Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

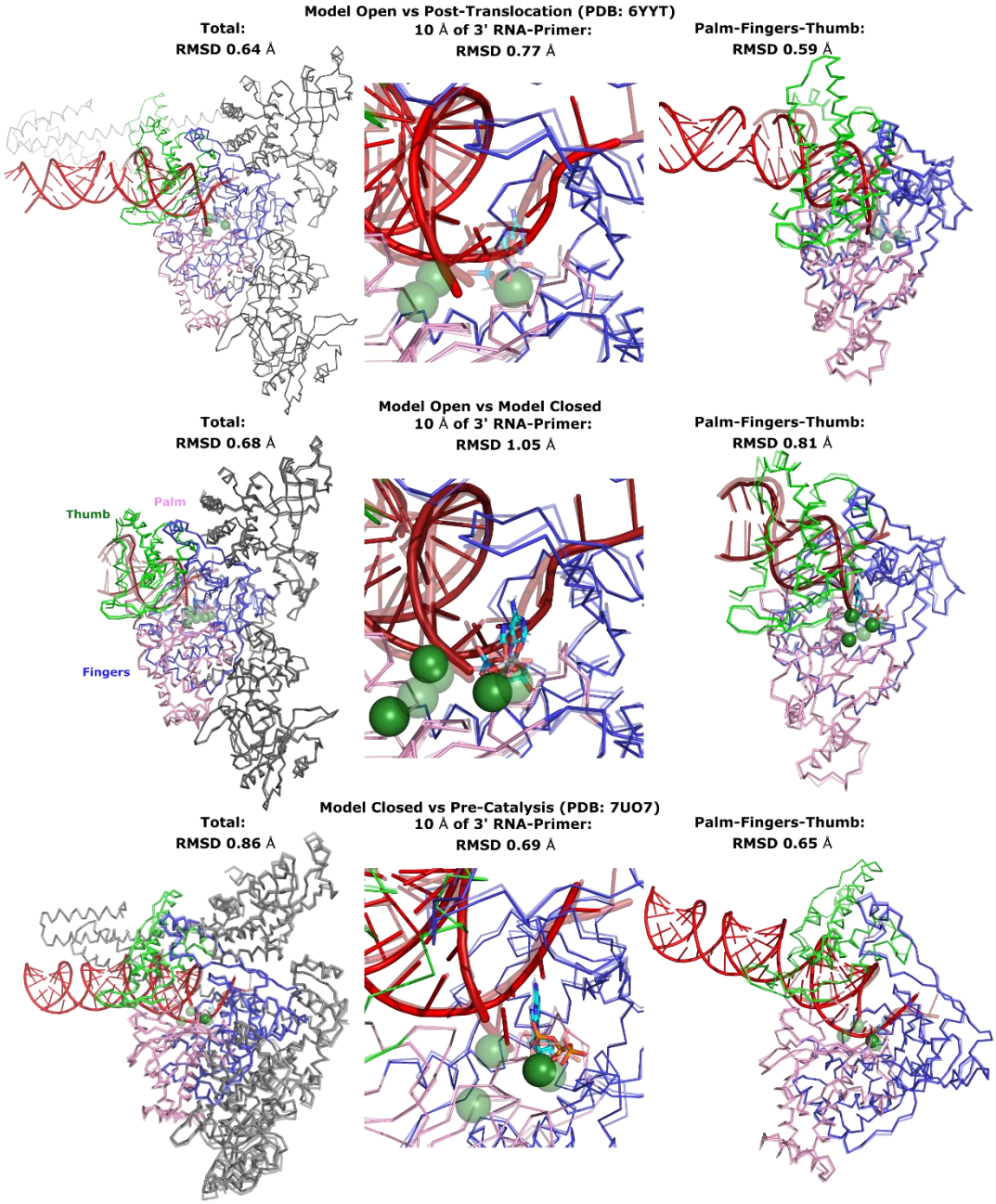

Figure S1. Alignment with RMSD variation of the following comparisons of SARS-CoV-2 RdRp modelled structures used in simulation: modelled open conformation (transparent, PDB:7BV2)[1] versus the more recent post-translocation structure (PDB:6YYT)[2], modelled open conformation (transparent) versus modelled closed conformation (PDB:7BTF)[3], modelled closed conformation (transparent) versus the more recent pre-catalytic structure (PDB:7UO7)[4]. Key subdomain colors have been correlated to main text **Figure 1**: pink-palm, green-thumb, blue-fingers. Left: RMSD of alignment of total protein and total protein structures aligned. Center: RMSD of alignment of all protein residues within 10 Å of the 3' RNA-primer. Right: RMSD of alignment of palm-fingers-thumb subdomain alignment with palm-fingers-thumb subdomain structure (right) and active site structure (center).

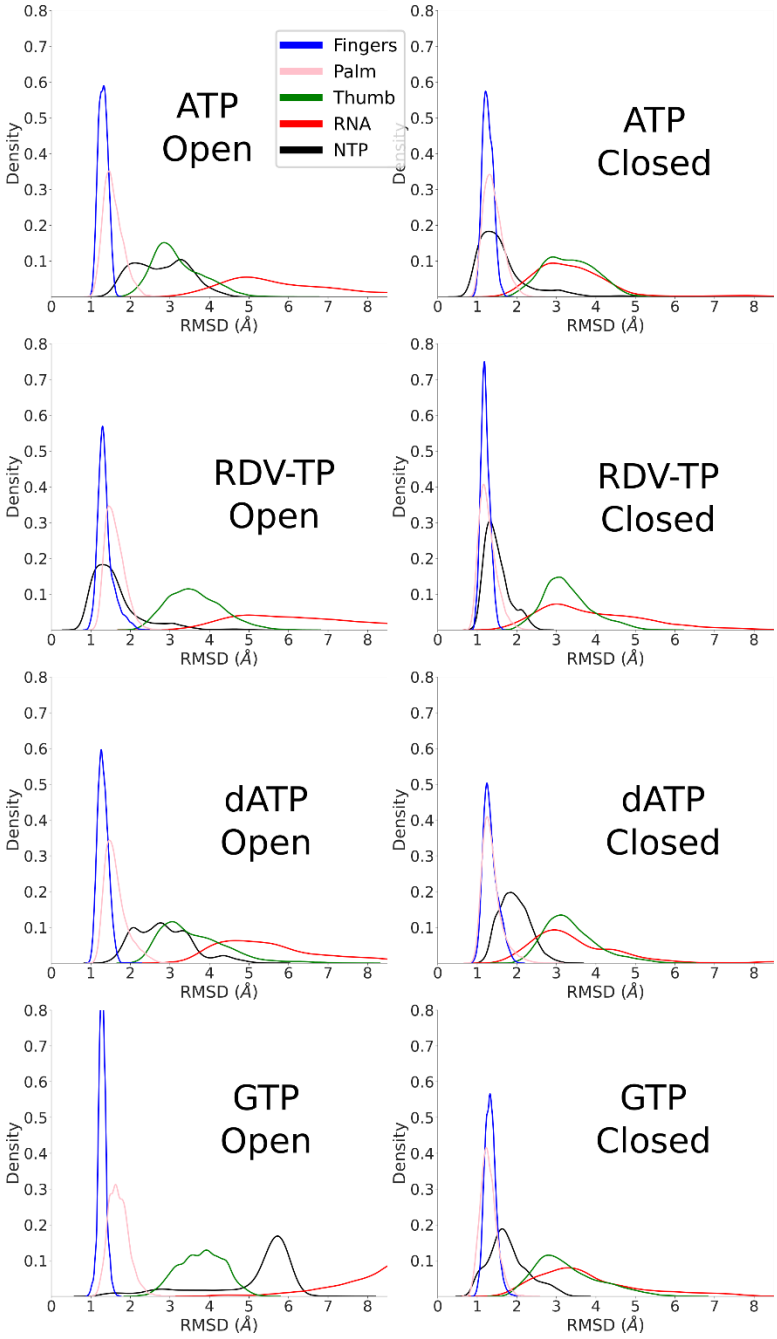

Figure S2. The root-mean-square deviations (RMSDs) of SARS-CoV-2 RdRp structural subdomains (backbone atoms), RNA (phosphate backbone), and NTP (heavy atoms) measured from equilibrium ensemble MD simulations (10x100 ns each system). The RMSDs are shown from *top to bottom* for cognate ATP, drug analog RDV-TP, non-cognate dATP and GTP upon initial binding (active site open; *left*) and insertion (active site closed; *right*) states. The subdomains are shown in different colors: fingers (blue). Palm (pink), thumb (green), RNA (red), and NTP (black).

