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

Lin-Tai Da[1], Dong Wang[2*], and Xuhui Huang[1*]

[1]Department of Chemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

[2]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA 92093-0625, USA

E-mail: xuhuihuang@ust.hk or dongwang@uscd.edu

Protonation states after the catalysis for PP_i and the TL residue H1085

In order to model the PP_i bound complex, we need to consider the protonation states of the PP_i group and H1085 right after the catalysis. Two mechanisms of the chemical reaction may lead to different protonation states depending on the role of H1085. Specially, if H1085 helps stabilize the transition state structure, it will remain as the protonated form with the unprotonated PP_i (PP_i^{-4}-HIP, see SI Fig. S5A). If H1085 donates the proton to the PP_i, it will take one of its two unprotonated forms (PPH_i^{-3}-HID and PPH_i^{-3}-HIE, see SI Fig. S5B and C). Experimental studies suggest that H1085 may play both roles [1-5]. The system after the catalysis may contain a mixture of different protonation forms.

In addition to PP_i^{-4}-HIP, we also modeled the other two protonation forms: PPH_i^{-3}-HID and PPH_i^{-3}-HIE. In these two systems, the active site residues of the energy minimized modeled structures also take almost identical positions as those in the crystal structure (RMSD ~0.8Å, see SI Fig. S1), which shows our models are reasonable starting points. Further molecular dynamics (MD) simulations suggest that the PP_i group tends to move toward the exit of the active site in all three systems, though the magnitude of the movement in PP_i^{-4}-HIP is slightly larger, which may due to stronger interactions between H1085 and PP_i (see SI Fig. S7). These results indicate that these three systems display similar dynamic behaviors to help PP_i to escape the active site. Furthermore, it has been suggested that in solution at physiological conditions, the major form of the PP_i should carry -4 charge when bounding to an Mg^{2+} ion [6]. Therefore, we only select the PP_i^{-4}-HIP protonation form for further simulations due to the limitation of computing resources.

References

(1) Kaplan, C. D.; Larsson, K. M.; Kornberg, R. D. *Mol. Cell* **2008**, *30*, 547-556.

(2) Castro, C.; Smidansky, E. D.; Arnold, J. J.; Maksimchuk, K. R.; Moustafa, I.; Uchida, A.; G tte, M.; Konigsberg, W.; Cameron, C. E. *Nat. Struct. Mol. Biol.* **2009**, *16*, 212-218.

(3) Zhang, J.; Palangat, M.; Landick, R. *Nat. Struct. Mol. Biol.* **2009**, *17*, 99-104.

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

(5) Yuzenkova, Y.; Zenkin, N. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 10878-10883.

(6) Zhang, Y.; Deng, H.; Schramm, V. L. *J. Am. Chem. Soc.* **2010**, 467-484.

SI Figures

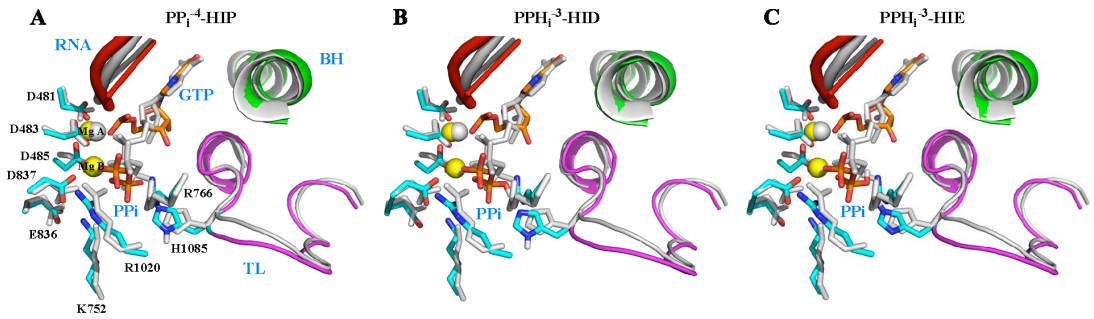
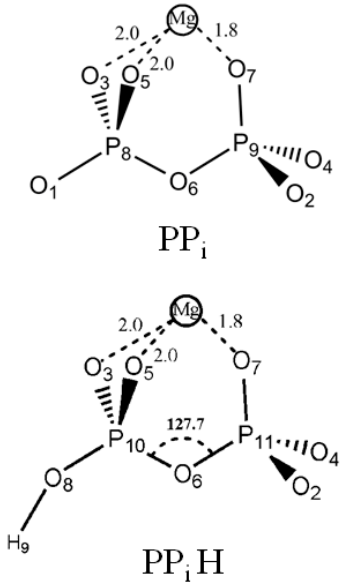


