Title

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

Journal

Physical Chemistry Chemical Physics, 25(19)

ISSN

0956-5000

Authors

Long, Chunhong
Romero, Moises Ernesto
Dai, Liqiang
et al.

Publication Date

2023-05-17

DOI

10.1039/d2cp05883a

PCCP
Physical Chemistry Chemical Physics

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: C. Long, M. E. Romero, L. Dai and J. Yu, *Phys. Chem. Chem. Phys.*, 2023, DOI: 10.1039/D2CP05883A.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/pccp

View Article Online
DOI: 10.1039/D2CP05883A

Energetic vs entropic stabilization between Remdesivir analogue and cognate ATP upon binding and insertion into active site of SARS-CoV-2 RNA dependent RNA polymerase

Chunhong Long,[a] Moises Ernesto Romero,[b] Liqiang Dai [c] and Jin Yu[*d,b]

The SARS-CoV-2 RNA dependent RNA polymerase (RdRp) serves as a highly promising antiviral drug target for Remdesivir analogue (RDV-TP or RTP). In this work, we mainly employed alchemical all-atom simulations to characterize relative binding free energetics between nucleotide analogue RTP and natural cognate substrate ATP upon initial binding and pre-catalytic insertion to the active site of SARS-CoV-2 RdRp. Natural non-cognate substrate dATP and mismatched GTP were also examined for computation control. We first identified significant differences of dynamical responses between nucleotide initial binding and subsequent insertion configurations to the open and closed active site of the RdRp, respectively, though the RdRp protein conformational changes between the active site open and closed states are subtle. Our alchemical simulations indicated that upon initial binding (active site open), RTP and ATP show similar binding free energies to the active site while in the insertion state (active site closed), ATP is more stabilized (~ -2.4 kcal/mol) than RTP in free energetics. Additional analyses show, however, that the RTP is more stabilized in binding energetics than ATP, in both the insertion and initial binding states, with RTP more stabilized due to electrostatic energy in the insertion state and due to vdW energy in initial binding state. Hence, it appears that natural cognate ATP still excels at association stability with the RdRp active site due to that ATP maintains sufficient flexibilities e.g., in base pairing with the template, which exemplifies entropic contribution to the cognate substrate stabilization. The findings highlight the importance of substrate flexibilities in addition to energetic stabilization in antiviral nucleotide analogue design.

Introduction

The 2019 novel coronavirus (COVID-19 coronavirus (CoV) or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)) has spread rapidly to cause serious outbreaks and eventually a global pandemic. COVID-19 has caused a global health crisis then for the past three years. Up to now, more than 596 million cases have been diagnosed worldwide [1]. Antiviral agents are urgently needed to treat COVID-19 patients. Viral RNA-dependent-RNA polymerase (RdRp) is a core protein of the viral replication machinery that is buried inside the viral capsule and is functionally conserved, rendering it resistant to mutations [2]. Due to its critical role in the viral RNA synthesis and highly conserved core structure, the viral RdRp serves as a highly promising antiviral drug target for both nucleotide analogue and non-nucleoside inhibitors [3, 4]. Remdesivir (or RDV), the first US-FDA proved drug (named VEKLURY) for treating COVID-19 [5], works as a prodrug that is metabolized into a nucleotide analogue (RDV-TP or RTP) to compete with natural nucleotide substrates of RdRp to be incorporated into the viral RNA gnome, and to further terminate the RNA synthesis [6-9]. Such a nucleotide analogue thus well surpasses nucleotide selectivity of the RdRp. The nucleotide selectivity of RdRp serves as a primary fidelity control method in viral genome replication and transcription [10-12]. The selectivity can be regulated throughout the nucleotide addition cycle (NAC) from nucleotide initial binding to insertion into the active site, then catalysis, and possibly extended to the post-catalytic steps, via stepwise kinetic checkpoints [11-13]. As RTP is capable of incorporation to the viral RNA, it must pass through all the kinetic checkpoints in the NAC [14, 15].

Notably, in the single subunit viral RNAP, such as the T7 RNAP adopting a hand-like structure, the nucleotide insertion often takes place slowly or becomes rate limiting in the NAC [16, 17], accompanied by substantial conformational changes from an open form to closed [18, 19]. The rate-limiting nucleotide insertion step thus plays an important role for nucleotide selectivity or fidelity control of polymerase [20]. The open (NTP initial binding or pre-insertion) and closed (substrate insertion) structures of T7 RNAP had been captured previously[21, 22], which are characterized by notable rotational movements of the flexible fingers subdomain relative to the stabilized palm subdomain. Seven highly conserved structural motifs are located within these two subdomains. For the high-resolution cryo-EM structures of SARS-CoV-2 RdRp or nsp12 (along with accessory factors nsp7 and nsp8), there are still seven conserved structural motifs A-G located in the palm and fingers subdomains [23, 24] (see **Figure 1A**). The conformational changes between the open and closed state SARS-CoV2 RdRp aligned with fingers domain

[a] *School of Science, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China*
[b] *Department of Chemistry, University of California, Irvine, CA 92697 USA*
[c] *Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou,325001,China*
[d] *Department of Physics and Astronomy, NSF-Simons Center for Multiscale Cell Fate Research, University of California, Irvine, CA 92697, USA*
*Corresponding author
Email: jin.yu@uci.edu

View Article Online
DOI: 10.1039/D2CP05883A

become quite subtle (RMSD~1.7 Å for all C_α atoms) while for T7 RNAPs aligned with palm subdomain, the RMSD is ~2.9 Å.

For RdRps such as that from the poliovirus (PV), open and closed forms had been suggested [25]. When there is no nucleotide substrate bound to the active site, RdRp adopts an open conformation. As the nucleotide substrate binds and inserts into the active site, the RdRp active site transits to a closed conformation which is ready for the following catalysis step. In the PV RdRp with an encircled active site, the incoming nucleotide forms base-pairing with the template nucleotide at +1 and stacking with the RNA3' priming nucleotide upon initial binding to the active site, but the triphosphate moiety of the incoming nucleotide is not fully loaded while the ribose and sugar part are loaded yet. The closure of the active site is achieved indeed by the ribose hydroxyl recognition after the NTP-template base-pairing along with the structural changes mainly in the palm domain including motif A-E (residue 69 to 96, 192 to 269, and 286 to 381) [25](see **Figure 1B**).

Recently, both the open and closed forms of the SARS-CoV-2 RdRp structures have also been captured from high-resolution cryoEM studies [23, 24, 26, 27]. For example, an open form of the active site reveals from an apo structure of the RdRp (nsp12), in complex with cofactor nsp7 and nsp8, but in the absence of RNA strands and the nucleotide substrate [23]. Meanwhile, a closed form structure of the RdRp appears e.g., with an RDV analogue just incorporated post-catalytically to the 3'-end of the primer RNA chain, having magnesium ions remained bound at the active site [24]. Based on the two resolved structures (PDB:7BTF[23] and 7BV2[24]). we constructed respectively an open (substrate initial binding) and a closed (substrate insertion) conformation of the SARS-CoV-2 RdRp with RTP bound, pre-catalytically (see **Figure 1D**). Then, according to these two structures, we also modeled the open and closed state conformations with the natural cognate nucleotide ATP bound pre-catalytically (see **Figure 1C**), and then with the non-cognate nucleotide dATP or GTP bound (see **Figure 1E-F**), replacing ATP.