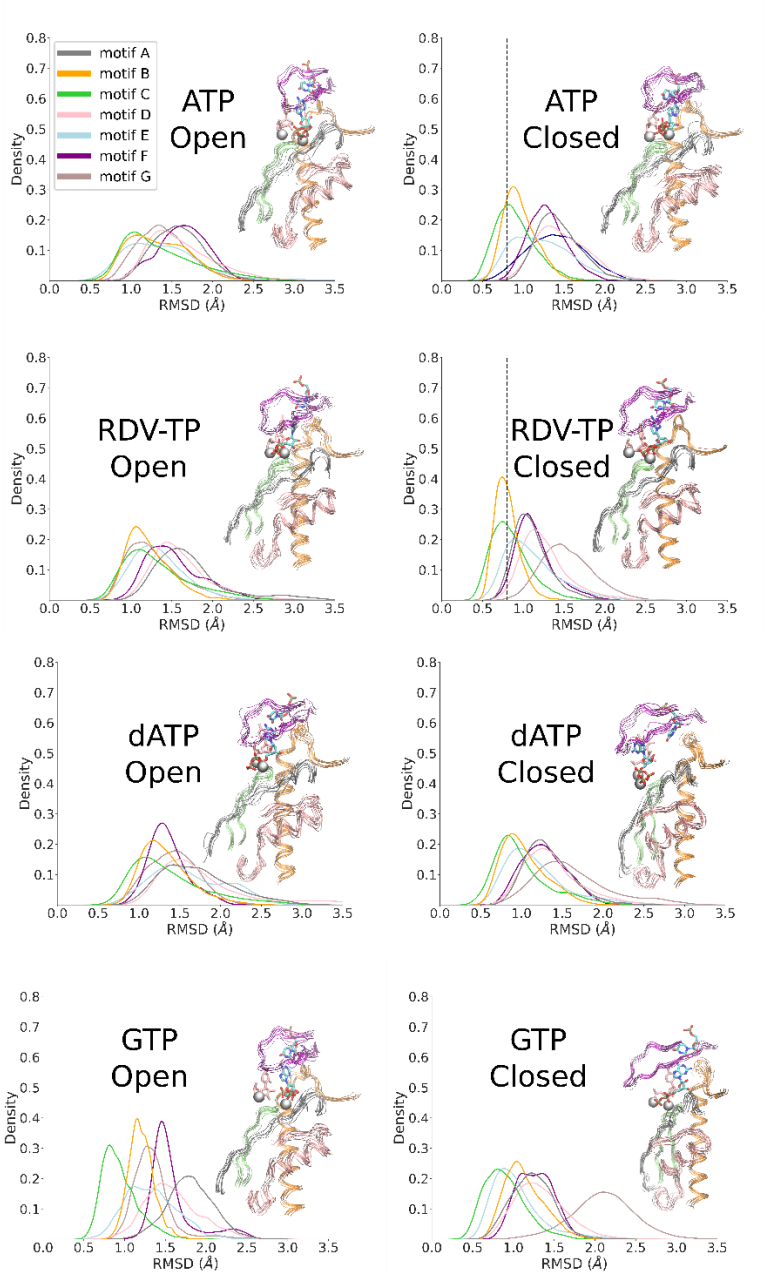

Figure S3. The root-mean-square deviations (RMSDs) of RdRp structural motifs (backbone atoms) measured from equilibrium ensemble MD simulations (10x100 ns each system). The motif RMSDs are shown from *top to bottom* for ATP, RDV-TP, dATP, and GTP upon initial binding (active site open; *left*) and insertion (active site closed; *right*) states. The key motif RMSD are displayed in different colors: motif A (gray), Motif B (orange), Motif C (green), Motif D (Pink), and motif F (purple). Structural representations of those motifs along with NTP, uracil template nucleotide, and two catalytic MG ions are shown for each simulation system. The dotted black line indicates the reference group of motifs (B & C) for systems of inserted ATP/RDV-TP.

Table S1. Average Motif RMSD from the initial binding equilibrium ensemble simulations using the closed state energy-minimized structure as reference. Units are in Angstrom.

Open State	ATP	RDV-TP	dATP	GTP
Motif A	1.6±0.3	1.7±0.5	1.7±0.4	1.8±0.3
Motif B	1.3±0.4	1.2±0.3	1.3±0.3	1.2±0.2
Motif C	1.4±0.5	1.3±0.4	1.4±0.5	0.9±+0.2
Motif D	1.6±0.4	1.6±0.4	1.8±0.6	1.6±0.3
Motif E	1.4±0.5	1.3±0.3	1.6±0.5	1.3±0.3
Motif F	1.6±0.3	1.5±0.3	1.4±0.2	1.6±0.3
Motif G	1.4±0.3	1.3±0.3	1.4±0.3	1.3±0.2

Table S2. Average Motif RMSD from the insertion equilibrium ensemble simulations using the closed state energy-minimized structure as reference. Units are in Angstrom.

Closed State	ATP	RDV-TP	dATP	GTP
Motif A	1.4±0.3	1.1±0.2	1.3±0.3	1.1±0.2
Motif B	1.0±0.2	0.8±0.2	1.0±0.3	1.1±0.2
Motif C	0.9±0.2	0.8±0.2	1.0±0.3	0.8±0.2
Motif D	1.5±0.3	1.3±0.3	1.4±0.4	1.4±0.3
Motif E	1.2±0.4	1.1±0.3	1.2±0.4	1.0±0.3
Motif F	1.3±0.2	1.1±0.2	1.3±0.3	1.3±0.2
Motif G	1.4±0.3	1.5±0.3	1.6±0.5	2.1±0.4

Table S3. Umbrella Sampling Parameters: force constant k for each NTP insertion path and total number of windows used. † specifies the system with forcing on the template +1 nucleotide.

NTP	Forward k $\left(\frac{kcal}{mol\,Å^2}\right)$	Backward k $\left(\frac{kcal}{mol\,Å^2}\right)$	# of windows
GTP	501	501.9	24
GTP†	250	501.9	24
dATP	501	250.95	21
dATP†	501	250.95	35
ATP†	501	501	26
RDV-TP	125	125	21

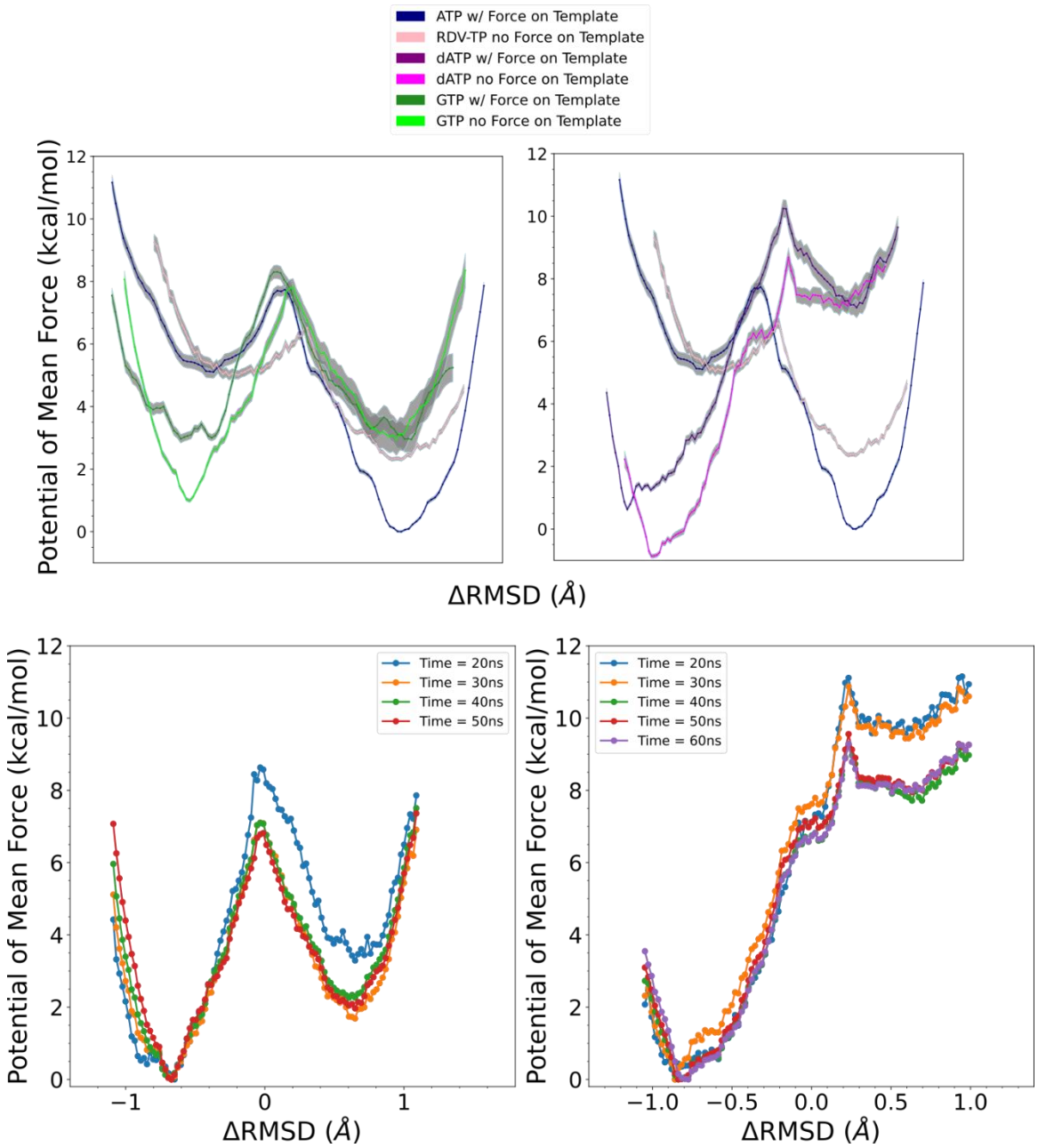

Figure S4. The potentials of mean force (PMFs) calculated for various NTPs from initial binding (active site *open*) to the insertion (*closed*) state via umbrella sampling simulations. The difference of RMSDs with respect to *open* and *closed* reference structures, i.e., DRMSD º RMSD (X, X_{open})- RMSD(X,X_{closed})[5], was used as the reaction coordinate in the PMF construction. The *upper left* panel shows the PMFs for GTP, with (dark green) and without (light green) force on the template +1 nucleotide. In both cases, PMFs are shown in comparison with the PMFs obtained for cognate ATP (blue) and drug RDV-TP (pink)[5]. The *upper right* panel shows the PMFs for dATP, with (dark purple) and without (magenta) force on the template +1 nucleotide. The *lower left* and *lower right* panels display convergence plots of PMFs for GTP and dATP, respectively, in current umbrella sampling simulations, without force implemented to the template +1 nucleotide.

