag time. The implied time scale plots for both the 274-microstates and four-macrostates MSM reach the plateau at a lag time from 3 to 4 ns, indicating the model is Markovian (see Supporting Information Figure S4). Thus, we chose a lag time of 3.5 ns to build our model.

The detailed procedure of constructing the MSM is as follows:

3.1. Splitting the Conformations into Microstates. We first used the *k*-centers algorithm implemented in MSMBuilder software[65] to cluster all the conformations into 300 microstates. We hope that conformations within each microstate are all geometrically similar so that they are also kinetically similar. Indeed, the average distance between conformations of all microstates to their centers is only ∼1 Å. Here the distance is defined as the rmsd of three atoms of the PP_i group (the bridge oxygen atom and two phosphate atoms) after aligning the system based on the bridge helix. We then leave out the disconnected or almost disconnected microstates (<30 counts to other states with 1 ps of lag time), which may be due to insufficient sampling or noise. This drops 26 microstates only containing a total of 0.01% of all the conformations. Next, we plotted the implied time scales for the 274-microstate MSM and found that they level off at a lag time from 3 to 4 ns, implying that the model is Markovian (see Supporting Information Figure S4A). Therefore, we further confirm that the microstates are sufficiently small to guarantee that conformations in the same state are kinetically similar.

3.1.1. Calculating the Transition Probability Matrix. To calculate a transition probability matrix, we first counted the number of transitions between each pair of states at some interval or lag time to generate a transition count matrix ($\mathbf{N}$). N_{ij} denoted to the number of transitions observed from state *i* to state *j*. In this study, we adopted a sliding window of the lag time on each 6 ns trajectory with an interval of 1 ps between stored conformations to count the transitions. Since the system may quickly go back and forth at the energy barriers, a hard boundary between states may lead to an overestimation of the number of transitions. We thus adopt a similar procedure as described in Silva et al.[66] to eliminate the recrossing events by only counting transitions from state *i* to state *j* if the system stayed in state *j* for at least 300 ps after the transition. Finally, the transition probability matrix was generated by normalizing rows of the transition count matrix.

3.2. Lumping Microstate into Macrostate. For this system, the implied time scale plots for the 274-microstate MSM display a clear gap in between the third and fourth longest time scales, indicating that there exists four macrostates. We then applied the PCCA+ algorithm[70,72] implemented in MSMBuilder[65] to lump all the microstates into four macrostates. The implied time scales were plotted again for the four-macrostate MSM (see Supporting Information Figure S4B). The results are consistent with the microstate MSMs that all the implied time scale curves level off when the lag time is in between 3 and 4 ns, indicating the model is Markovian. Therefore we select the lag time of 3.5 ns in our further calculation of equilibrium state populations and kinetic properties.

3.3. Calculation of the Mean First Passage Time. The mean first passage time (MFPT) is defined as the average time taken to go from the initial to the final state, as described in ref 73. The MFPT from state *i* to state *f* ($MFPT_{if}$) is determined as

$$MFPT_{if} = \sum_j P_{ij} \times (t_{ij} + MFPT_{jf})$$

where P_{ij} is the transition probability from state *i* to state *j*, t_{ij} is the lag time used to construct the transition probability matrix T_{ij}, and $MFPT_{jf}$ is the mean first passage time of the state *j* to final state *f*. For each transition, a set of linear equations can be solved under the boundary condition that $MFPT_{ff} = 0$.

RESULTS AND DISCUSSION

Formation of the PP_i Destabilizes the Closed Conformation of the Trigger Loop. We first modeled the PP_i-bound pol II complex by converting the substrate-bound form of the pol II complex (the GTP-bound and prechemistry complex, PDB ID, 2E2H) into a product-bound complex (PP_i-bound and postchemistry complex) (see Methods). In our model, the TL residue H1085 is protonated and the PP_i group is unprotonated (see Supporting Information Figure S5A and Supporting Information text for discussions of the protonation states). The positions of active-site residues in our modeled structure after energy minimization are nearly identical to those in the crystal structure (rmsd < 0.8 Å, see Supporting Information Figure S1A), indicating our model is a reasonable starting point. In addition, we found that the fluctuations of the TL-tip residues (Rpb1 residues Gln1078 to Ala1087, root-mean-square fluctuation (RMSF) = 0.78 Å) are smaller than those of the whole Rpb1 subunit (RMSF = 1.63 Å). These results are consistent with the B-factor values in the X-ray structure (PDB ID, 2E2H[13]), where the average B-factor of the TL-tip residues (∼115.15) is lower than that of the Rpb1 subunit (∼133.16). Next, we compared the dynamics of the system before (GTP-bound state) and right after chemistry (PP_i-bound state) by performing MD simulations. The results

show that the TL-tip residues undergo much larger fluctuations at the nanosecond time scale in the PP_i-bound complex than in the GTP-bound complex (see Figure 2A).