In order to understand how RTP binds and inserts to the active site of SARS-CoV-2 RdRp or evade the nucleotide selectivity in comparison with the natural cognate ATP, a recent all-atom molecular dynamics (MD) study has been conducted employing the umbrella sampling method on determining the nucleotide insertion free energy profile or potential of mean force from initial binding to insertion, prechemically, connecting the open and closed states of RdRp [28]. It has been demonstrated that the free energy barrier of ATP insertion is as low as ~ 2.6 kcal/mol, while for RTP, the insertion barrier can even drop to ~1.5 kcal/mol. It was also found that RTP can initially bind around the active site primarily via base stacking with the template nucleotide rUTP at +1. In contrast, ATP binds initially to the active site still in the Watson–Crick base pairing with the template rUTP. Additionally, it was shown that the closed insertion state of the bound nucleotide, no matter for ATP or RTP association, is much more stabilized than the open state in the SARS-CoV-2 RdRp (by ~ -5.1 kcal/mol and -2.7 kcal/mol) [28]. Nevertheless, even though the two free energy profiles or PMFs were obtained for the ATP and RTP insertion processes, respectively, it was not clear whether ATP or RTP is more stabilized upon initial binding (open state) or in the insertion (closed state). That is to say, the relative binding free energetics between RTP and ATP in certain binding or insertion state of RdRp remained unknown. Hence, in current work, we want to determine the relative binding free energy between RTP and ATP, in both the open and closed states of RdRp, which correspond to nucleotide initial binding and insertion equilibrium, respectively.

Accordingly, in current studies, we employed mainly alchemical MD simulations to demonstrate relative stabilities or calculate the relative binding free energies between RTP and ATP, for both the initial binding (open) and stably insertion (closed) states in the SASR-CoV-2 RdRp. In addition, we included natural non-cognate substrates dATP and GTP as controls for our calculations, to also calculate their relative binding free energetics with respect to the cognate ATP, i.e., upon modeling them in the initial binding and insertion states as well. We found notable energetic stabilization of RTP in the insertion state to the RdRp active site, while marginal flexibilities of ATP upon insertion to the active site seem to bring entropic advantages to this natural substrate, such that the binding free energetics of ATP still turns out to be lower than that of the its analog RTP.

View Article Online
DOI: 10.1039/D2CP05883A

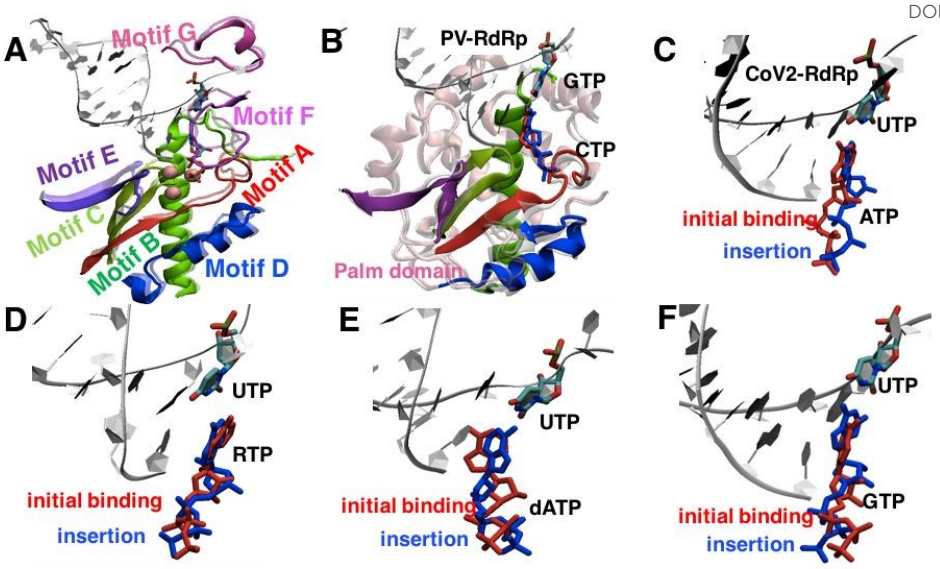

Figure 1. The core structures around the active site of the RNA dependent RNA polymerase (RdRp) from SARS-CoV-2 and Polio virus (PV) in complex with an initial binding and an insertion NTP. (A) The open (non-transparent colored) and closed (transparent) RdRp structures aligned according to RNA construct upstream, with motif A-G shown and colored, respectively, and magnesium ions also shown in pink (see **SI Figure S1** for further detail). (B) The initial binding (open) and insertion (closed) structures of CTP bound PV RdRp (PDB:3ola and 3ol7) [25], with RMSD between the initial binding and insertion CTP 1.05 Å. (C-F) The equilibrated structures of modeled initial binding (open) and insertion (closed) structures of various NTP bound with SARS-CoV-2 RdRp, for (C) cognate ATP (RMSD: 4.37 Å), (D) RTP analog (RMSD: 2.52 Å), and non-cognate (E) dATP (RMSD: 4.57 Å) and (F) mismatch GTP (RMSD: 4.74 Å).

Methods

1. Building open/closed structures of SARS-CoV2 for cognate (ATP) and non-cognate (RTP/dATP/GTP) species.

The high-resolution elongation complexes of SARS-CoV-2 viral RNA dependent RNA polymerase (RdRp) are available in post-catalysis state with RDV-MP (PDB:7BV2). RTP and the natural nucleotide substrate ATP were fitted into the active site where an incorporated RDV-MP occupies in the complex, so that to obtain a pre-catalytic closed insertion complexes for RTP and ATP, respectively. The open-state elongation complex with an initially bound RTP (or ATP) was obtained from the apo nsp12 structure (PDB:7BTF) along with additionally incorporated RNA strands and then RTP (or ATP), by fitting the apo structure with the above constructed RdRp closed complex. Then for both open and closed complex structures, the cognate ATP was mutated into dATP or GTP to generate the corresponding initial binding and insertion structures. As a result, there are four complex structures, i.e., for ATP, RTP, dATP and GTP, for both an open initial-binding and a closed insertion state of the SARS-CoV-2 RdRp elongation complex, or 8 structures in total (see **Figure 1C-F**). Before calculating the alchemical free energy, each structure was energy minimized and conducted with 100-ns equilibrium MD simulations.

Note that in the open or initial binding state of the RTP, a base-stacking configuration of RTP with respect to the template (rUTP) was identified recently [28], which appears to be more stabilized than the base-pairing configuration of RTP-rUTP. Accordingly, we used the base-stacking configuration of RTP in the open initial binding complex structure of RTP (see **Figure 1D**).

2. Calculating the alchemical free energy by using the Free Energy Perturbation (FEP) method.

The relative binding free energy between nucleotide RTP (or dATP/GTP) and the cognate nucleotide ATP at the open or initial binding state as well as at the closed or insertion state can be evaluated through the thermodynamic cycle [29-32] (see **Figure 2**): $\Delta\Delta G_b \equiv \Delta G_b(RTP) - \Delta G_b(ATP)$, where $\Delta G_b(RTP) = \Delta G_3 - \Delta G_1 + \Delta G_2$ and $\Delta G_b(ATP)$

View Article Online
DOI: 10.1039/D2CP05883A

$= \Delta G_3 - \Delta G_1 + \Delta G_2$, with $\Delta G_2 = 0$ due to that the dummy is set to be vanished in all cases [29]. The relative binding free energy calculations for dATP and GTP with respect to ATP are also using the same thermodynamic cycle as RTP. The alchemical free energy was calculated using the free energy perturbation (FEP) method [33]. The free energy change of transforming a system from state A (dummy or void) to state B (a full ligand) is a function of a coupling or perturbation parameter λ $(0 \leq \lambda \leq 1)$, which indicates the level of change between state A and state B. The extent to which the Hamiltonian or energy function has been perturbed indicates how the system has been transformed. Simulations conducted at different values of λ thus allow us to get ΔG_{AB} . Note that the simulations were then conducted in free solution (for ΔG_1) and inside the protein complex (for ΔG_3), respectively. Consequently, the transformation free energy ΔG in each of the environments was calculated with the following equations:

$$\Delta G_\lambda = -\frac{1}{\beta} In\langle \exp\{-\beta[H(\lambda + \Delta\lambda) - H(\lambda)]\}\rangle \qquad (1)$$

$$\Delta G = \sum_\lambda \Delta G_\lambda \qquad (2)$$

where $\frac{1}{\beta} = k_B T$, and $H(\lambda) = (1-\lambda)H_1 + \lambda H_2$, as H_1denotes the potential energy of the initial state and H_2 denotes the potential energy of the final state. The coupling parameter λ value from 0 to 1 is set with an increment of 0.05.