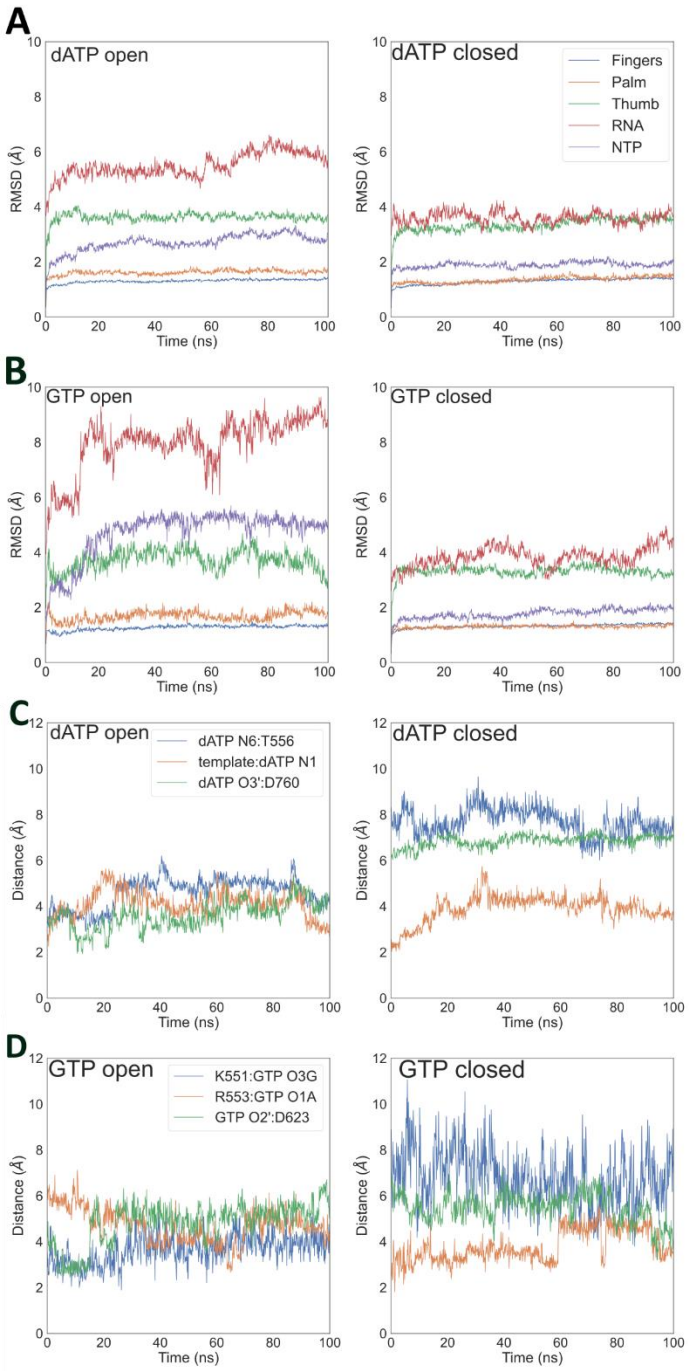

Figure S5. A-B: Open and closed conformation equilibration results for subdomain RSMD (mean of 10 trials) for dATP (A) and GTP (B) after fingers subdomain alignment, mean of 10 trials. ATP and RDV-TP RMSD are available in our previous work figure S12 and S14.[5] C: Open and closed conformation key dATP H-bond donor-acceptor distances as identified in umbrella sampling initial binding (see main text **Figure 7** and **SI Figure S2**), mean of 10 trials. D: Open and closed conformation key GTP H-bond donor-acceptor distances as identified in umbrella sampling initial binding (see main text **Figure 7** and **SI Figure S2**), mean of 10 trials.

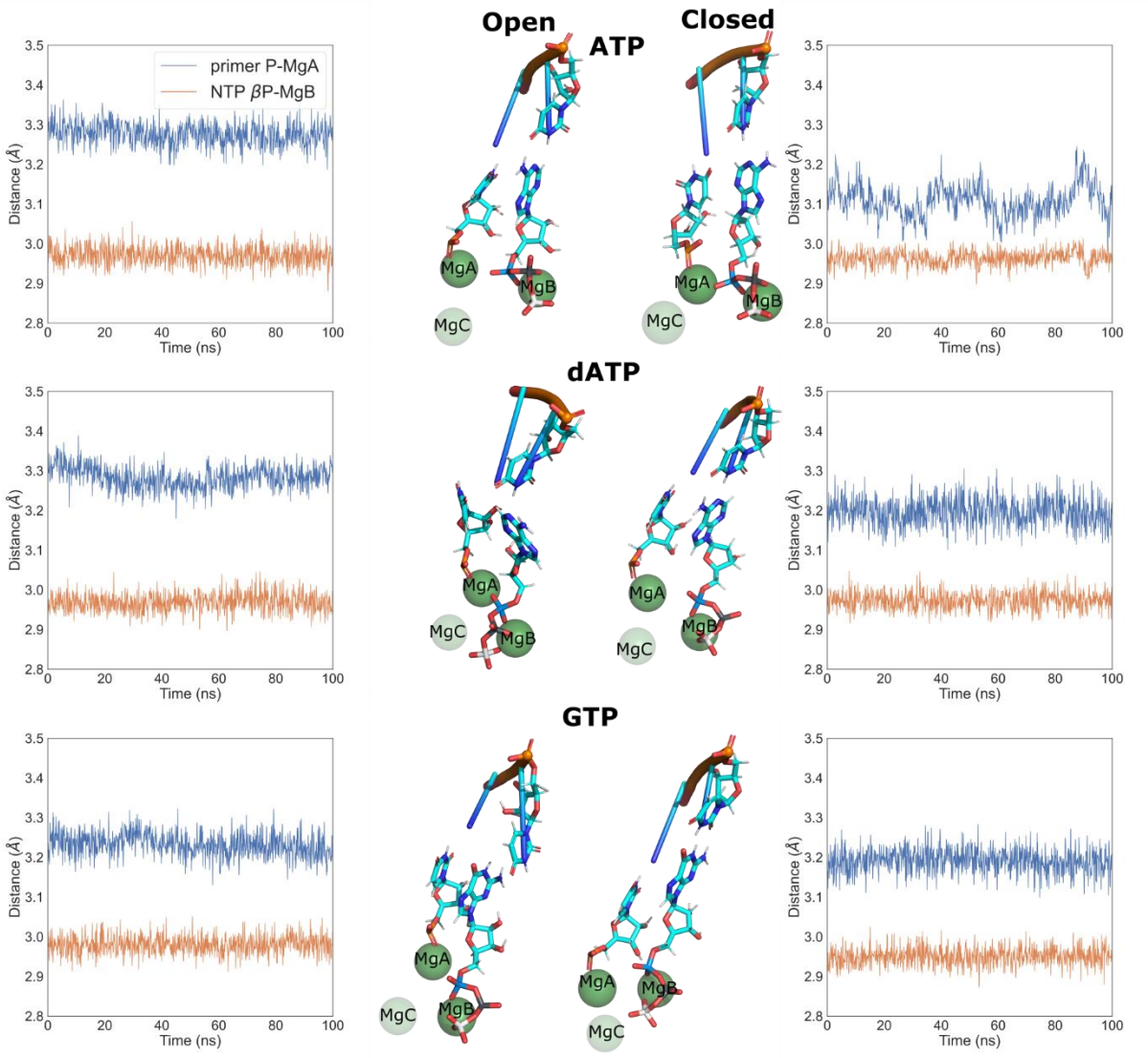

Figure S6. Distance coordination of Mg^{2+} ions: MgA-3' RNA Primer P and MgB-NTP βP; for cognate ATP, non-cognate dATP, and non-cognate GTP, with MgA and MgB the two-metal ions essential for catalysis.[6] Left: mean of 10 trials for open initial binding conformation model. Center: example equilibrium structures of labelled Mg^{2+} with NTP-template and 3' RNA primer-RNA primer template shown. A third MgC ion is also shown, which was captured from an early version PDB structure (7BV2)[1]. Right: mean of 10 trials for closed insertion conformation model. See Romero et al, 2021[5] for a discussion of RDV-TP Mg^{2+} coordination.