Figure S1. Energy minimized structures of PP_i bound pol II complexes with different protonation states of H1085 and PP_i: (A). PP_i^{-4}-HIP, (B). PPH_i^{-3}-HID and (C). PPH_i^{-3}-HIE are superimposed with the crystal structure of GTP bound complex (PDB ID: 2E2H, color gray). For PP_i bound complexes, the bridge helix, trigger loop (TL), RNA chain, Mg^{2+} ions and the PPi group are shown in green, magenta, red, yellow, and organe/red respecitively. The RMSD values of energy minimized structures of different PP_i bound complexes compared to the GTP bound X-ray structure are 0.77 Å, 0.84 Å and 0.83Å for PP_i^{-4}-HIP, PPH_i^{-3}-HID and PPH_i^{-3}-HIE respectively (residues around 20Å of the Mg^{2+} in the active site are included in the calculations).



PP_i

ATOM	RESP Charge
O1	-0.83
O2	-0.85
O3	-1.03
O4	-0.85
O5	-1.03
O6	-0.64
O7	-1.01
P8	1.51
P9	1.30
Mg10	1.44

PP_iH

ATOM	RESP Charge
Mg1	1.55
O2	-0.81
O3	-0.96
O4	-0.81
O5	-0.96
O6	-0.70
O7	-1.02
O8	-0.66
H9	0.45
P10	1.50
P11	1.41

Figure S2. The calculated RESP charges for the $(Mg\text{-}PP_i)^{2-}$ group and its protonated form: $(Mg\text{-}PP_iH)^{1-}$.

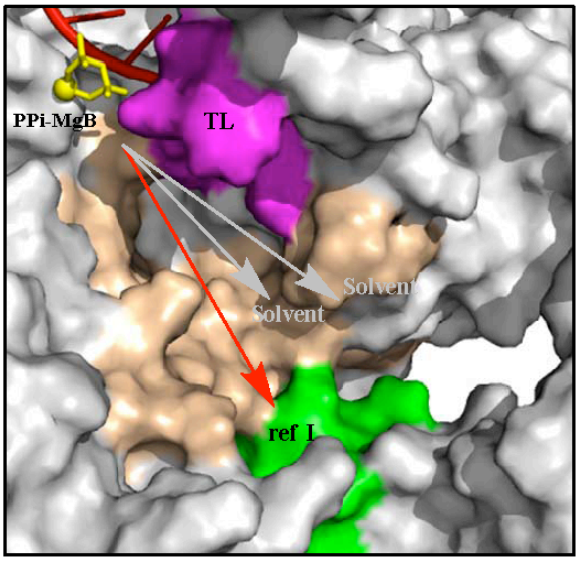


Figure S3. Independent steered Molecular Dynamics(MD) simulations are performed along three directions to generate initial release pathways. One direction (red arrow) is along the wall of the secondary channel using the center coordinate of the C_α atoms from Rpb 1 residues 880-882 and 953-955 (green) as the reference (ref I). In the other two directions (grey color), two other groups: Rpb1 residues 716-720 (group 1) and 726-731 (group2) are chosen, and the center of the ref I and group 1, ref I and group 2 are used as the other two references (ref II and ref III). All the steered simulations are conducted in the same conditions with the force constant of 0.5 kJ $mol^{-1}Å^{-2}$ and pulling rate of 0.01Å/ps (see Methods section for more detailes).