Note that the electrostatic and vdW interaction were simultaneously changed during the simulation.

In the free-solution simulations, NTP was solvated with ~4020 TIP3P waters in a cubic box with a size of ~50 Å; to neutralize the system and make the salt concentration 0.1M, 12 Na^+ and 8 Cl^- were added. There were totally ~12,060 atoms in the final system. The simulations in the protein complex were solvated with ~ 183,800 TIP3P waters in a cubic box with a size of ~180 Å. The minimum distance from the protein complex to the wall of simulation box was 18 Å. The protein complex simulations are specified as in the next subsection.

For both the free-solution and protein bound simulations, we followed the same procedures for each window: (i) 50,000 steps of energy minimization with a steepest decent algorithm; (ii) 100-ps MD simulations within the canonical ensemble; (iii) 100-ps NPT equilibration with atomic position restraints on the heavy atoms of the protein and nucleic acids; (iv) the productive runs with the restraints removed and within the NPT ensemble, carried out for 5 ns in the protein and free-solution systems. There were totally 21 windows for one complete alchemical simulation; each window was supplied to 5-ns FEP runs. Hence, $0.84\ \mu s$ free-solution simulation data and $0.84\ \mu s$ protein bound simulation data were collected in total to evaluate the relative binding free energy between RTP/dATP/GTP and cognate ATP, for both open state and closed state CoV-2 RdRp complex, with each NTP species 105-ns alchemical simulation.

For ATP, we have supplied each window to 10-ns FEP run, and we found that the binding free energy values calculated between the 5-ns and 10-ns FEP runs are almost the same, i.e., with $\Delta G_b(ATP)$ obtained ~ -11.42kcal/mol for 5-ns FEP runs and -11.65kcal/mol for 10-ns FEP runs, only about 0.23 kcal/mol difference. In addition, we also measured the RMSD value of fingers subdomain in CoV-2 RdRp during 0-5 ns and 5-10 ns MD simulations (see **SI Figure S2A&B**). Similarly, the RMSD value of the palm subdomain in CoV-2 RdRp during 0-5 ns and 5-10 ns MD simulations were obtained (**SI Figure S2C&D**). The 5-ns and 10-ns results are largely consistent. Hence, we used 5-ns MD simulations per window for all other NTP alchemical simulations.

View Article Online
DOI: 10.1039/D2CP05883A

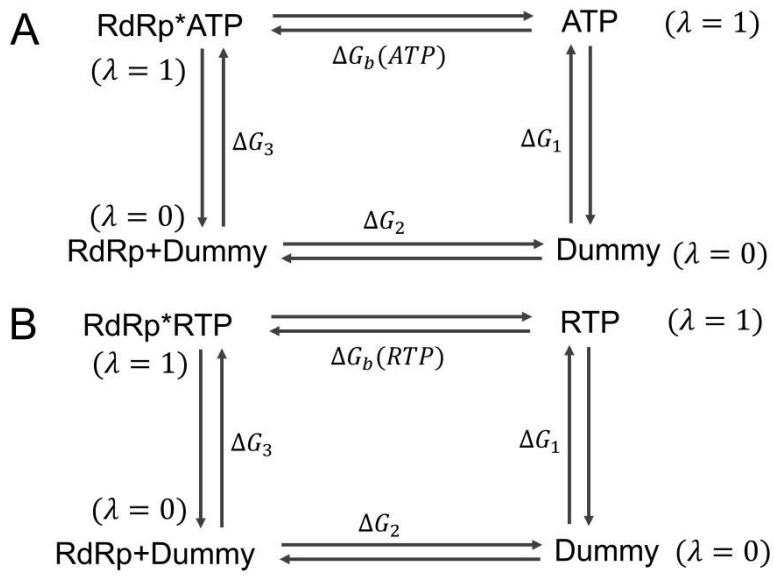

Figure 2. The computational scheme to calculate the relative binding free energy between RTP (or non-cognate nucleotide species dATP/GTP) and the cognate ATP in association with RdRp. (A) The thermodynamic cycle is used to calculate the binding free energy $\Delta G_b(ATP) = \Delta G_3 - \Delta G_1 + \Delta G_2$ of the cognate ATP. (B) The thermodynamic cycle is used to calculate the binding free energy $\Delta G_b(RTP) = \Delta G_3 - \Delta G_1 + \Delta G_2$ of RTP (or non-cognate dATP/GTP). The relative binding free energy is thus calculated by using: $\Delta\Delta G_{binding} = \Delta G_b(RTP) - \Delta G_b(ATP)$, with $\Delta G_2 = 0$ as the dummy is set as void for each species.

3. Setup of the equilibrium MD simulation.

All MD simulations were performed using GROMACS2019 package[34], with the Amber14sb protein force field [35] and Parmbsc1 nucleic acid force field [36] used. For the NTPs, triphosphate parameters calculated previously were used[37]. For the magnesium ions, amber14sb force field were still used. For equilibrium MD simulation, the RdRp complex was solvated with explicit TIP3P water[38] in a cubic box with a size of ~157Å, and the minimum distance from the complex to the wall was set to 15Å. To neutralize the system and make the salt concentration 0.1M, 387 sodium ions and 351 chloride ions were added. Three magnesium ions were kept from the cryo-EM structure [PDB: 7BV2 (from the early version1)][24]. The full equilibrium simulation systems contained ~382,000 atoms. For all simulations, the cut-off of van der Waals (vdW) and the short-range electrostatic interactions were set to 10Å. The particle-mesh-Ewald (PME) method was used to evaluate the long-range electrostatic interactions [39, 40]. All MD simulations were run at 1 bar and 310 K using the Parrinello-Rahman Barostat and the velocity rescaling thermostat [41-43], respectively. Before the equilibrium NPT simulation, we minimized the system for 50,000 steps using the steepest-descent algorithm followed by 2-ns NVT MD simulation. The time step was 2 fs and the neighbor list was updated every 10 fs. In total for each initial binding and insertion state, 10x100-ns equilibration simulations for each NTP species (ATP/RTP/dATP/GTP) were launched independently for a total of 8-μs of simulation time.

Results

We first present the alchemical and relative binding free energy calculations of RTP and non-cognate dATP/GTP with respect to the cognate ATP in the precatalytic insertion state (active site closed), and then in the initial binding state (active site open). We also show samplings on the nucleotide geometries such as the distance between insertion nucleotide (ATP/RTP/dATP/GTP) and template rUTP from the equilibrium ensemble MD simulations (10x100-ns for each NTP). We additionally show the hydrogen bonding (HB) patterns between key residues around the active site and the bound nucleotide (ATP/RTP/dATP/GTP), upon the nucleotide insertion (closed) and initial binding (open).