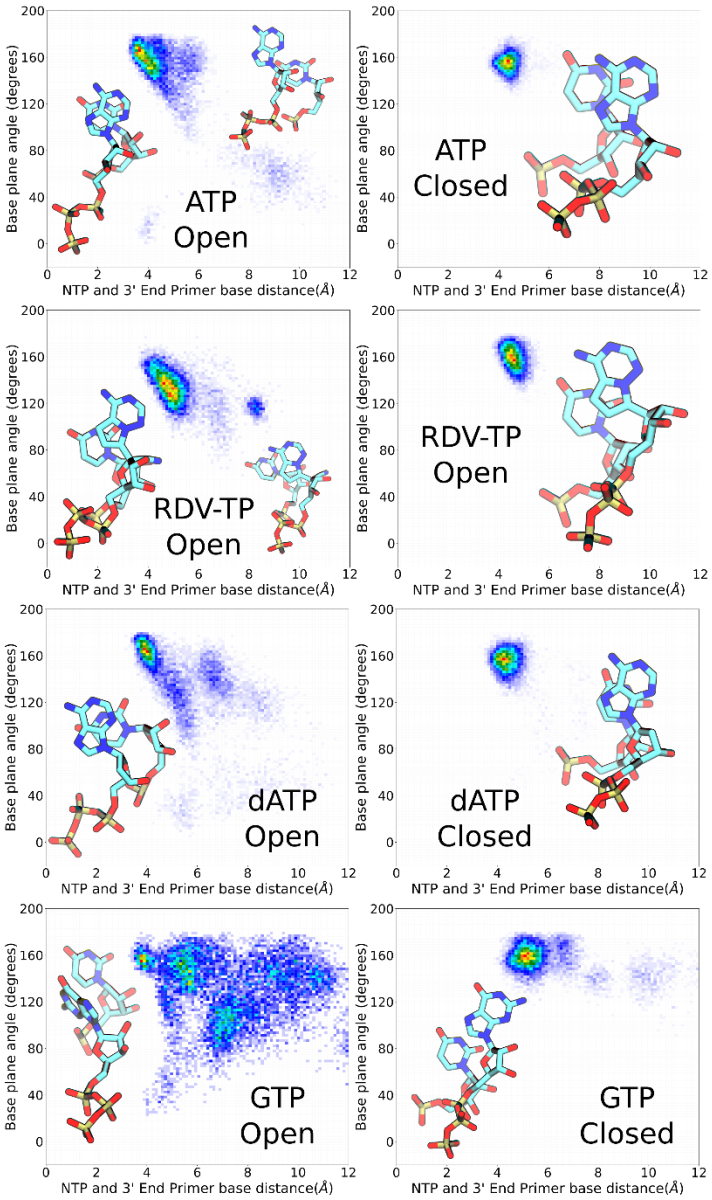

Figure S7. NTP and 3'-end primer association geometries sampled from equilibrium ensemble simulations. The geometric measures (see Methods) are shown between the 3'-end primer and individual incoming NTP *from top to bottom*: ATP, RDV-TP, dATP, and GTP are demonstrated, upon initial binding (*left*) and insertion (*right*) for each NTP species. Licorice representations of the NTP and 3'-end primer show the dominant geometries for each simulation system. Distance is measured by the center of mass between the bases. Base plane angle is measured using the C1'-C2-C5 (3' end primer) and C1'-C7-C5 (NTP).

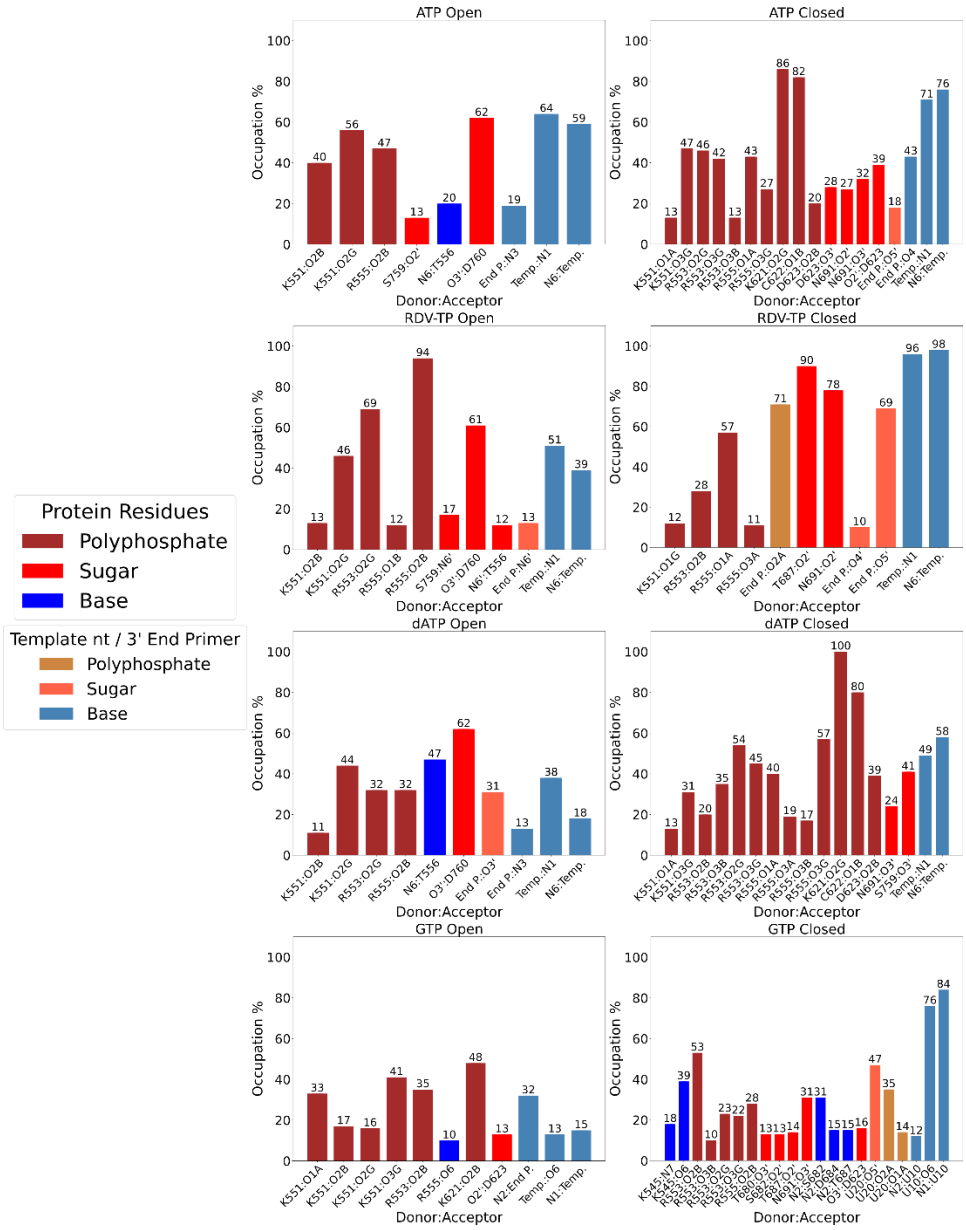

Figure S8. The hydrogen bonding (HB) occupancy from the equilibrium ensemble simulations for each NTP, shown from *top to bottom*: ATP, RDV-TP, dATP, and GTP, upon initial binding (open, *left*) to insertion (closed, *right*). Each unique HB interaction >10% population is considered. Two color-code sets are used: protein-NTP interactions are shown brown (polyphosphate), red (sugar), and blue (base); template-nt / 3' end primer-NTP interactions are shown light brown (polyphosphate), light red (sugar), and light blue (base).