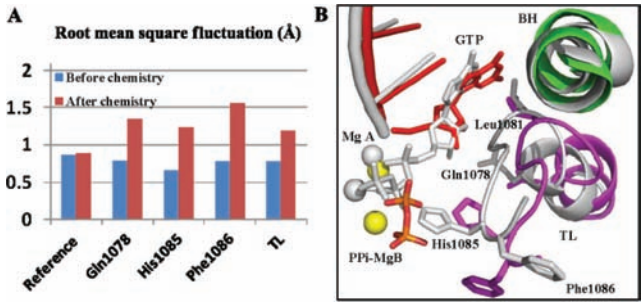


Figure 2. (A) Root-mean-square fluctuations (RMSFs) for the GTP-bound (before chemistry) and PP_i-bound (after chemistry) states of the pol II complex. The reference system includes heavy atoms within 25 Å of the $Mg^{2+}A$ in the active site. The TL system contains the tip of the trigger loop (from Rpb1 Gln1078 to Ala1087). The RMSF values are obtained from two and five independent simulations for the GTP-bound and PP_i-bound system, respectively. (B) A representative conformation of the PP_i-bound state from MD simulations (the same color code as in Figure 1 is used, and the bridge helix is shown in green) is superimposed with the GTP-bound state from the X-ray structure (PDB ID, 2E2H, gray). RMSF values are averaged over different residues, and a single residue RMSF value is defined as RMSF $= [(1/T)\sum_{t_0}^{T}(x_i(t) - \bar{x}_i)^2]^{1/2}$, where $\bar{x}_i$ is the time-averaged position of the C_α atom of the residue i.

These large fluctuations in the PP_i-bound complex lead to the breaking of various contacts between the TL and 3′-end of the RNA chain. For example, the interaction between the critical TL residue Q1078 (at which the single substitutions Q1078N or Q1078S are lethal[39]) with 3′−OH of the substrate is lost (see Figure 2B). In addition to the TL residues, we also found that other critical interactions between the pol II and 3′-end of RNA chain are lost after the formation of PP_i. For example, the critical residue Rpb1 N479 (its substitution can greatly decrease the elongation rate[39]) breaks its interaction with the GTP 2′−OH. Other TL residues such as H1085 and F1086 also become more flexible as reflected by significantly larger RMSF values after chemistry (see Figure 2A). These lost or weakened interactions may destabilize the interaction network between the 3′-end of the RNA chain and pol II, especially the tip region of the TL. Indeed, at the time scale of several nanoseconds, TL-tip residues already exhibit significantly larger motions than most of the enzyme on average (see Figure 2A). The newly formed PP_i also establishes new contacts with pol II residues such as Rpb1 K752. In spite of these newly formed or broken contacts, PP_i preserves most of its interactions with pol II in our nanosecond MD simulations, especially with positively charged residues such as Rpb2 R766 and R1020 (see Supporting Information Figure S1). These salt bridges and other interactions can still hold the PP_i in the active site, and thus we did not observe any event in which the PP_i escapes the active site in our nanosecond MD simulations.

PP_i Release through the Secondary Channel Pore Region Occurs via Hopping. As discussed in the previous section, nanosecond MD simulations are not able to capture the PP_i release due to their insufficient length. In order to overcome this time scale gap, we have applied MSM to model the PP_i release at the microsecond time scale from shorter nanosecond unbiased MD simulations. MSM has recently shown to be a powerful tool to model long time scale dynamics from many short MD simulations in protein folding and other conformational changes by us[65,66,74,75] and other groups.[67−69,76,77] In specific, we performed steered MD simulations to generate the initial PP_i release pathways using an external force to pull the PP_i away from the active site. Next, in order to eliminate the bias of the external force on the release dynamics, we used conformations from these initial pathways as starting points for 122 unbiased 6 ns MD simulations. Finally, data from these short MD simulations were used to build an MSM, which can identify the metastable states and capture the long time scale dynamics of PP_i release (see Supporting Information Figure S6 and the Methods section for details).

We performed 18 steered MD simulations to pull the PP_i out of the active site along the secondary channel, which consists of a pore region and a funnel-shaped region as shown in Figure 1. The pore region, surrounded by Rpb1 and Rpb2 domains including the closed TL, is so narrow that it barely allows the PP_i to pass through. Because computational simulation of PP_i passage through the entire length of the secondary channel is not currently feasible, we focused on PP_i release through the pore region, which is likely to be the rate-limiting step for the entire PP_i release process. In order to generate multiple alternative initial release pathways, we altered the pulling direction. In particular, we pulled PP_i along three directions from the active site, through the pore, and into the funnel region: one along the wall, and the other two along the central region of the funnel (see Supporting Information Figure S3 for pulling directions). In all the simulations, the $Mg^{2+}B$ ion is released together with the PP_i, even though the pulling force was only applied to PP_i. The results also show that the release of the PP_i can induce significant conformational changes in the closed state trigger loop and facilitate its transition to the open state. Thus, even if the external forces in these steered MD simulations alter the energy landscape, they still provide reasonable initial release pathways for our following studies.