1. The relative binding free energy and energetic contributions in the nucleotide insertion state (active site closed)

We have calculated the relative binding free energy of the nucleotide analogue RTP as well as non-cognate nucleotide dATP/GTP with respect to cognate nucleotide ATP in the active site closed insertion state, and the results are shown below (see **Table 1**).

Nucleotide species	$\Delta\Delta G_{binding}\ (kcal/mol)$ Closed state	$\Delta\Delta G_{binding}\ (kcal/mol)$ Open state
RTP	2.36 ± 0.69	-0.23 ± 0.16
dATP	1.36 ± 0.14	1.51 ± 0.17
GTP	2.95 ± 0.66	1.70 ± 0.55

Table 1. The relative binding free energy for the bound RTP and non-cognate nucleotide species with respect to cognate ATP at the active site of SARS-CoV-2 RdRp in the closed state and open state, respectively. More details are provided in **SI Table S1 & S3.**

The relative binding free energy between RTP and ATP is ~2.4 kcal/mol, between non-cognate dATP and ATP is ~1.4 kcal/mol, and between the base mis-matched GTP and ATP is ~3 kcal/mol (more details can be found in **SI Table S1**). It is as expected that the mismatched GTP shows notable instability upon insertion, even with wobble pairing formed between GTP and template rUTP. In comparison, instability is smaller for the non-cognate dATP with respect to ATP than GTP, as dATP only has one oxygen atom removed for the deoxyribose sugar comparing to the cognate ATP. It was however not clear whether RTP would be more or less stabilized than ATP. The calculation here shows that RTP is still destabilized comparing to ATP in the insertion state.

Note that during the alchemical simulations on the protein complex, one can see that a growth of an NTP ligand in the active site would trigger deviations of the dihedrals of template +1 rUTP and the primer RNA 3'-end differently, for respective simulations with ATP/RTP/dATP/GTP (see **SI Figure S3).** Interestingly, one notices that in the closed active site or the NTP insertion state, the primer RNA3'-end often have larger dihedral deviations than the template +1 rUTP, mainly for the growing ligand of RTP and non-cognate GTP.

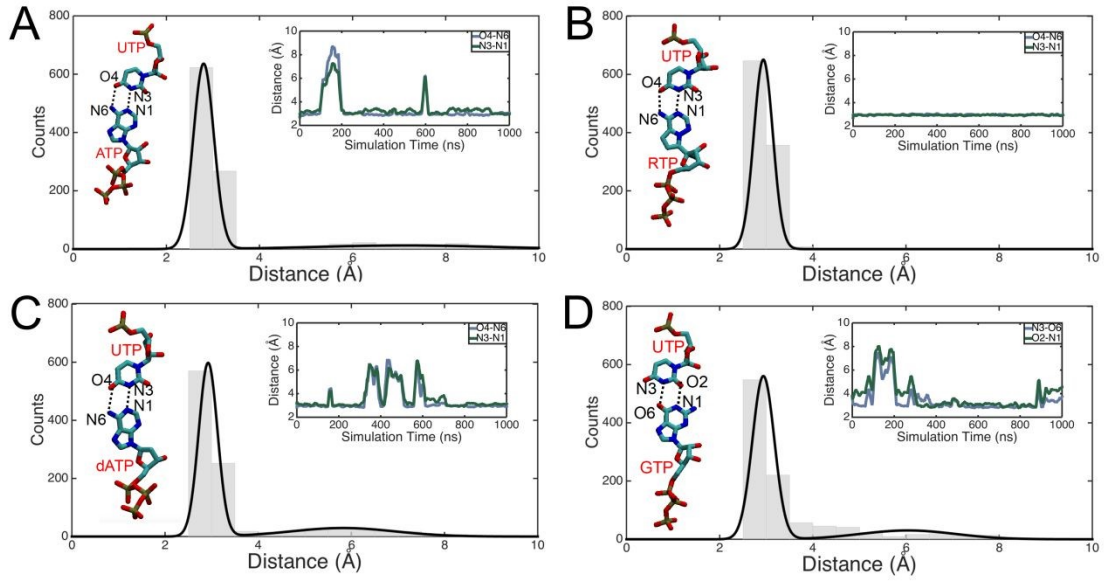

Figure 3. The histograms on measured distances between nucleotide ATP/RTP/dATP/GTP base and template nt (rUTP) base on atoms for base pairing. The base pairing percentiles are ~90.8%, ~99.2%, ~82.5%, 77.3% for (A) ATP, (B) RTP, (C) dATP, and (D) GTP, respectively. In the insets *left*, the structures of insertion nucleotide and the associating template rUTP are shown. In the insets *right*, the distances between atoms forming hydrogen bonding interactions (or base pairing) are shown over the simulation time.

From the equilibrium ensemble simulations, we examined the distance geometries between the inserted NTP and template rUTP, in particular, by measuring the atomic distances which would be involved in the NTP-rUTP base

View Article Online
DOI: 10.1039/D2CP05883A

pairing interactions when the distances are sufficiently small (see **Figure 3**). We noticed that the base pairing between the inserted RTP and the template rUTP are much more stabilized (~99.2% base pairing) than all the natural substrates, i.e., no matter the cognate ATP (~90.8%), or the non-cognate dATP/GTP (~82.5% or 77.3%). The higher stability demonstrated on RTP than ATP in base pairing with the template rUTP is not quite expected based on the relative binding free energy calculation above (in **Table 1**), i.e., as RTP is destabilized for > 2 kcal/mol than ATP in binding free energy upon insertion. Hence, one needs to identify and justify which interactions or components indeed stabilize ATP relative to RTP upon insertion.