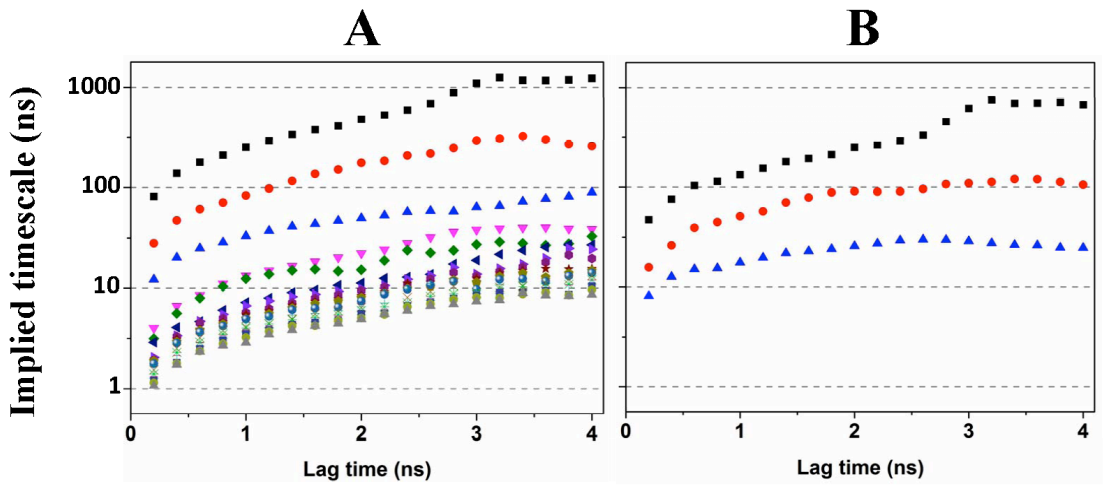


Figure S4. Implied timescale plots as the function of the lag time for Markov State Model (MSM) constructed from (A) 274 microstates and (B) 4 macrostaes.

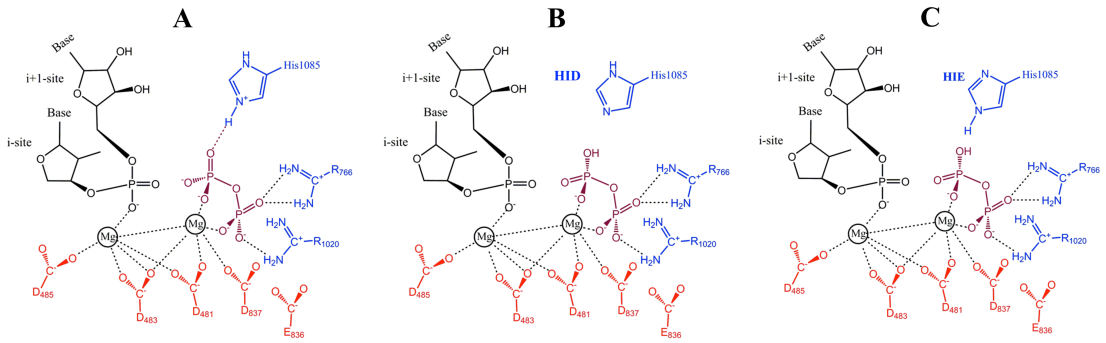


Figure S5. Chemical structures of three PP_i bound pol II complexes with different protonation states of H1085 and PP_i: (A) PP_i^{-4}-HIP, (B) PPH_i^{-3}-HID, and (C) PPH_i^{-3}-HIE.