We constructed a four-state MSM from a total of 122 6 ns unbiased MD simulations derived from the initial release pathways (see the Methods section for details of model construction). On the basis of the results, we propose a hopping model for PP_i release, in which the $(Mg{-}PP_i)^{2-}$ ion jumps between three intermediate sites that are composed of positively charged residues. First, we found two positively charged residues (Rpb1 K752 and TL H1085) that clamp the $(Mg{-}PP_i)^{2-}$ ion and facilitate its exit from the active site to the first intermediate state (from state S1 to S2 in Figure 3B). Next, the $(Mg{-}PP_i)^{2-}$ ion transits from the first to the second intermediate state (from state S2 to S3 in Figure 3B). During this process, PP_i breaks the interactions with K752 and H1085 and forms interactions with another two positively charged residues: Rpb1 K619 and K518. Next, the $(Mg{-}PP_i)^{2-}$ ion can transfer to the fourth metastable state in our simulations in which K620 and K880 form interactions with the $(Mg{-}PP_i)^{2-}$ ion (state S4 in Figure 3B). The relative equilibrium populations among states S1−S4 calculated from the MSM are 24.7%, 16.7%, 33.4%, and 25.5%, respectively. PP_i is then expected to diffuse into the funnel region of the secondary channel and finally into the solvent. In our simulations of PP_i release, we did not observe pol II translocation.

On the basis of the above analysis, the four positively charged residues K752, 619, 620, and 518 are likely to play important roles in the PP_i release. Although all four residues are well-conserved among eukaryotes (from yeast to human, Figure 3C

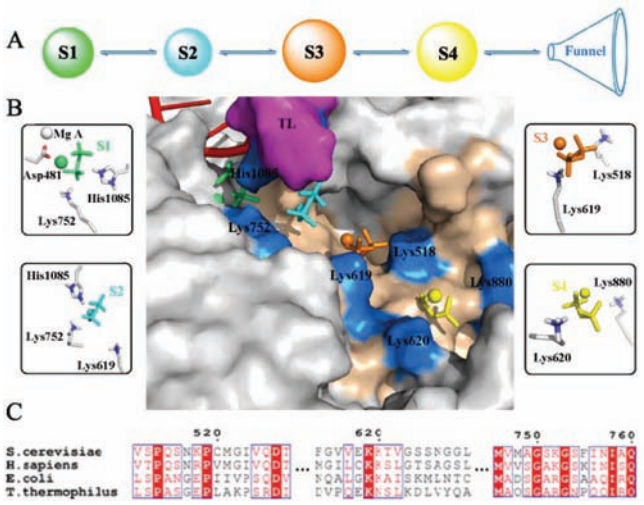


Figure 3. PP_i release pathway through the pore region of the secondary channel identified by Markov state model (MSM). (A) Four metastable states (S1−S4) on the releasing pathway are displayed in circles. The size of these circles is proportional to their equilibrium populations obtained from MSM. (B) The locations of the four states in the channel are displayed with the same color code as in part A. Key interactions between PP_i and the pol II in each state are also shown in part B. Several positive residues including K518, K619, K620, and K752 are found to play critical roles in facilitating the PP_i release. (C) Multiple sequence alignment of these positively charged residues among different species. The sequence alignment was performed using the online software ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

and Supporting Information Figure S8), only K619 and K752 are conserved among prokaryotes. K620 is altered to asparagine (N) in *Thermus thermophilus* (Figure 3C), and K518 is altered to the glutamate (E) in both *T. thermophilus* and *E. coli* (Figure 3C). These results suggest that, during the PP_i release, K619 and K752 may play common roles in different species, whereas K518 in eukaryotic pol II may have different functions compared with its counterpart residue in bacterial RNAP.

K619 and K752 are located at the previously described nonspecific binding site that can be occupied by NTPs (so-called "E site").[10] Previous work has predicted that the E site can enhance the rate of NTP diffusion into the pore region of the secondary channel because of the electrostatic attractions between the negatively charged NTP and positively charged E site.[46] For PP_i release, we hypothesize that these positively charged residues in the E site may play a similar role to promote the release of the negatively charged $(Mg{-}PP_i)^{2-}$ ion from the active site. In addition to K619 and K752, the protonated H1085 also contributes to this electrostatic interaction with the $(Mg{-}PP_i)^{2-}$ in our models (see Figure 3B).

Simulations of Pol II Mutants Reveal Individual Roles of Key Residues in Facilitating PP_i Release. On the basis of our MSM constructed from wild-type (WT) MD simulations, three important residues, H1085, K752, and K619, were identified to be involved in the interactions with PP_i during its release from the pore region of the secondary channel. To elucidate the specific roles of these three residues in different stages of PP_i release, we next performed simulations with pol II containing single substitutions.