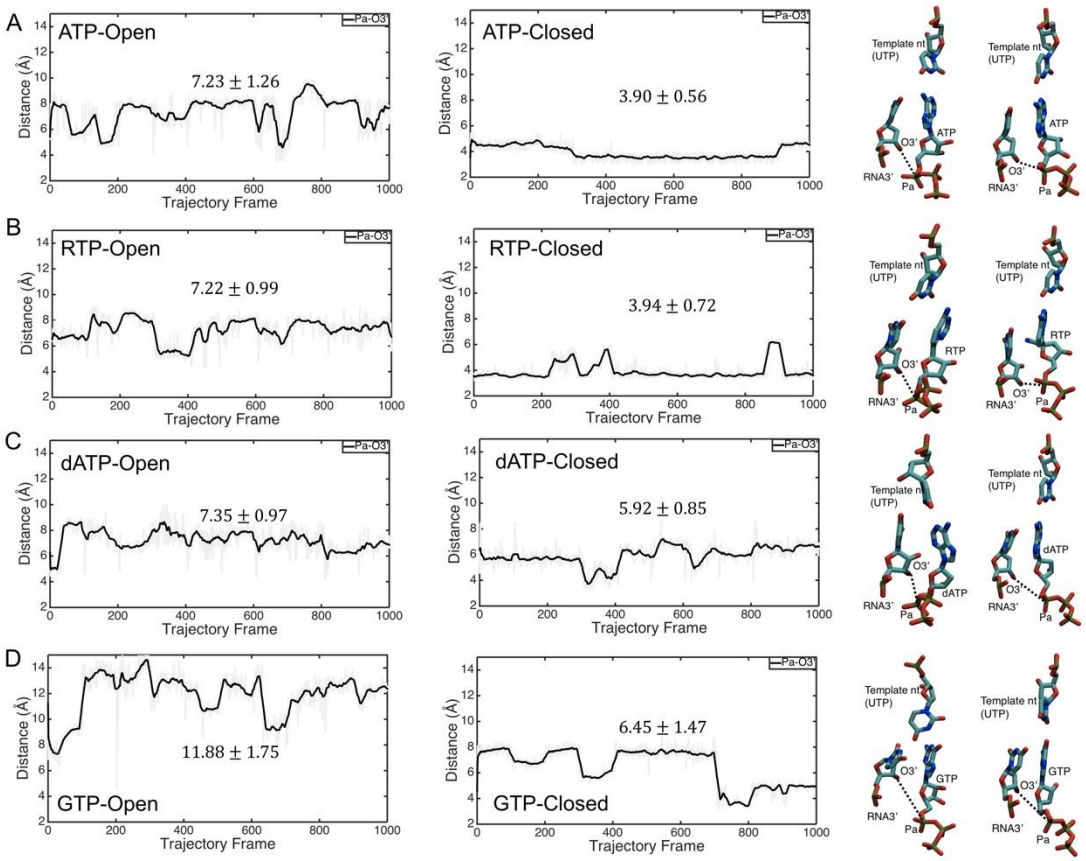

Figure 4. The distances between the bound nucleotides and the 3'-terminal of the nascent RNA strand, for the active-site open (*left*) and closed (*right*) states. The distances measured from the ensemble equilibrium simulations between the $\mathrm{O3'}$ atom of the $3'-\mathrm{end}$ RNA and the P_α atom of the bound nucleotide in the open and closed state are labeled in each diagram, for (A) ATP, (B) RTP, (C) dATP, and (D) GTP . The structures of RNA 3'-end and the template rUTP as well as the bound nucleotides are shown (on the right), with the distances measured labeled.

We then measured the distance between P_α atom of the insertion nucleotide (ATP/RTP/dATP/GTP) and the O3' atom of the 3'-terminal of nascent RNA strand in the closed state. We found that the distance between P_α and O3' of ATP is smallest and most stable (3.90 ± 0.56Å), while that for RTP, dATP, and GTP are 3.94 ± 0.72Å, 5.92 ± 0.85Å, 6.45 ± 1.47Å, respectively (see **Figure 4**). Hence, it appears that the association between cognate ATP and the 3'-end of the RNA primer (e.g. via base stacking) is comparatively stabilized, with RTP slightly less stable, and the GTP insertion the least stabilized. The large deviations of GTP from the 3'-end RNA may dominantly contribute to the overall instability of GTP here.

In addition, we also measured the hydrogen bonding (HB) occupancies or probabilities between the amino acids around the active site and the insertion or template nucleotide (see ATP and RTP in **Figure 5**). The criteria for the HB formation were set as that the distance between the donor- acceptor atoms was less than 3.5Å, and the donor-hydrogen-acceptor angle was larger than 140° [44]. In the closed state, the average HB association occupancies between the amino acids and the insertion or template nucleotide are shown slightly higher for RTP (~35.1%) than for ATP ($\sim 32.5\%$). RTP form four comparatively strong HBs (>50% occupancies) while ATP form three around the

active site. Among them, there are two strongest HBs (>90%) with RTP, i.e., between T687-OG1 and RTP-O2' atom, and between N691-ND2 and RTP-O2' atom (on the sugar), while no such very highly occupied HBs for ATP. Indeed, the two strongest HBs with RTP are formed with its 1'-ribose cyano substitution. In addition, K545-NZ and S682-OG also form HBs with template nt-O4 atom, respectively, at a medium level of occupancies (56.5% and 77.0%) but still higher than that with ATP. The HB associations around the non-cognate dATP and mis-matched GTP are also shown (in **SI Figure S4**), with dATP having similar average HB occupancies (~33.6%) around as ATP, while for GTP it is much lower (~27.4%).

Hence, from both results on base pairing between the insertion nucleotide and template rUTP, and the HB patterns around the active site amino acids and nucleotides, we found that RTP is more stabilized than the natural substrates. Only in association with the 3'-end, the inserted cognate ATP appears to be slightly more stabilized than the inserted RTP.

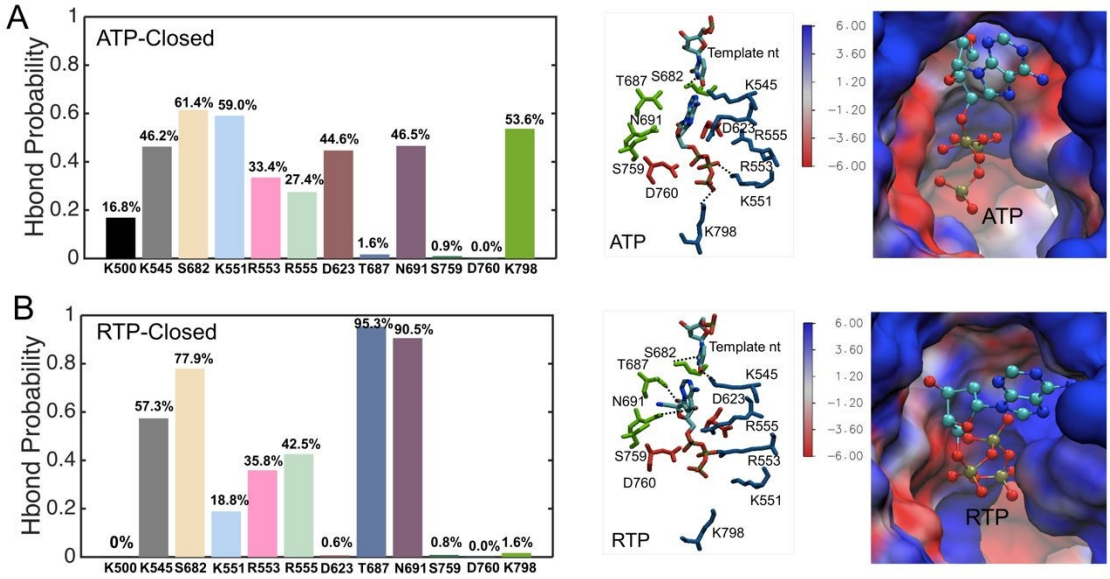

Figure 5. The hydrogen bonding (HB or Hbond) association patterns between the active site amino acids and the insertion nucleotide in the closed state of SARS-CoV-2 RdRp, (A) for ATP and (B) for RTP, with the Hbond occupancies calculated from the ensemble equilibrium simulations (*left*), the molecular views around the active site (*middle*), and the electrostatic potential map generated by solving the Poisson-Boltzmann equation using the APBS [45] in VMD [46] (*right*). In the map, the protein surface is colored based on the electrostatic potential from low (negative) by red to high (positive) by blue, the values in the color bar are in units of $k_BT/|e|$.