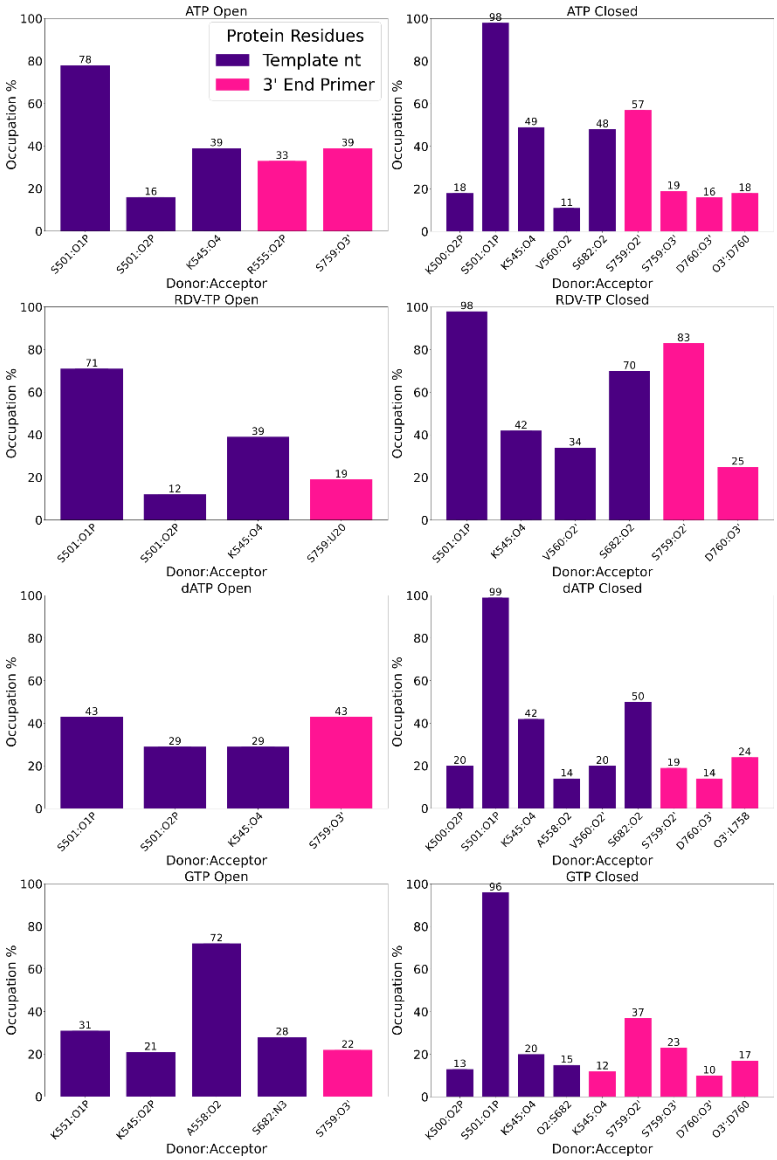

Figure S9. The hydrogen bonding (HB) occupancy from the equilibrium ensemble simulations for protein-template nt Uracil (purple) and protein-3' end primer (pink) for each NTP simulation system, shown from *top to bottom*: ATP, RDV-TP, dATP, and GTP, upon initial binding (open, *left*) to insertion (closed, *right*). Each unique HB interaction >10% population is considered.

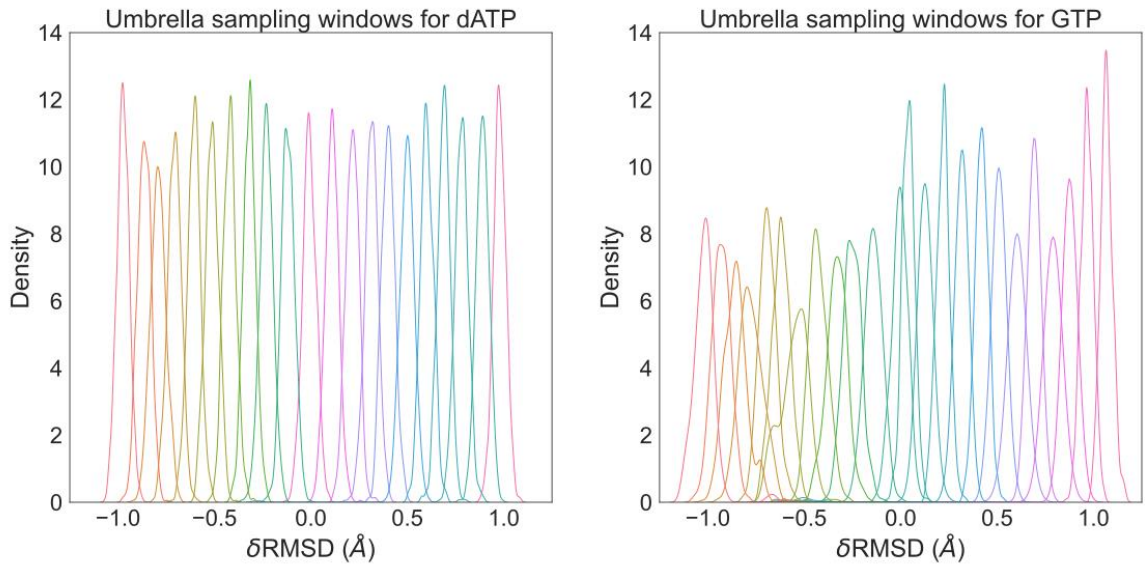

Figure S10. Distribution of $\delta RMSD$ from umbrella sampling of dATP (right) and GTP (left) used to generate converged PMFs. For dATP, 21 windows of 60 ns are used with $501\left(\frac{kcal}{mol\,Å^2}\right)$ force constant on windows 1-11 and $250\left(\frac{kcal}{mol\,Å^2}\right)$ force constant on windows 12-21. For GTP, 24 windows of 50 ns are used with $501\left(\frac{kcal}{mol\,Å^2}\right)$ force constant on all windows.

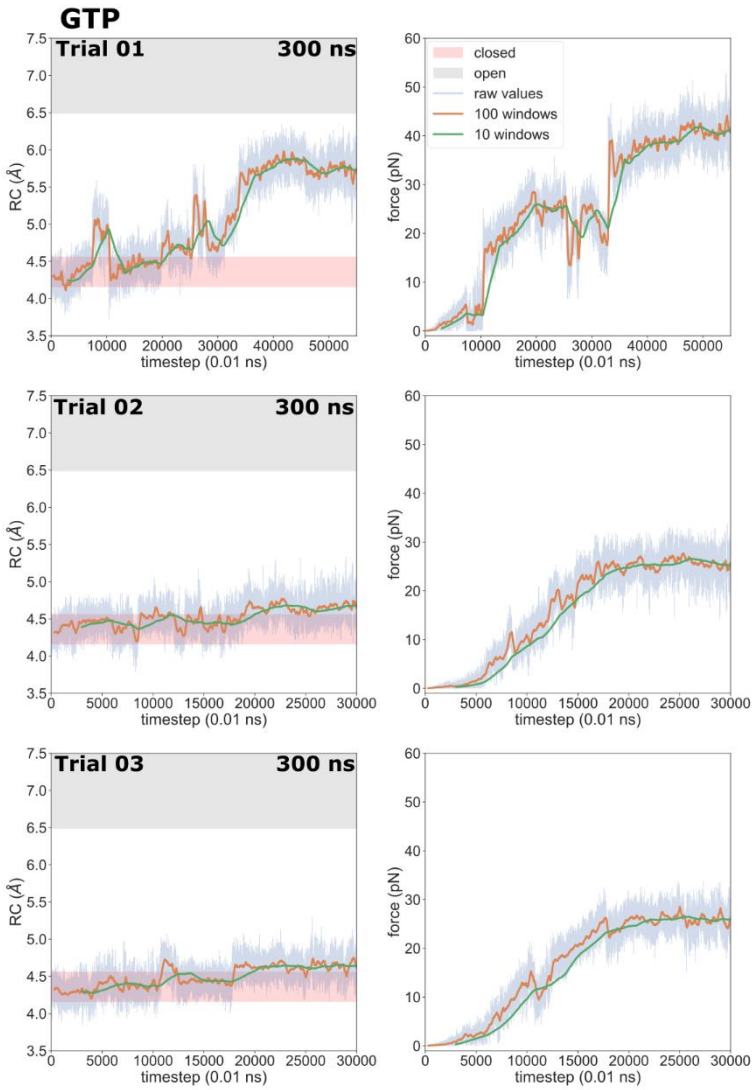

Figure S11. Results from steered MD (SMD) pulling GTP from the insertion (active-site closed) state well towards the initial binding (open) state at a rate of 1 Å/ns (force constant 2.4 $\frac{kcal}{mol\,Å^2}$). The reaction coordinate (RC) in pulling simulations, defined as the distance between the GTP and the active center (the center of mass of all αC within 10 Å of the 3' RNA primer), is shown on the left panels. Instantaneous force applied in the SMD simulations is shown on the right panels. Raw data values, 100 window smoothed curves, and 10 window smoothed curves are drawn. ±1 standard deviation of the RCs in the open and closed wells from umbrella sampling are included (as the gray and pink bars on the left panels). Trial simulation 01 was run to a total of 550 ns. The other two trials 02 and 03 were run to a total of 300 ns.