We designed and simulated several single pol II mutants to examine PP_i passage among the different metastable states (see Figure 4A−C). In particular, we focused on simulations of PP_i release for H1085F and K752A (for S1 → S2 and S2 → S3

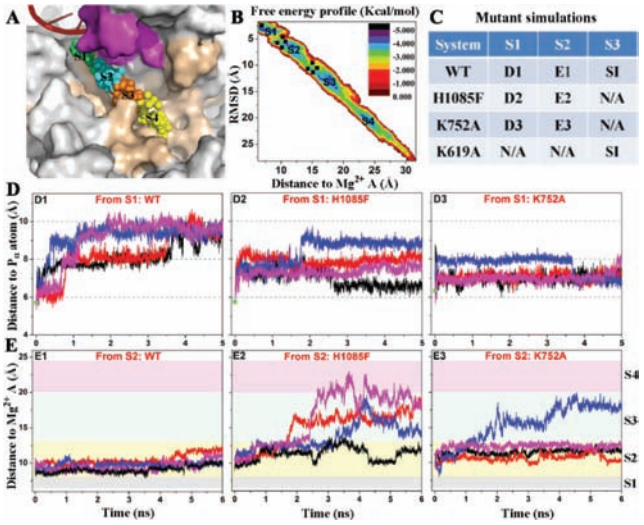


Figure 4. Single-mutant simulations of three key residues, H1085, K752, and K619, to elucidate their individual roles in the PP_i release. (A) The center of mass of the PP_i of 100 randomly selected conformations from each metastable state is superimposed. (B) Projection of the PP_i free energy landscape on two reaction coordinates: the distance between the PP_i group and $Mg^{2+}A$ and the rmsd of the PP_i compared to its bound conformation (see Supporting Information Figure S1A). Mutant simulations are initiated from different points of the energy landscape as shown in black squares. (C) A table summarizes the mutant simulations with the elements matching the labels of the simulation results displayed in part D. (D) and (E) Distance as a function of simulation time for different mutant simulations. For simulations starting from S1 (D1−D3), distance between the PP_i group and the P_α atom of the terminal RNA is shown. For simulations starting from the S2 and S3 states (E1−E3 and Supporting Information Figure S9, parts A and B), distance between the PP_i group and the $Mg^{2+}A$ is shown. In panels E1−E3, the four metastable states are highlighted in different colors: S1 (gray), S2 (light yellow), S3 (light blue), and S4 (pink).

transitions), and K619A (for the S2 → S3 transition). We chose these substitutions because H1085 and K752 are involved in the interactions with PP_i at the S1 and S2 states, whereas K619 is mostly involved in the interaction with PP_i at the S3 state during PP_i release (see Figure 3B). To visualize the selected initial conformations of our simulations, we projected the free energy landscape of the PP_i release onto two reaction coordinates: the distance between the PP_i group and the $Mg^{2+}A$ in the active site and the rmsd value between the PP_i conformation and that in the initial modeled PP_i-bound complex (see Figure 4B). This free energy profile also displays four minimums, which are consistent with the four metastable states identified by our MSM.

For the S1 → S2 transition, our single-variant simulations indicate that residues H1085 and K752 can facilitate this transition and help PP_i escape the active site. We use the distance between the P_α atom of the RNA terminal nucleotide and the P_β atom of the PP_i ($d_{\alpha\beta}$) to measure the distance between the PP_i group and the active site. In the WT simulations, PP_i tends to escape the active site with $d_{\alpha\beta}$ dramatically increasing from 5.7 Å to around 10 Å within 5 ns in all four simulations (see Figure 4D1). By losing its interactions with the residue H1085, the PP_i group in the

dx.doi.org/10.1021/ja210656k | J. Am. Chem.Soc. 2012, 134, 2399−2406

H1085F mutant shows much weaker tendency to leave the active site, with the $d_{\alpha\beta}$ value ending at around 8 Å in most simulations (see Figure 4D2). More strikingly, the PP_i group in the K752A mutant becomes even more stable in the active site, with the $d_{\alpha\beta}$ value at around only 7 Å in three out of four simulations (see Figure 4D3). These results suggest that both H1085 and K752 can help PP_i exit the pol II active site and transition from S1 to S2 state.

In sharp contrast, K752 and H1085 play quite different roles in the S2 → S3 transition. All four 6 ns simulations initiated from S2 in the WT stay in the same S2 state, indicating that S2 is metastable (see Figure 4E1). Only one of four K752A mutant simulations allows PP_i to escape S2, which suggests that the interaction between K752 and the PP_i group only plays a minor role in keeping the PP_i group in S2 (see Figure 4E2). However, the H1085F mutant decreases the metastability of PP_i in S2, with three of four simulations further release to the S3. Therefore, strong interactions between H1085 and the PP_i group can keep it metastable in S2. Further release of the PP_i from S2 is coupled with a tip-opening motion of TL in which the side chain of the H1085 moves away from the closed state. This will eventually disrupt the interactions between PP_i and H1085.