Further, we employed g-mmpbsa [47, 48] to calculate the binding energetics of ATP/RTP/dATP/GTP by in the active site using the equilibrium ensemble simulations (see **SI Table S2**). Note that the entropy component is not included in the calculation. The dielectric constant used in the calculation is set at ~10 in order to accommodate the highly charged nucleic acid-protein complex environment. Accordingly, the calculations show that the inserted RTP has a lowest binding energetics (~ -82 kcal/mol) comparing with the others, with ATP ~ -74 kcal/mol, dATP ~ -66 kcal/mol, and GTP ~ -61 kcal/mol. The electrostatic interaction stabilization toward RTP turns out to be most notable, comparing to ATP/dATP/GTP (see **SI Table S2**). Accordingly, the amino acids around the active with electrostatic potential calculated from APBS [45] are displayed in **Figure 5** *right*, from which we can see the RTP has less negative potential but more extensively positive potential around, in comparison with ATP. The relative insertion nucleotide relative binding energetics between RTP/dATP/GTP with respect to cognate ATP (from MM/PBSA calculations) contributed by individual amino acids and nucleotides in SARS-CoV-2 RdRp are shown in **SI Figure S5**. One finds that K411, E436, D499, R513 contribute most to stabilize RTP, while K411, R555, D914 contribute most to destabilize dATP, and K411, K438, R555, and D901 to destabilize GTP. From these results, it seems that K411 and R555 are key to differentiate the NTP species or select against the non-cognate species. Note that R555 locates around the active site while K411 stays comparatively far, i.e., with a distance to NTP about 16 Å. Hence, it appears that some remote residue can still make impacts to nucleotide selectivity.

View Article Online
DOI: 10.1039/D2CP05883A

In brief, the above results largely support that RTP is energetically stabilized. Consequently, it must be the entropy component in the binding free energetics that contributes to stabilize more or favor ATP. One did see that ATP was captured with more flexibilities, e.g. in base paring with the template rU and in HB interactions with local residues around the active site. That is to say, the relative binding free energy (~2.4 kcal/mol) between RTP and ATP, with ATP being more stabilized, would be attributed substantially by entropic advantages of the ATP inserted to the active site, while energetically or the enthalpy contribution to RTP is more significant than ATP.

2. The relative binding free energy (open state)

In the RdRp active site open state upon the initial nucleotide binding, we also calculated the relative binding free energy of RTP and non-cognate dATP/GTP with respect to the cognate ATP (see **Table 1**). More details are provided in **SI Table S3**. The relative binding free energy between RTP and ATP indicated that RTP is only slightly more stable than ATP (~ -0.2 kcal/mol), or say, the initial binding free energy of RTP and ATP are almost identical. Note that during the alchemical simulations, the responses of the template rUTP and 3'-end toward a growing NTP are shown (**SI Figure S6**), which shows that for systems with ATP/RTP/GTP bound in open state, the dihedral angle of template +1 rUTP have larger deviations than those of RNA3'-end, however, for dATP bound system, the dihedral angle of template +1 rUTP and the primer RNA 3'-end react nearly the same (see **SI Figure S6**).

Notably, the alchemical calculation results conducted here for the open and closed state of ATP/RTP in association with SARS-CoV-2 RdRp active site turn about to be quite consistent with the potential of mean force or free energetic calculations on the ATP/RTP from initial binding to insertion to the RdRp active site conducted previously [28] (see **SI Figure S7**).

For non-cognate dATP and mismatched GTP studied in control, they still appear less stabilized than ATP upon the initial binding, for ~ 1.5 kcal/mol and ~1.7 kcal/mol, respectively (see **Table 1**). The destabilization however seems less than that in the closed state, in particular, for GTP. Such a result may be due to the flexibility of the nucleotide upon initial binding (active site open), such that the non-cognate one is less strongly selected against than being forced into the insertion state (active site closed).

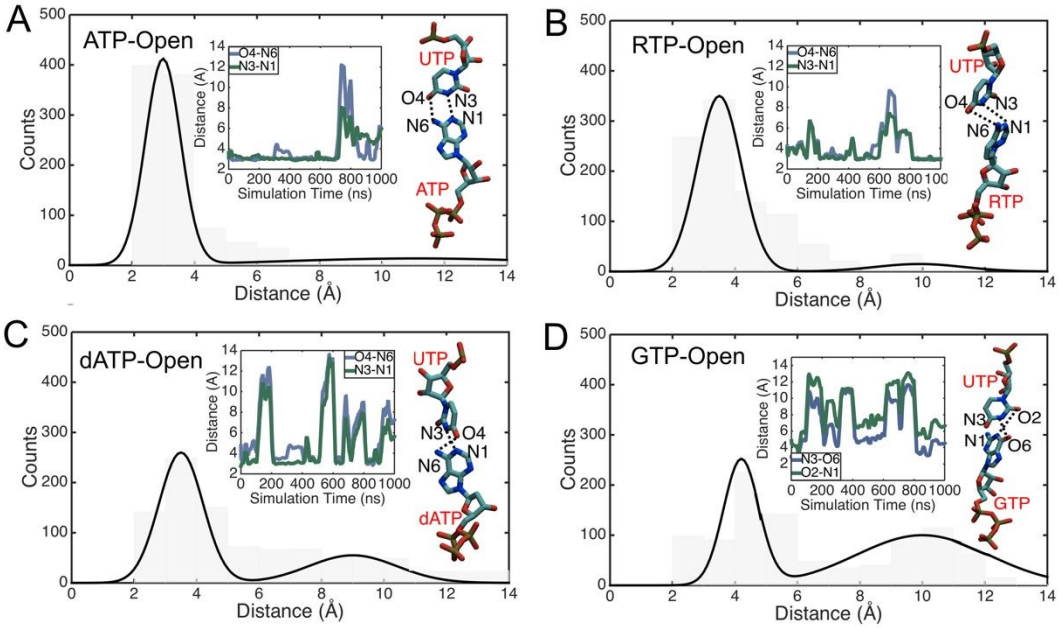

Figure 6. The histograms of measured distances between the bases of the initial binding nucleotides and the template rUTP base in the active site open state. (A) for ATP, (B) for RTP, (C) for dATP, and (D) for GTP. In the insets right, the structure of the initial binding nucleotide and the template rUTP are shown. In the insets middle, the distances between atoms forming hydrogen bond interactions are shown over the simulation time. Note that in the active site open state, ATP can form base pairing with the template UTP, while RTP mainly forms base stacking with the template UTP.

We also sampled from the equilibrium ensemble simulations on the open state RdRp for the initial binding configurations of incoming nucleotides or NTPs. Again, we examined the distance geometries between the

View Article Online
DOI: 10.1039/D2CP05883A

incoming NTP and template rUTP, by measuring the atomic distances that would possibly be involved in the NTP-rUTP base pairing (see **Figure 6**). Note that in the open state, RTP base stacking with template rUTP seems to be preferred [28], we thus measured the atomic distances involved in the base stacking between RTP and template rUTP. We also showed the distances for the base pairing of RTP-rUTP (see **SI Figure S8**). For dATP and GTP, the distances for potential base pairing could only be occasionally shortened, or otherwise become very large due to lack of base pairing interactions. For dATP and GTP, only ~32% and ~16% populations show shortened distances, while for ATP and RTP, the base pairing percentiles are ~71 % and ~53 %, respectively.

In addition, we also calculated the distance between the P_α atom of the intial binding nucleotides (ATP/RTP/dATP/GTP) and the O3' atom of the 3'-terminal of nascent RNA strand in the open state RdRp (see **Figure 4**). Notably, the distances are much larger and with more fluctuations in the open state than in the closed state. For ATP, RTP and dATP, the average distance in the open state is ~ 7.2 to 7.4 Å. The distance reduces to ~ 3.9 Å for ATP and RTP in the closed state, and for dATP it reduces to ~5.9Å. In addition, one can also see that the distance is particularly large for GTP (~11.9Å open and ~6.5Å closed), indicating a strong resistance of the RdRp complex to the base-mismatched GTP, here shown interestingly as being repelled from the 3'-end of the RNA primer.