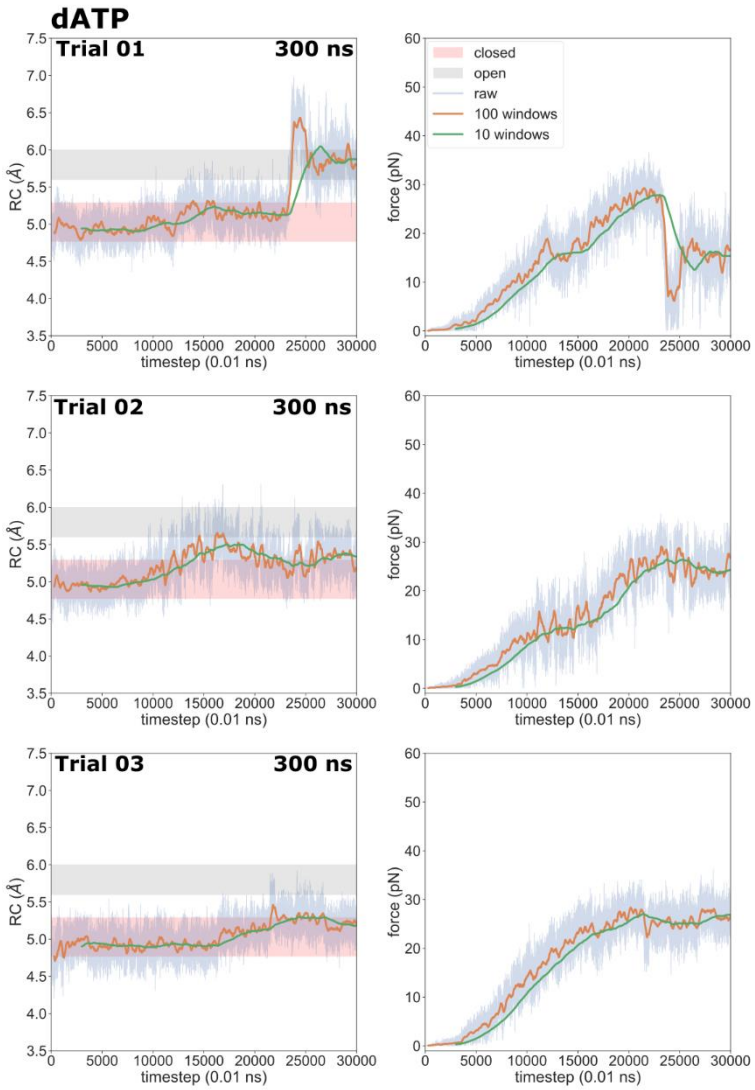

Figure S12. Results from steered MD (SMD) pulling dATP from the insertion (active-site closed) state well towards the initial binding (open) state at a rate of 1 Å/ns (force constant 2.4 $\frac{kcal}{mol\,Å^2}$). The reaction coordinate (RC) in pulling simulations, defined as the distance between the dATP and the active center (the center of mass of all αC within 10 Å of the 3' RNA primer), is shown on the left panels. Instantaneous force applied in the SMD simulations is shown on the right panels. Raw data values, 100 window smoothed curves, and 10 window smoothed curves are drawn. ±1 standard deviation of the RCs in the open and closed wells from umbrella sampling are included (as the gray and pink bars on the left panels). All trial simulations were run to a total of 300 ns.

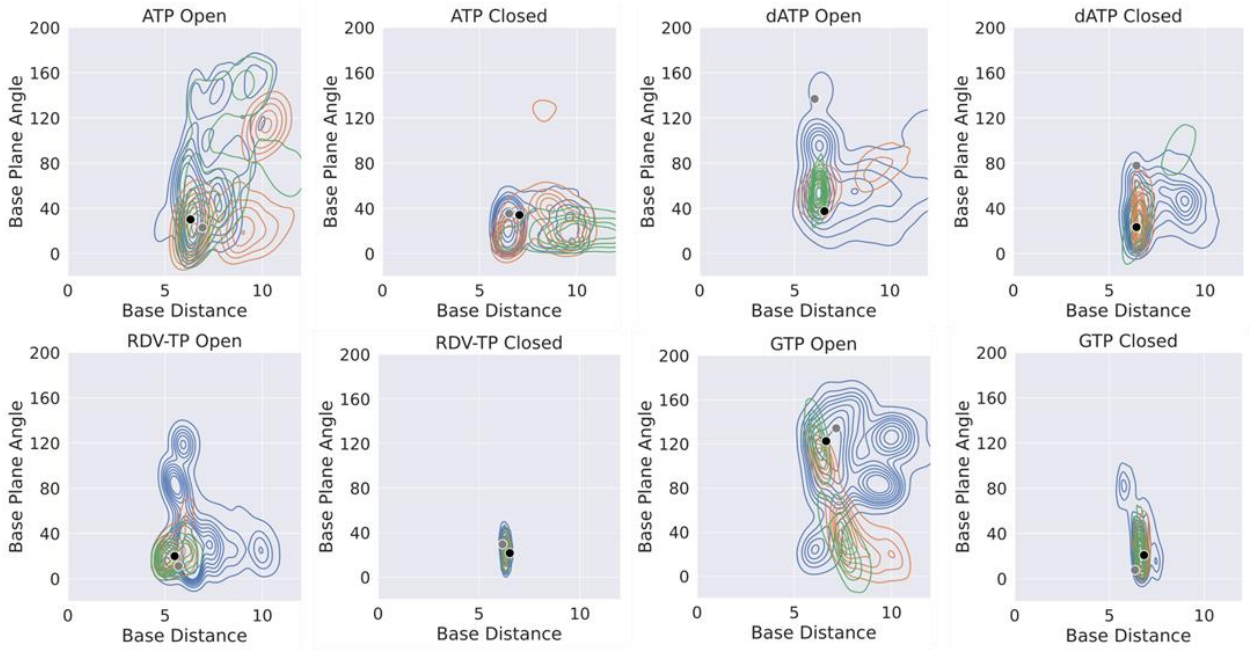

Figure S13. The NTP-template association geometry distributions obtained from the umbrella sampling simulations (for PMF calculations) in comparison with that from ensemble equilibrium simulations for various NTP species. A kernel density estimate has been used to visualize the data. Each simulation system (*open* to *closed*, for ATP, RDV-TP, dATP and GTP as in main **Figure 2** and **Figure 4**), the equilibrium ensemble distribution is shown (blue) along with that obtained from the umbrella sampling (w/ force on template in orange; w/out force on template in green). The black dot indicates the reference state used to generate the initial paths for the umbrella sampling, and grey dot the reference state used in the alchemical calculations from previous study [7].

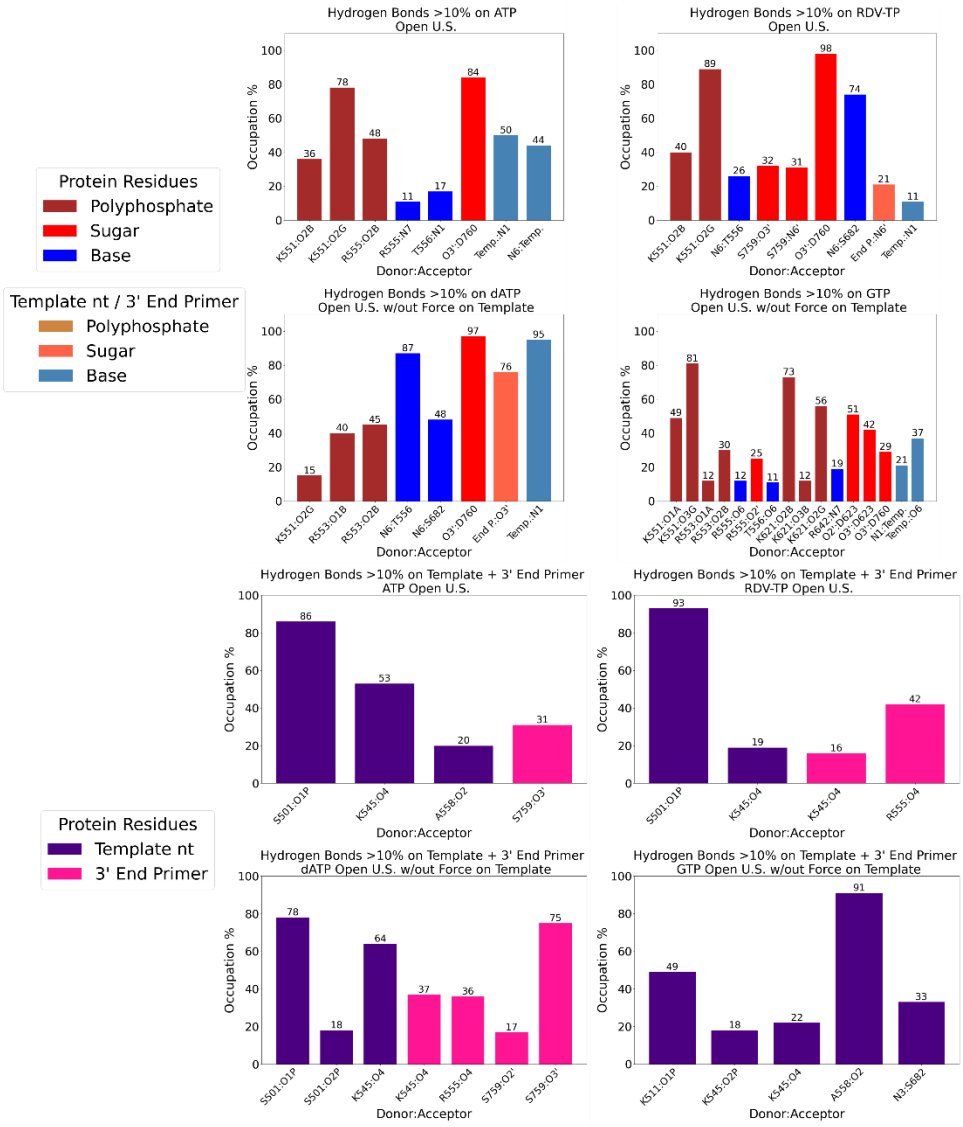

Figure S14. The hydrogen bond (HB) occupancy in umbrella sampling trajectories representing the initial binding (open) state minima for each NTP simulation system. The upper panel shows the HBs for each NTP from protein/template +1 nt/3'-end primer. The lower panel shows the protein HBs on the template +1 nt/3'-end primer. For each NTP (ATP, RDV-TP, dATP, and GTP), only unique HB interactions with a population greater than 10% are considered.