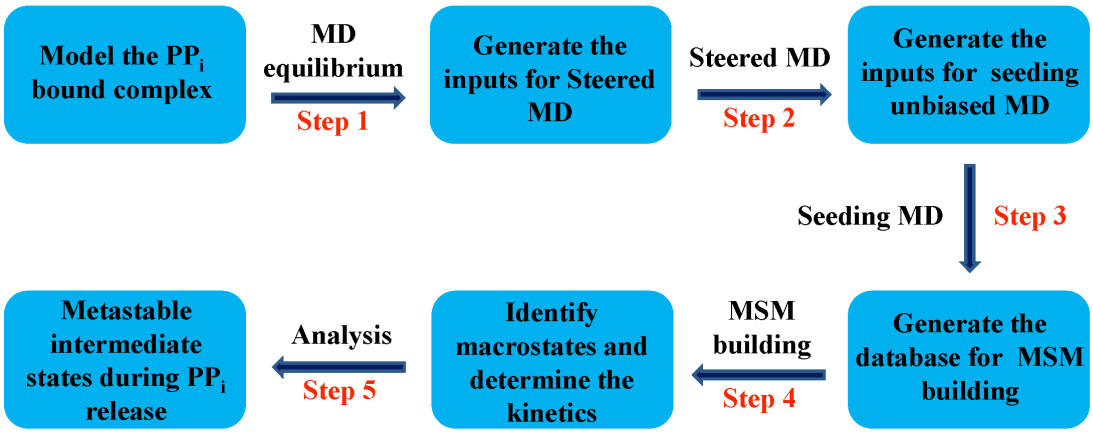


Figure S6. Flow chat of our simulation methodology.

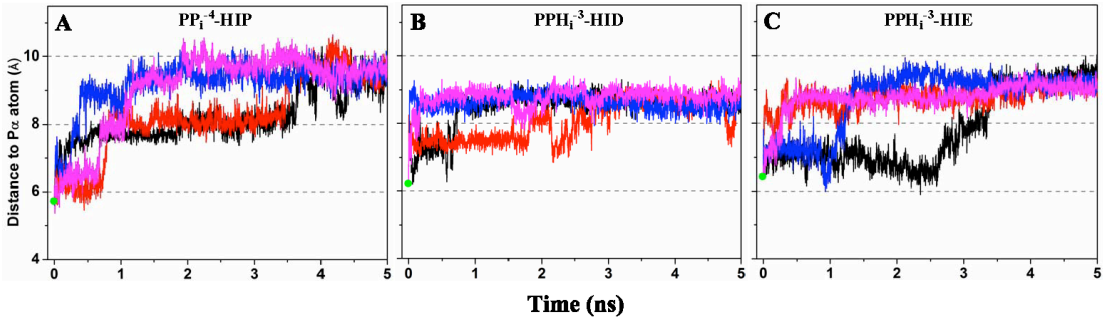


Figure S7. Distance as a function of simulation time for PP_i bound pol II complexes with different protonation states of H1085 and PP_i: (A). PP_i^{-4}-HIP, (B). PPH_i^{-3}-HID and (C). PPH_i^{-3}-HIE. The distance between the P_α atom of the primer nucleotide of RNA chain and the P_β atom of the PP_i group is used. Initial conformationas are labeled using green circles.