We next investigated the role of K619 in PP_i movement from the S2 to S3 state. With the substitution K619A, two out of four simulations initiated from the S3 but close to the boundary with S2 spend a significant fraction of time back in the state S2 (see Supporting Information Figure S9B). In contrast, none of the simulations initiated from the same four points in the S3 go back to the state S2 in the WT simulations (see Supporting Information Figure S9A). Moreover, one of the WT simulations even reaches the state S4 (the blue curve in Supporting Information Figure S9A). Another WT simulation (black curve in Supporting Information Figure S9A) also samples the region close to the boundary between the S3 and S4. These results clearly show that the positively charged residue K619 facilitates PP_i movement from the state S2 to S3 and further transfer to the state S4. Therefore, K619 may serve as the hopping site for the state S3 in our hopping model for PP_i. Interestingly, a number of other enzymes employ similar hopping mechanisms for facilitating diffusion across long distances, such as efficient long-distance electron translocation within proteins[78,79] or searching for the target binding site/damage site along the DNA for DNA-binding proteins, restriction enzymes, and DNA repair proteins.[80,81]

PP_i Release Is Coupled with the Tip Motion of the Trigger Loop, but Not with Translocation. It has been proposed that release of PP_i leads to opening of the TL.[28,47] However, it is not clear how these two motions are correlated. Does PP_i release precede or follow TL opening? Our results suggest that these two motions are highly coupled (see Figure 5). In particular, the dynamic fluctuations of the TL can help the PP_i group to escape the active site, while release of PP_i may also be facilitated by the opening motion of the TL. After the catalysis, the large intrinsic motion of the TL promotes movement of PP_i from the active site to the first metastable state S2. Specifically, simulations of mutants suggest that the tip part of the trigger loop facilitates PP_i to transfer from the S1 to S2 state by the interactions between the TL residue H1085 and PP_i (see Figure 5, parts A and B). Substitution H1085 with Phe tends to trap PP_i in the active site.

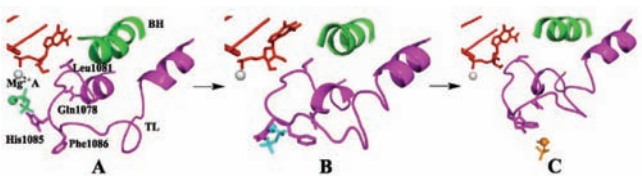


Figure 5. Coupling between the trigger loop motion and PP_i release. The locations of the $(Mg-PP_i)^{2-}$ group in parts A, B, and C are the same as the S1, S2, and S3 in Figure 3, respectively, and the representation of the pol II is the same as in Figure 2B.

On the other hand, once PP_i leaves the active site, its further release can help the opening of the TL by dragging down the TL-tip residue H1085 through electrostatic interactions between this residue and the PP_i. Without this histidine residue, the PP_i group can release much faster in the state S2 (see Figure 4E). Our results also suggest that PP_i release may only aid the early stages of TL opening and that the overall time scale of PP_i release is significantly shorter than the time scale for the swinging motion of the TL to a full open state. We draw this conclusion because we did not observe the full open state for the TL in our simulations of PP_i release. In particular, the average transition time for PP_i along the pore region from the state S1 to S4 is only around 1.5 μs, calculated as MFPT from our MSM (see Supporting Information Table S1). Importantly, we did not observe any pol II translocation event in our PP_i release simulations over this time scale. Averaged pol II elongation rates are typically around 1 or a few kb min^{-1}. Recently, Maiuri et al.[82] have discovered a superefficient human pol II system with an elongation rate of ~100 kb min^{-1} (~600 μs/base). Even in this highly efficient system, the time scale of PP_i release (~1.5 μs) is still around 2 orders of magnitude faster than the nucleotide addition. Therefore, the PP_i release in pol II is unlikely to be coupled with the translocation step due to the faster dynamics of PP_i release. This conclusion is consistent with recent experimental results.[49,50]

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

This study is focused on the molecular mechanisms of PP_i release along the pore segment of the secondary channel in pol II. We discovered four critical metastable states during the PP_i release and proposed a hopping model (see Figures 3 and 6). Each of the metastable states between S2 and S4 serves as a

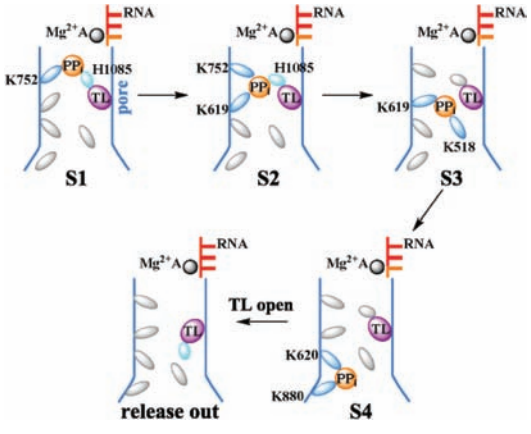


Figure 6. Cartoon model of PP_i release and its couplings with the opening of the trigger loop.

hopping site with two or three pol II positively charged residues acting as the anchor points to provide favorable interactions with the $(Mg-PP_i)^{2-}$. In particular, three residues, Rpb1 residues K752, K619, and H1085 on the trigger loop, were found to play most important roles during PP_i release along the pore region. K752 and K619 residues are located in the entry or E site, previously proposed for NTP loading. Batada *et al.* shows that loss of the E site can greatly limit the rate of NTP diffusion into the active site, assuming an entry route through the secondary channel.[46] In this study, we investigated the reverse process, PP_i release, and also found these two residues can greatly help the PP_i release. Specifically, K752 helps PP_i to escape from the active site (state S1 to S2), whereas K619 is involved in the later stage of the release (state S2 to S3). We also discovered that the critical TL residue H1085 can play similar role as K752 to help the PP_i group escape the active site. Additionally, H1085 serves as a modulation point between the TL and PP_i group so that the further release of the PP_i helps the opening of the TL tip by dragging down the side chain of H1085. These transient interactions of the TL with PP_i in its distinct hopping sites could play a key role in guiding the TL to unfold without becoming trapped in conformations or by interactions that could potentially slow transcription. Our results also suggest that the PP_i release occurs before full opening of TL due to the faster dynamics of PP_i. Thus, our results suggest PP_i release is not coupled with translocation.