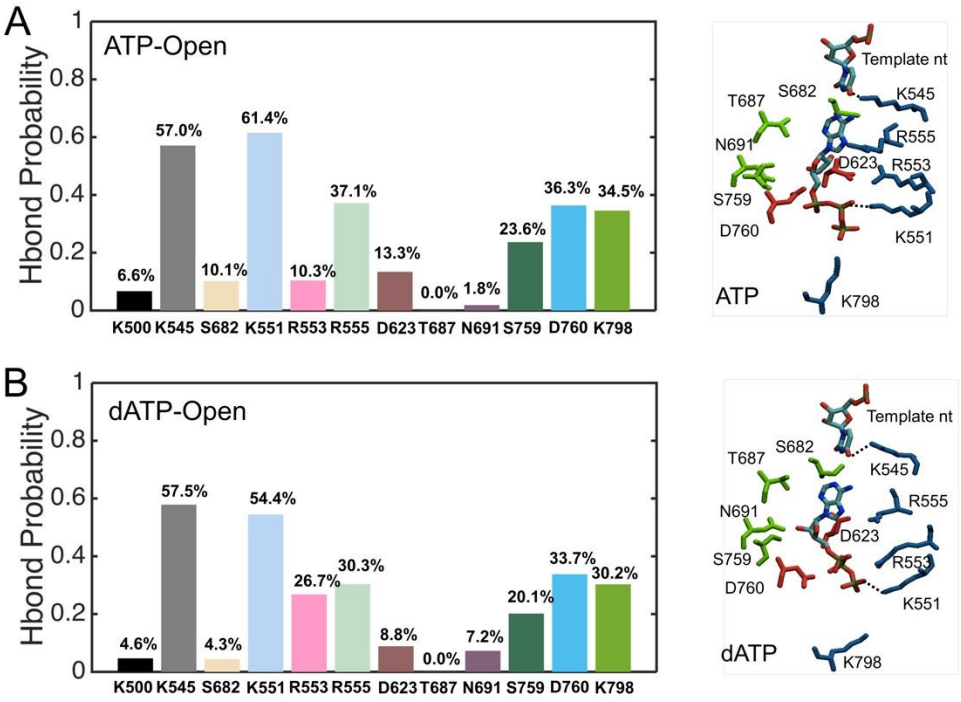

Figure 7. The hydrogen bonding (HB or Hbond) association patterns between the active site amino acids and the initially binding nucleotide in the open state of SARS-CoV-2 RdRp 2, (A) for ATP and (B) for dATP, with the Hbond occupancies calculated from the ensemble equilibrium simulations (*left*) and the molecular views of the active site (*right*).

Further, we examined again the HB occupancies around the active site for the initially bound nucleotides. Interestingly, one can see that the HB patterns are to some extent similar between the cognate ATP and non-cognate dATP (see **Figure 7**), having two HBs formed between K545-NZ and rUTP-O4 and between K551-NZ and ATP/dATP-OG, and with both occupancies higher than 50%. The HB patterns for ATP and RTP are also comparable (see **SI Figure S9**). For example, HBs formed between R555-NH1 and OB atom of both ATP and RTP on the phosphate, with the HB occupancy between R555 and RTP (~55.8 %) larger than that between R555 and ATP (~30.3%). Again, the average HB associations are lowest around non-cognate GTP in initial binding (~16.7%; see **SI Figure S9**), in comparison with that around ATP (~24.3%, RTP (~31.50%) and dATP (~23.2%).

Last, we employed g-mmpbsa [47, 48] again to calculate the binding energetics of ATP/RTP/dATP/GTP for the initial binding nucleotides using the equilibrium ensemble simulations (see **SI Table S4**). Note that the calculations show that the initially bound RTP still has a lowest binding energetics (~ -62.4 kcal/mol) comparing with the others, with ATP ~ -60.0 kcal/mol, dATP ~ -60.9 kcal/mol, and GTP ~ -56.0 kcal/mol. Now the energetic stabilization for RTP mainly comes from the vdW interactions indeed (see **SI Table S4 and SI Figure S9**). Meanwhile, the calculations

View Article Online
DOI: 10.1039/D2CP05883A

show that the initial binding ATP and dATP with similar binding energetics. The total relative binding energetic (RTP/dATP/GTP relative ATP) decompositions to individual residues are shown in **SI Figure S10**. We can find that K411, E802, D846 contribute most to stabilize RTP relative to ATP, while K551 contribute to destabilize dATP slightly relative to ATP, K438, D484, K551, D618, K849 contribute most to destabilize GTP relative to ATP. Still, K411 remains far from the active site, with a distance to NTP about 21 Å.

Discussion

In this work, we studied the SARS-CoV-2 RdRp nucleotide binding stabilities and selectivity at the active site in both the open and closed configurations, pre-catalytically. We mainly conducted all-atom alchemical MD simulations to calculate the relative binding free energetics between nucleotide analogue RTP and natural cognate nucleotide ATP, as well as that between natural non-cognate dATP/GTP and cognate ATP for comparison. Our simulation models were constructed based on high-resolution cryo-EM structures solved for SARS-CoV-2 RdRp or nsp12 along with cofactors nsp7 and nsp8 [23, 24], with the nucleotide initial binding to the active-site in an open form, based on the apo RdRp structure (PDB: 7BTF), and with a bound nucleotide inserted to the closed active-site, modified on top of a product RdRp complex obtained right upon chemical addition of an RDV analog (PDB: 7BV2).

For the closed form RdRp with a comparatively stabilized nucleotide inserted to the active site, pre-catalytically, our work shows that the binding free energy for the cognate ATP is lower than that of RTP, i.e. more stabilized for ~ -2.4 (± 0.7) kcal/mol. In comparison, binding energetics calculated via MMPBSA method (using GROMACS plugin g-mmpbsa ([47, 48]), in the absence of entropy contributions, indicated that RTP is more stable than ATP, energetically. For close examinations, we noticed that the HB interactions around the inserted RTP appear to be more stabilized than those with ATP. In particular, there are two highly stabilized HBs (>90% occupancies) formed between T687-OG1 or N691-ND2 and RTP-O2' on the ribose, which are absent for the inserted ATP.

Essentially, we found that the base-pairing between the template rUTP at +1 and the inserted RTP is highly stable. In comparison, such template base pairing with the inserted cognate ATP, though stabilized in general, reveals occasional flexibilities of the ATP base for fluctuations (~9% non-base paired). On the other hand, one finds that the distance between P_α atom of the inserted ATP and O3' atom of the 3'-terminal of nascent RNA strand appears to be slightly more stabilized (3.90 ± 0.56Å) than that for the inserted RTP (3.94 ± 0.7Å), suggesting that flexibility of the inserted ATP in association with the 3'-end RNA primer is nevertheless restricted. In addition, the MMPBSA binding energetics calculated largely support that the inserted RTP is mostly stabilized, energetically via electrostatic interactions, with ~ -8.1 kcal/mol lower than ATP. The above results consistently suggest that the inserted RTP shows prominent binding energetic stability and rigidity into the active site, which was captured in a closed form right after catalysis but before the active site re-opening [24]. Consequently, the alchemical binding free energy stabilization of the inserted ATP, i.e., ~ -2.4 kcal/mol calculated with respect to the inserted RTP, can only be interpreted by entropic rather than energetic contributions. The entropic advantages of the inserted ATP are revealed from its limited but noticeable flexibility in base pairing with template rUTP, as well as from a HB interactions around the active site. Due to ribose modifications, highly stabilized HBs form additionally from T687 and N691 to RTP but not to ATP. Hence, both template base pairing and HBd network seem to energetically stabilize RTP more but entropically favor ATP in the insertion state. Meanwhile, we note that the high-resolution cryoEM structure of the SARS-CoV-2 RdRp complexes we utilized to construct the closed insertion state was captured originally with RDV in the post-catalytic product state [24], which might also favor the RTP insertion energetically but non-optimized for ATP. Further validations can be made when the high-resolution RdRp complex structure is captured with the cognate ATP insertion to the active site.