Table S4. Hydrogen bond interactions stabilizing/trapping cognate ATP, cognate RDV-TP, non-cognate dATP, and non-cognate GTP upon initial binding.

Initial binding hydrogen bonding (HB) interactions	**ATP**	**dATP**	**RDV-TP**	**GTP**
Template uracil	With ATP base two HBs; with motif-F K545 and motif-G S501	With dATP base one single HB; motif F K545 and motif G S501	Stacking with RDV-TP base; motif F K545 and motif G S501	Motif F K545&A558 on template base; Motif G K511 on template backbone
NTP-base	With template two HBs each ~50% occupancy	One HB with template nt+1; motif F T556 & K545	Motif B S682; stacking with template nt+1	
NTP-sugar	With motif-C D760	3'-OH with motif C D760; with 3'-end primer		With motif-A D623
3'-end primer	Transient HB with motif C S759	With dATP-sugar; motif C S759; motif F K545/R555		
NTP-phosphate	motif F (K551/R553/R555)		motif F K551	Motif F K551/R553 salt-bridge; Motif A D623, K621

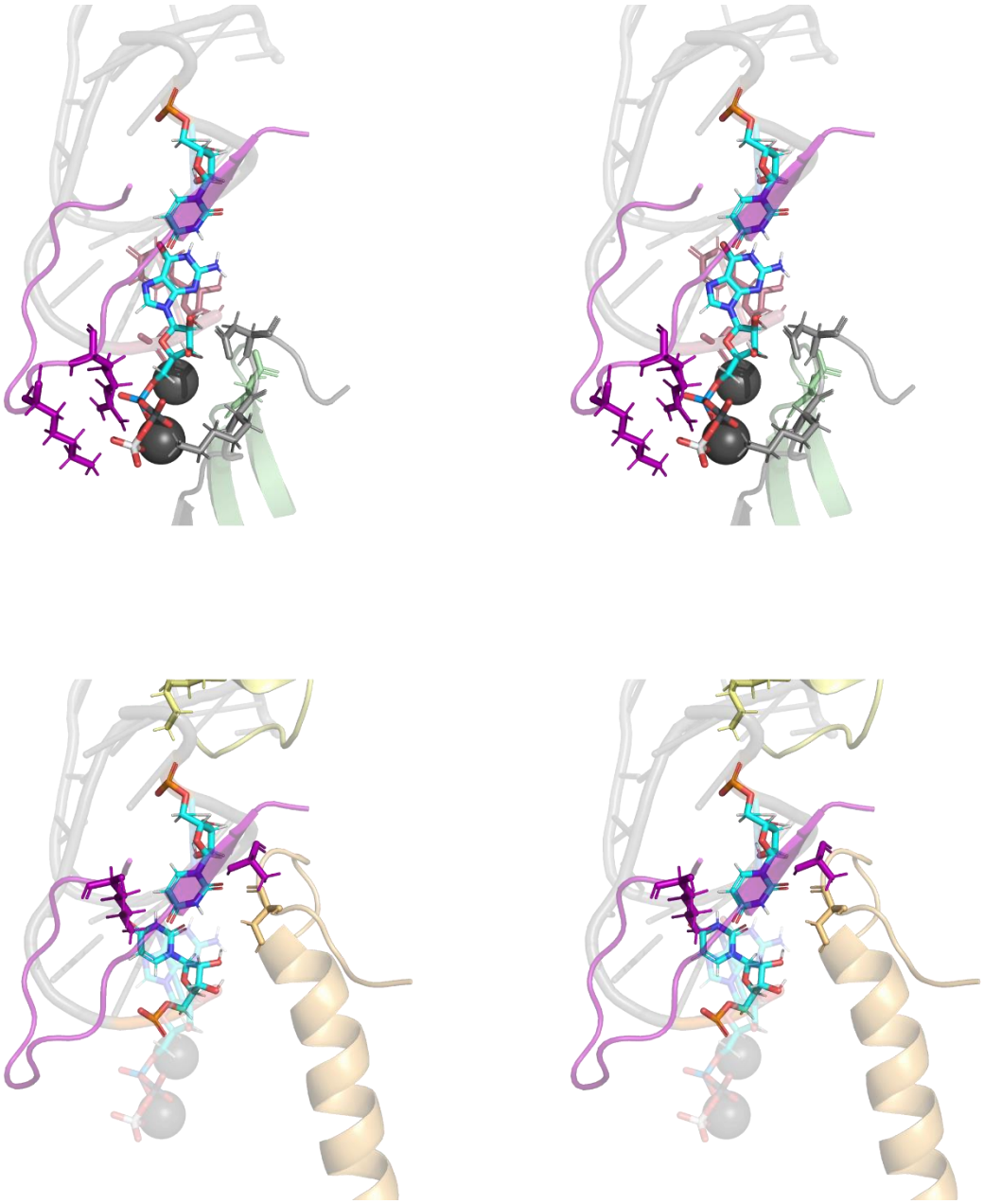

Figure S15. Stereoscopic views of main **Figure 6A** (top) and **Figure 6B** (bottom). GTP, 3’ RNA primer, template, motifs A, B, C, F and G are shown. The color scheme is the same as **Figure 6**.

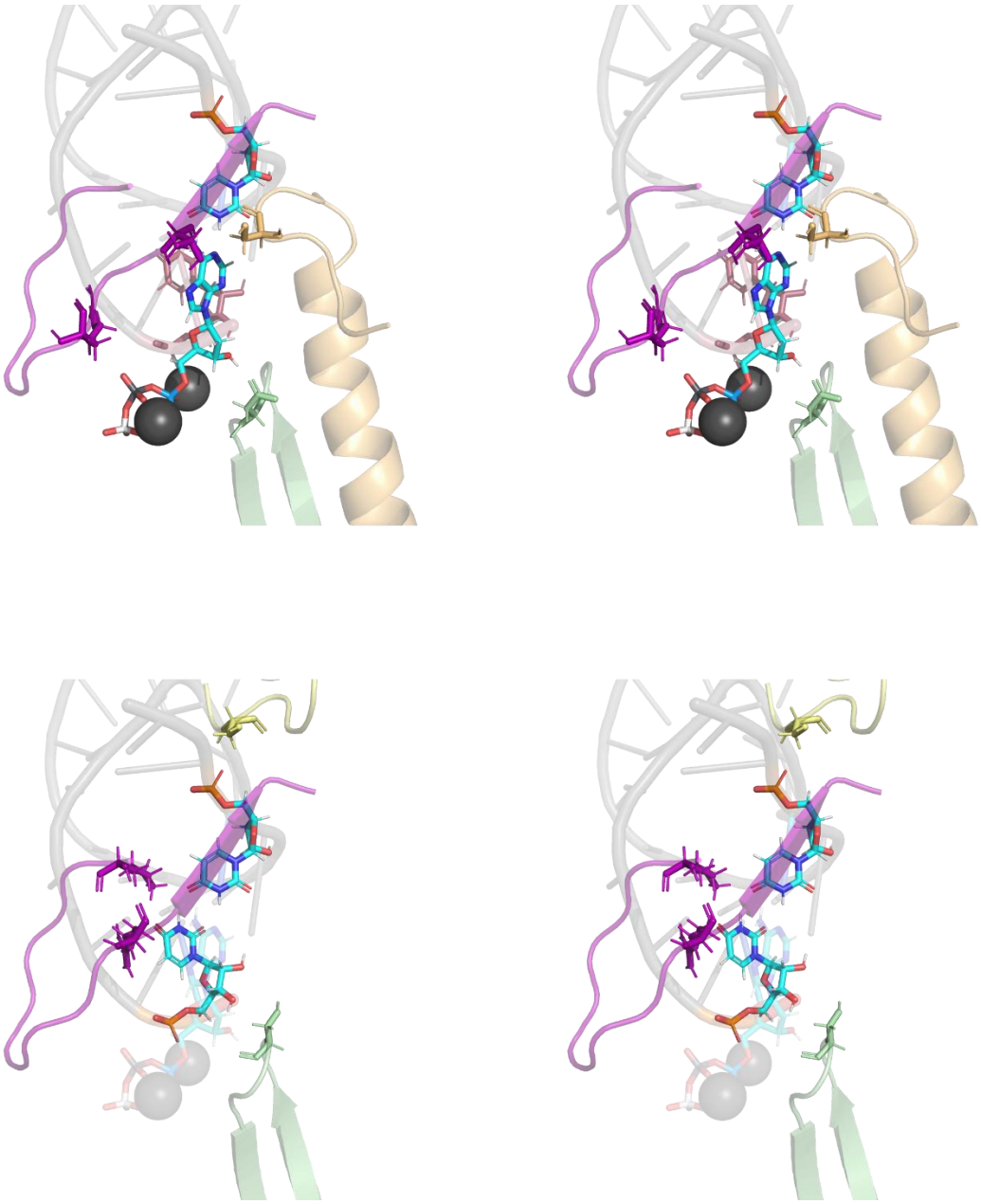

Figure S16. Stereoscopic views of main **Figure 7A** (top) and **Figure 7B** (bottom). dATP, 3’ RNA primer, template, motifs B, C, F and G are shown. The color scheme is the same as **Figure 7**.