Site-directed mutagenesis experiments can be designed to test key predictions of our computational results. For example, substitutions of K752 and K619 to hydrophobic/acidic residues may disrupt the positively charged pocket identified from our simulations and slow down the rate for both PP_i release and the reverse process, NTP addition. Alternatively, changes such as increasing the positive charges in hopping sites may be able to trap stable intermediate states during the PP_i release. Finally, chemical cross-linking between the trigger loop and pol II may also help trap the state with the TL closed while PP_i still occupies the active site.

It will also be of interest to compare the PP_i release mechanisms between eukaryotic and prokaryotic RNA polymerases. Crystal structures of the RNAP from *T. thermophilus* shows that the secondary channels in the eukaryote and prokaryote are quite different.[26] Specifically the secondary channel of the bacterial RNAP is shorter compared to the yeast channel, which may lead to faster PP_i release. Moreover, the bacterial RNAP contains an extra positively charged TL residue, R1239 in *T. thermophilus* (R933 in *E. coli*), in addition to the H1242 (H936 in *E. coli*). Mutagenesis studies have shown that changing this arginine to alanine decreases the elongation rate.[41] Another striking difference between the bacterial and yeast RNA polymerase is that the positively charged residue K518 in yeast is replaced with a negatively charged residue E776 in *T. thermophilus* or E497 in *E. coli* (see Figure 3C). This change in the pore region of the secondary channel may also lead to different mechanisms for PP_i release in eukaryotic and prokaryotic RNA polymerases.

■ ASSOCIATED CONTENT

Supporting Information

Discussions about the protonation states of PP_i and the TL residue H1085 after the catalysis (text and Figures S1, S5, and S7), RESP charges for fitting the force field of the $(Mg-PP_i)^{2-}$ group and its protonated form $(Mg-PP_iH)^{-}$ (Figure S2), three pulling directions in the steered MD simulations (Figure S3), implied time scales plots of MSMs (Figure S4), flowchart of our simulation methodology (Figure S6), comparisons of the location of three conserved residues in the secondary channel of the yeast RNA pol II and bacterial RNAP (Figure S8), mutation simulation results to elucidate the role of Rpb1 residue K619 on PP_i release (Figure S9), and MFPT results (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

xuhuihuang@ust.hk; dongwang@uscd.edu

■ ACKNOWLEDGMENTS

X.H. acknowledges the Hong Kong Research Grants Council (661011, DAG11SC02G, and HKUST6/CRF/10). D.W. acknowledges the NIH (GM085136), start-up funds from Skaggs School of Pharmacy and Pharmaceutical Sciences, UCSD, and an Academic Senate Research Award from UCSD. Computing resources were provided by the CCB computer cluster in HKUST and the Dawning TC5000 supercomputing cluster in Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. We thank Dr. Robert Landick for his extensive critical reading of the manuscript.