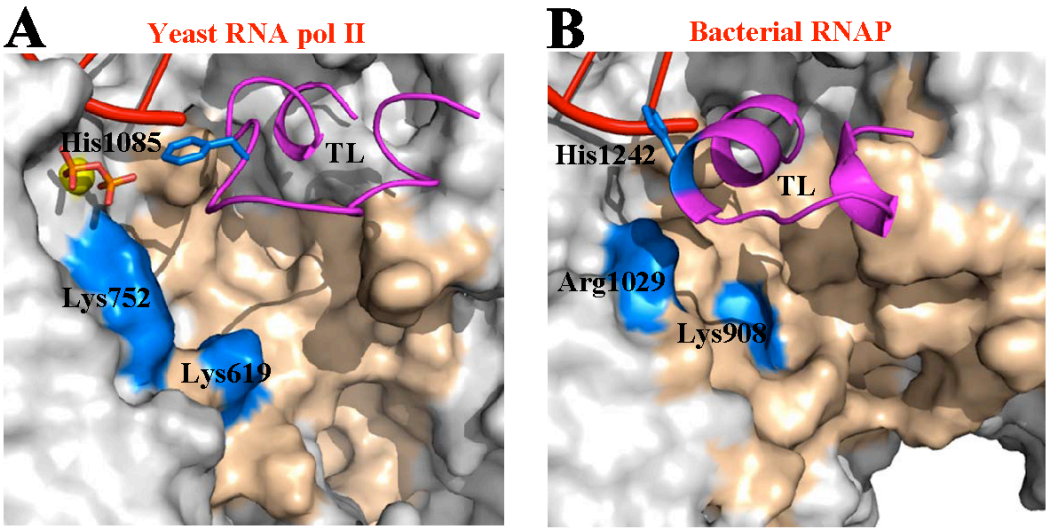


Figure S8. Side-by-side comparisons of three conserved residues in the secondary channel of the yeast RNA pol II (*S. cerevisiae*) (A) and bacterial RNAP (*T. thermophilus*) (B). In both systems, the secondary channal, trigger loop (TL) and the RNA chain are shown in wheet surface, magenta and red cartoon respectively. Three pairs of equilvent and conserved residues: (pol II His1085, RNAP His1242), (pol II Lys752, RNAP Arg1029), and (pol II Lys619, RNAP Lys908) are highlighted in blue.

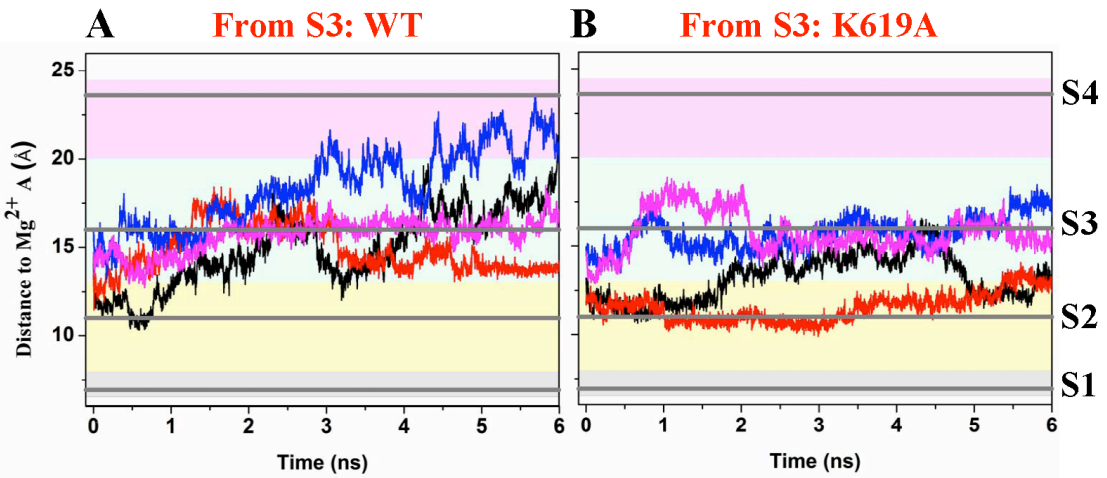


Figure S9. Distance as a function of time for simulations initiated from state S3 for (A) wild type and (B) single mutant K619A. See Fig. 4 for other details.

Table S1. Mean First Passage Time (MFPT) obtained from our MSMs for transitions between different metastable states. See SI section 4 for details of the MFPT calculations.

	S1 to S2	S2 to S3	S3 to S4	S1 to S4
MFPT (ns)	838	111	547	1496