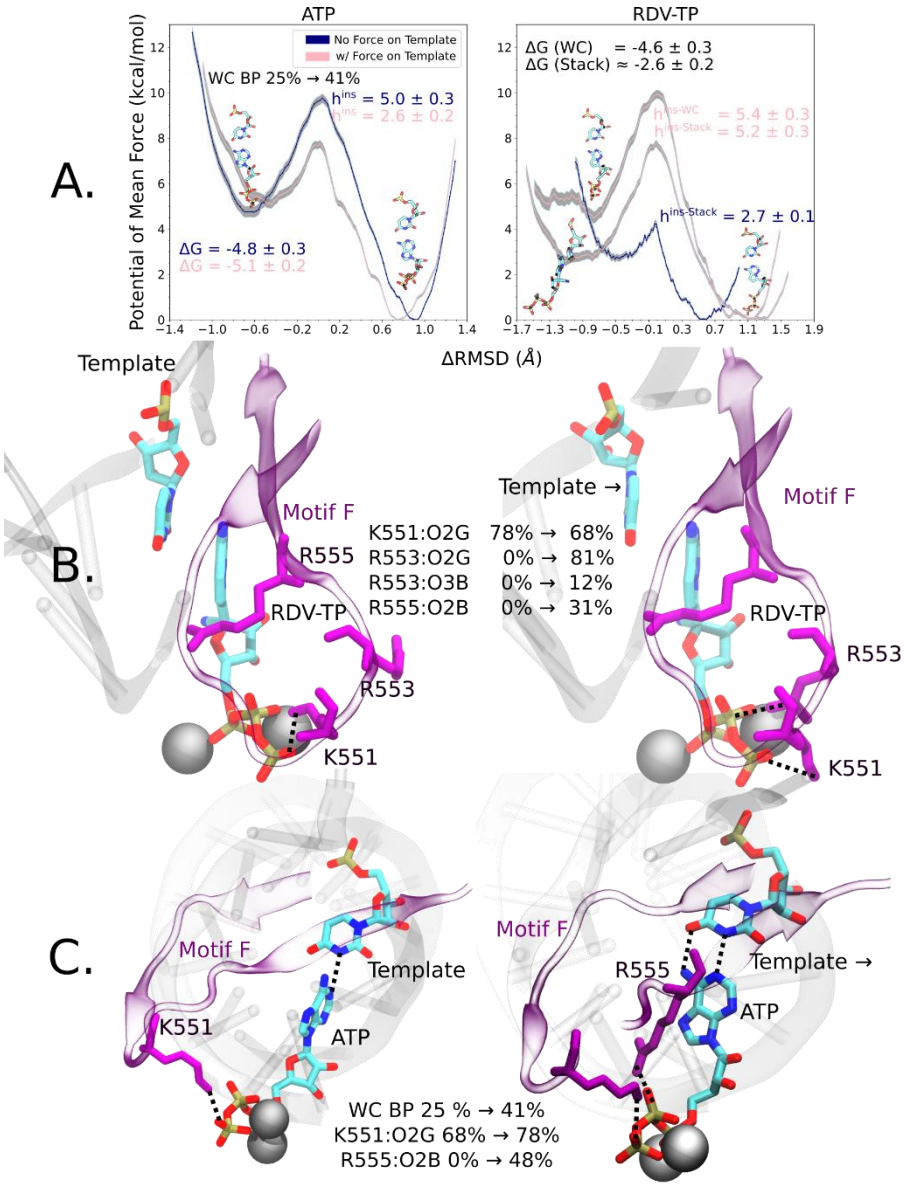

Figure S17. A summary of previous work on the insertion PMFs of ATP and RDV-TP[5]. *A:* The PMFs of insertion demonstrate that the free energy barrier (h^{ins}) can vary depending on whether there is force implemented on the template +1 nt: w/force, insertion of ATP/RDV-TP results in a lowered/increased barrier comparing with the case w/o force. Labelled energy values are reported in (kcal/mol). *B:* It is shown that starting from the base stacking configuration between RDV-TP and template +1 nt w/o force (*left*), the interactions between the motif F residues (K551, R553, R555) and the polyphosphate are destabilized (i.e., to lower the RDV-TP insertion barrier) comparing to w/force (*right*). *C:* It is shown that implementing the force on the template stabilizes the Watson-Crick base pairing between ATP and the template and enhances the interactions between the motif F residues (K551 and R555) and the polyphosphate, facilitating the insertion (lower the ATP insertion barrier).

Table S5. Sequence alignment of total RdRp and key motifs for SARS-CoV-2 compared to corresponding RdRps from Poliovirus (PV), Enterovirus-71 (EV), and Hepatitis C (HCV). Percent sequence alignment is reported for total as calculated with UniProt.[8] Motif sequences with bold overlap are also reported from literature[9,10,11] (PV and HCV) or from UniProt sequence alignment of motif ±10 residues of PV (EV).

SARS-CoV-2	PV	EV	HCV
7BV2[1]	3OLB[9]	6KWQ[12]	4WTA[13]
Total Sequence Alignment	**15.30%**	**16.30%**	**20.30%**
Motif A PHLMGWDYPKCDRAM	EEK**L**FAF**DY**TGY**D**ASL	PGS**L**FAF**DY**SGY**D**ASL	-- P**MG**FS**Y**DTRHF**D**STV
Motif B GGTSSGDATTAYANSV FNICQAVTANVNALLS	**GG**MP**SG**CSG**T**SIF**NS**M I**N**NLIIR**T**LLLKTYKG	**GG**MP**SG**CSG**T**SIF**N** **S**M I**N**NIIIRALLIKTFKG	RCRA**SG**VL**TT**SCG **N**TL TCYLK**A**SA**A**CRA**A** KLQ
Motif C FSMMILSDDAVVCFN	LK**M**IAYG**DD**VIASYP H	LN**M**VAYG**DD**VLAS YPF	CT**M**LVNG**DD**L**VV**I CE
Motif D ASQGLVASIKNFSV L YYQNNVFMSE	VDASLLAQSGKDYG**L** TMTPA	IDCLELAKTGKEYG **L** TMTPA	TQEDAASLRVFTEA MTRYSAPPGDPP
Motif E HEFCQHTMLV	VT**F**LKRFFRAD	AT**F**LKRGFLPD	ITSCSSNVSVA
Motif F LKYAIS-----AKN- RAR	V**K**DELRKTKVEQGKS **R**LI	**L**P**Y**STUV**K**DEL**R**SI	A**K**NEVPEKGGR- KPA**R**LI
Motif G DKSAGFPFNKWGK ARL	ALDLST**SAG**Y	ALDLHT**SAG**Y	LTPPH**SA**KSK

1. Wanchao Yin *et al.*, Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **368**, 1499-1504 (2020).
2. Hillen, H.S., Kokic, G., Farnung, L. *et al.* Structure of replicating SARS-CoV-2 polymerase. *Nature* **584**, 154–156 (2020).
3. Yan Gao et al., Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **368**, 779-782 (2020).
4. Malone, B.F., Perry, J.K., Olinares, P.D.B. *et al.* Structural basis for substrate selection by the SARS-CoV-2 replicase. *Nature* **614**, 781–787 (2023).
5. Romero, M. E. *et al.* Probing remdesivir nucleotide analogue insertion to SARS-CoV-2 RNA dependent RNA polymerase in viral replication. *Mol Syst Des Eng* **6**, (2021).
6. T. A. Steitz & J. A. Steitz, A general two-metal-ion mechanism for catalytic RNA, *Proc. Natl. Acad. Sci.* **90**, 6498-6502 (1993).
7. Chunhong Long, Ernesto Romero, M., Liqiang Dai & Jin Yu. Energetic vs. entropic stabilization between a Remdesivir analogue and cognate ATP upon binding and insertion into the active site of SARS-CoV-2 RNA dependent RNA polymerase. *Physical Chemistry Chemical Physics* **25**, 13508–13520 (2023).
8. The UniProt Consortium, UniProt: the Universal Protein Knowledgebase in 2023, *Nucleic Acids Research*, **51**, 523-531 (2023).
9. Gong, P., Peersen, O. B., Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci.* **107**, 22505-22510 (2010).
10. Wang, M., Li, R., Shu, B. *et al.* Stringent control of the RNA-dependent RNA polymerase translocation revealed by multiple intermediate structures. *Nat. Commun.* **11**, 2605 (2020).
11. Sesmero, E. & Thorpe, I., Using the Hepatitis C Virus RNA-Dependent RNA Polymerase as a Model to Understand Viral Polymerase Structure, Function and Dynamics. *Viruses*. **7**, 3974-3994 (2015).
12. Wei Shi, Han-Qing Ye, Cheng-Lin Deng, Rui Li, Bo Zhang, Peng Gong, A nucleobase-binding pocket in a viral RNA-dependent RNA polymerase contributes to elongation complex stability, *Nucleic Acids Research*, **48**, 1392-1405 (2020).
13. Todd C. Appleby *et al.*, Structural basis for RNA replication by the hepatitis C virus polymerase. *Science* **347**, 771-775 (2015).