■ REFERENCES

(1) Lee, T. I.; Young, R. A. *Annu. Rev. Genet.* **2000**, *34*, 77−137.
(2) Hurwitz, J. *J. Biol. Chem.* **2005**, *280*, 42477−42485.
(3) Shilatifard, A.; Conaway, R. C.; Conaway, J. W. *Annu. Rev. Biochem.* **2003**, *72*, 693−715.
(4) Fuda, N. J.; Ardehali, M. B.; Lis, J. T. *Nature* **2009**, *461*, 186−192.
(5) Kornberg, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12955−12961.
(6) Cramer, P.; Bushnell, D. A.; Fu, J.; Gnatt, A. L.; Maier-Davis, B.; Thompson, N. E.; Burgess, R. R.; Edwards, A. M.; David, P. R.; Kornberg, R. D. *Science* **2000**, *288*, 640−649.
(7) Cramer, P.; Bushnell, D. A.; Kornberg, R. D. *Science* **2001**, *292*, 1863−1876.
(8) Gnatt, A. L.; Cramer, P.; Fu, J.; Bushnell, D. A.; Kornberg, R. D. *Science* **2001**, *292*, 1876−1882.
(9) Kettenberger, H.; Armache, K. J.; Cramer, P. *Mol. Cell* **2004**, *16*, 955−965.
(10) Westover, K. D.; Bushnell, D. A.; Kornberg, R. D. *Cell* **2004**, *119*, 481−489.
(11) Westover, K. D.; Bushnell, D. A.; Kornberg, R. D. *Science* **2004**, *303*, 1014−1016.
(12) Temiakov, D.; Zenkin, N.; Vassylyeva, M. N.; Perederina, A.; Tahirov, T. H.; Kashkina, E.; Savkina, M.; Zorov, S.; Nikiforov, V.; Igarashi, N. *Mol. Cell* **2005**, *19*, 655−666.
(13) Wang, D.; Bushnell, D. A.; Westover, K. D.; Kaplan, C. D.; Kornberg, R. D. *Cell* **2006**, *127*, 941−954.
(14) Brueckner, F.; Hennecke, U.; Carell, T.; Cramer, P. *Science* **2007**, *315*, 859−862.
(15) Brueckner, F.; Cramer, P. *Nat. Struct. Mol. Biol.* **2008**, *15*, 811−818.
(16) Damsma, G. E.; Cramer, P. *J. Biol. Chem.* **2009**, *284*, 31658−31663.
(17) Kostrewa, D.; Zeller, M. E.; Armache, K. J.; Seizl, M.; Leike, K.; Thomm, M.; Cramer, P. *Nature* **2009**, *462*, 323−330.
(18) Sydow, J. F.; Brueckner, F.; Cheung, A.; Damsma, G. E.; Dengl, S.; Lehmann, E.; Vassylyev, D.; Cramer, P. *Mol. Cell* **2009**, *34*, 710−721.
(19) Sydow, J. F.; Cramer, P. *Curr. Opin. Struct. Biol.* **2009**, *19*, 732−739.
(20) Wang, D.; Bushnell, D. A.; Huang, X.; Westover, K. D.; Levitt, M.; Kornberg, R. D. *Science* **2009**, *324*, 1203−1206.

(21) Liu, X.; Bushnell, D. A.; Wang, D.; Calero, G.; Kornberg, R. D. *Science* **2010**, *327*, 206−209.
(22) Wang, D.; Zhu, G.; Huang, X.; Lippard, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 9584−9589.
(23) Cheung, A. C. M.; Cramer, P. *Nature* **2011**, *471*, 249−253.
(24) Vassylyev, D. G.; Sekine, S.; Laptenko, O.; Lee, J.; Vassylyeva, M. N.; Borukhov, S.; Yokoyama, S. *Nature* **2002**, *417*, 712−719.
(25) Vassylyev, D. G.; Vassylyeva, M. N.; Perederina, A.; Tahirov, T. H.; Artsimovitch, I. *Nature* **2007**, *448*, 157−162.
(26) Vassylyev, D. G.; Vassylyeva, M. N.; Zhang, J.; Palangat, M.; Artsimovitch, I.; Landick, R. *Nature* **2007**, *448*, 163−168.
(27) Tagami, S.; Sekine, S.; Kumarevel, T.; Hino, N.; Murayama, Y.; Kamegamori, S.; Yamamoto, M.; Sakamoto, K.; Yokoyama, S. *Nature* **2010**, *468*, 978−982.
(28) Brueckner, F.; Ortiz, J.; Cramer, P. *Curr. Opin. Struct. Biol.* **2009**, *19*, 294−299.
(29) Nedialkov, Y. A.; Gong, X. Q.; Hovde, S. L.; Yamaguchi, Y.; Handa, H.; Geiger, J. H.; Yan, H.; Burton, Z. F. *J. Biol. Chem.* **2003**, *278*, 18303−18312.
(30) Huang, X.; Wang, D.; Weiss, D. R.; Bushnell, D. A.; Kornberg, R. D.; Levitt, M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15745−15750.
(31) Kaplan, C. D.; Larsson, K. M.; Kornberg, R. D. *Mol. Cell* **2008**, *30*, 547−556.
(32) Toulokhonov, I.; Zhang, J.; Palangat, M.; Landick, R. *Mol. Cell* **2007**, *27*, 406−419.
(33) Yulia, Y.; Aleksandra, B.; Vasisht, T.; Mohammad, R.; Savva, Z.; Konstantin, S.; Nikolay, Z. *BMC Biol.* **2010**, *8*, 54−69.
(34) Sosunov, V.; Sosunova, E.; Mustaev, A.; Bass, I.; Nikiforov, V.; Goldfarb, A. *EMBO J.* **2003**, *22*, 2234−2244.
(35) Yang, W.; Lee, J. Y.; Nowotny, M. *Mol. Cell* **2006**, *22*, 5−13.
(36) Svetlov, V.; Nudler, E. *Curr. Opin. Struct. Biol.* **2009**, *19*, 701−707.
(37) Feig, M.; Burton, Z. F. *Biophys. J.* **2010**, *99*, 2577−2586.
(38) Svetlov, V.; Vassylyev, D. G.; Artsimovitch, I. *J. Biol. Chem.* **2004**, *279*, 38087−38090.
(39) Trinh, V.; Langelier, M. F.; Archambault, J.; Coulombe, B. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 12−36.
(40) Tan, L.; Wiesler, S.; Trzaska, D.; Carney, H. C.; Weinzierl, R. O. J. *J. Biol.* **2008**, *7*, 40−54.
(41) Zhang, J.; Palangat, M.; Landick, R. *Nat. Struct. Mol. Biol.* **2009**, *17*, 99−104.
(42) Yuzenkova, Y.; Zenkin, N. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 10878−10883.
(43) Zenkin, N.; Yuzenkova, Y.; Severinov, K. *Science* **2006**, *313*, 518−520.
(44) Castro, C.; Smidansky, E.; Maksimchuk, K. R.; Arnold, J. J.; Korneeva, V. S.; Gotte, M.; Konigsberg, W.; Cameron, C. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4267−4272.
(45) Castro, C.; Smidansky, E. D.; Arnold, J. J.; Maksimchuk, K. R.; Moustafa, I.; Uchida, A.; Götte, M.; Konigsberg, W.; Cameron, C. E. *Nat. Struct. Mol. Biol.* **2009**, *16*, 212−218.
(46) Batada, N. N.; Westover, K. D.; Bushnell, D. A.; Levitt, M.; Kornberg, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17361−17364.
(47) Yin, Y. W.; Steitz, T. A. *Cell* **2004**, *116*, 393−404.
(48) Bar-Nahum, G.; Epshtein, V.; Ruckenstein, A. E.; Rafikov, R.; Mustaev, A.; Nudler, E. *Cell* **2005**, *120*, 183−193.
(49) Abbondanzieri, E. A.; Greenleaf, W. J.; Shaevitz, J. W.; Landick, R.; Block, S. M. *Nature* **2005**, *438*, 460−465.
(50) Hein, P. P.; Palangat, M.; Landick, R. *Biochemistry* **2011**, *50*, 7002−7014.
(51) Guajardo, R.; Sousa, R. *J. Mol. Biol.* **1997**, *265*, 8−19.
(52) Xiong, Y.; Burton, Z. F. *J. Biol. Chem.* **2007**, *282*, 36582−36592.
(53) Golosov, A. A.; Warren, J. J.; Beese, L. S.; Karplus, M. *Structure* **2010**, *18*, 83−93.
(54) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. *Nucleic Acids Res.* **2003**, *31*, 3381−3385.
(55) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(56) Meagher, K. L.; Redman, L. T.; Carlson, H. A. *J. Comput. Chem.* **2003**, *24*, 1016−1025.
(57) Dupradeau, F. Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821−7839.
(58) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(59) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel Publishing Company: Dordrecht, The Netherlands, 1981.
(60) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(61) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(62) Berendsen, H. J. C.; Postma, J. P. M; Van Gunsteren, W. F.; DiNola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(63) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(64) Isralewitz, B.; Gao, M.; Schulten, K. *Curr. Opin. Struct. Biol.* **2001**, *11*, 224−230.
(65) Bowman, G. R.; Huang, X.; Pande, V. S. *Methods* **2009**, *49*, 197−201.
(66) Silva, D. A.; Bowman, G. R.; Sosa-Peinado, A.; Huang, X. *PLoS Comput. Biol.* **2011**, *7*, e1002054.
(67) Bowman, G. R.; Voelz, V. A.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2010**, *21*, 4−11.
(68) Noé, F.; Fischer, S. *Curr. Opin. Struct. Biol.* **2008**, *18*, 154−162.
(69) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. *J. Chem. Phys.* **2007**, *126*, 155101−155117.
(70) Deuflhard, P.; Weber, M. *Linear Algebra Appl.* **2005**, *398*, 161−184.
(71) Swope, W. C.; Pitera, J. W.; Suits, F. *J. Phys. Chem. B* **2004**, *108*, 6571−6581.
(72) Deuflhard, P.; Huisinga, W.; Fischer, A.; Schütte, C. *Linear Algebra Appl.* **2000**, *315*, 39−59.
(73) Singhal, N.; Snow, C. D.; Pande, V. S. *J. Chem. Phys.* **2004**, *121*, 415−425.
(74) Huang, X.; Bowman, G. R.; Bacallado, S.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19765−19769.
(75) Zhuang, W.; Cui, R. Z.; Silva, D. A.; Huang, X. *J. Phys. Chem. B* **2011**, *115*, 5415−5424.
(76) Swope, W. C.; Pitera, J. W.; Suits, F. *J. Phys. Chem. B* **2004**, *108*, 6571−6581.
(77) Sriraman, S.; Kevrekidis, I. G.; Hummer, G. *J. Phys. Chem. B* **2005**, *109*, 6479−6484.
(78) Giese, B.; Graber, M.; Cordes, M. *Curr. Opin. Chem. Biol.* **2008**, *12*, 755−759.
(79) Gray, H. B.; Winkler, J. R. *Biochim. Biophys. Acta* **2010**, *1797*, 1563−1572.
(80) Halford, S. E.; Marko, J. F. *Nucleic Acids Res.* **2004**, *32*, 3040−3052.
(81) Friedman, J. I.; Stivers, J. T. *Biochemistry* **2010**, *49*, 4957−4967.
(82) Maiuri, P.; Knezevich, A.; De Marco, A.; Mazza, D.; Kula, A.; McNally, J. G.; Marcello, A. *EMBO Rep.* **2011**, *12*, 1280−2185.