For dATP and GTP modeled into the insertion state, they are expectedly less stable than ATP by ~1.4 (± 0.1) kcal/mol and ~3.0 (± 0.7) kcal/mol, respectively. The associations can hardly be stabilized between dATP or GTP and the template rUTP. Although wobble base-pairing between rUTP and GTP were modeled and persisted in the equilibrium simulations (base pairing ~ 77%), they are supposedly similarly stable as the standard Watson-Crick base pairing [49-51]. The association of dATP/GTP with the 3'-terminal of nascent RNA is indeed significantly destabilized, especially for the inserted GTP. Therefore, the non-cognate nucleotides serve for the control showing that the current alchemical simulations well capture their destabilization effects. Note that the non-cognate dATP/GTP was artificially placed into the insertion site similarly as RTP/ATP, which would be hardly achieved in natural conditions due to likely high barriers to the non-cognate NTP insertion, as well as destabilization against the non-cognate NTP for initial binding [18, 19]. Consequently, there would be very low populations of the non-

cognate dATP/GTP to be actually inserted to the closed active site of RdRp. In contrast, though the inserted RTP in the closed active site is less stabilized than the cognate ATP, it is shown as well that the initial binding RTP in the open active site is similarly stabilized as ATP; additionally, it has already been shown that RTP inserts into the active site with a sufficiently low barrier (~1.5 kcal/mol), even lower than that of ATP (~ 2.6 kcal/mol) [28].

Furthermore, according to the equilibrium ensemble simulations conducted for the open and closed states of the SARS-CoV-2 RdRp with nucleotides bound to the active site, we found that the associations (i.e., base pairing) between the template nt +1 and incoming NTP vary significantly from the open to the closed states. The associations (i.e., stacking) of the bound nucleotides with the 3'-terminal of nascent RNA also distinguish well between the open and closed complexes. It has been noticed that the RdRp protein conformational changes between the open and closed forms are subtle, aligned with fingers subdomain, e.g., with an overall RMSD ~1.69 Å for all C_α atoms on average. Close inspections show that the major structural differences between the open and closed form are restricted around the active site or locate on the structural motifs A relative to C [25]. Our simulations show that indeed the bound nucleotides associate or respond quite differently in the open and closed forms of the active sites, i.e., with the nucleotide binding to the open active site much more flexible in association with the template nt+1, with the 3'-end of the RNA primer, and with the local amino acids, than the nucleotide upon insertion to the closed active site, no matter for the natural cognate ATP, nucleotide analog RTP or, or non-cognate dATP/GTP.

For the nucleotide initial binding to the open active site, the binding free energy of RTP is nearly the same as ATP. Putting the relative binding free energetics between RTP and ATP together in the open and closed RdRp complexes, our current alchemical calculations show quite consistent results with the umbrella sampling simulations calculating the free energy profiles of the RTP and ATP insertion [28]. The results consistently show that upon initial binding, RTP in stacking with the template rUTP can be similarly stabilized in binding free energetics as that of ATP, which is in base pairing with rUTP; into the insertion state with the active site closed, RTP becomes more stabilized than its initial binding state for about ~ -2.7 kcal/mol; while ATP upon its insertion becomes more stabilized than its initial binding state for about ~ -5.1 kcal/mol [28], hence ATP becomes more stabilized than the inserted RTP in the binding free energetics by ~ -2.4 kcal/mol, fully consistent with current alchemical calculations (see **SI Figure S7**).

On the other hand, in the open active site of RdRp for the initial nucleotide binding, it is expected that a comparatively large ensemble of configurations exists for incoming NTPs, with some of them captured in current equilibrium ensemble simulations. Nevertheless, the alchemical simulations conducted cannot sample much beyond the nucleotide configuration provided as input for the simulation, hence, the equilibrium ensemble of the initial binding NTP may not be sufficiently characterized. Interestingly, it has been noticed that the HB patterns around the open active site with the incoming ATP are somehow similar to that shown with the incoming dATP upon initial binding. As for the initially bound GTP, however, the base can be highly distinguishable from that of ATP and is thus recognized as a mismatch for rUTP. Only a small populations of wobble pairing could be formed between the initial binding GTP and the template rUTP, hence the GTP-rUTP association is hardly stabilized. Meanwhile, the distance between P_α atom of GTP and O3' atom of the 3'-terminal of nascent RNA also becomes significantly larger than that for the cognate ATP upon initial binding. Besides, the MMPBSA calculations also show that GTP is comparatively destabilized energetically and RTP is most stabilized energetically, while the other two, ATP and dATP show similar binding energetics. The observations thus indicate that GTP has been well detected upon initial binding to the active-site open state, and appears to be selected against already. Such nucleotide selectivity on nucleotide base but not much on ribose upon the initial nucleotide binding had been suggested indeed for the Poliovirus RdRp in its active-site open state [25].

Recently, several computationally intensive works have been conducted related to the RTP incorporation to SASR-CoV-2 RdRp active site for viral RNA synthesizing [52-57]. Among them, one of the all-atom MD simulation studies elucidated that the 1'-cyano group of the remdesivir analog can assist halting RdRp translocation and then inhibit viral RNA replication [52]. Another simulation study had been conducted for calculating the binding energetics of RTP and other nucleotide analogues[53]. They showed that nucleotide analogue compound-17 (-59.6 kcal/mol) binds more strongly than compound-8 (-46.3kcal/mol) and RTP (-29.7kcal/mol) with RdRp. The calculations were conducted by using the apo RdRp crystal structure in the absence of RNA strand[23]. In addition, a very recent MD simulation study was conducted to calculate the binding energies of RTP along with three designed analogues (R1T, R2T, R3T) by using MM-PBSA protocol, and showing that analogues R1T, R2T and R3T with higher affinities (-85.04 kcal/mol, -85.04 kcal/mol, -90.85 kcal/mol) than RTP (-68.24 kcal/mol) [54]**.** Such calculations were conducted by

View Article Online
DOI: 10.1039/D2CP05883A

using the RdRp structure modeled in the closed active site conformation, in the absence of nsp7 and nsp8 proteins. Another MD simulation work calculated the binding free energy of RDV-TP and three hits (SN00166900, SN00303170, SN00359915) which pose similarly to RDV-TP from the natural product database. The study showed that the binding free energies for RDV-TP at -18.90 kcal/mol, -35.33 kcal/mol for SN00166900, -45.23 kcal/mol for SN00303170, and -64.86kcal/mol for SN00359915 (about three times lower than that of RDV-TP) [55]. The corresponding simulations used the homology modeling structure of SARS-CoV-2 nsp12 (PDB:6M71) [23] and HCV RNA primer template (PDB: 4WTG) [58]. Moreover, another alchemical MD simulation study had been conducted calculating the relative binding free energy between RTP and ATP, which shows stabilization of RTP relative to ATP by ~ -2.8 kcal/mol [56]. The study was also conducted in the absence of RNA strand, using the RdRp structure constructed via homology modeling based on SARS-CoV RdRp. An additional MD study on the closed-state complex had been conducted by using advanced samplings [57]. The study showed that the binding free energy were -32.7 kcal/mol, -15.2 kcal/mol and -29.6 kcal/mol for remdesivir, favilavir and ribavirin, respectively, which were obtained by multistate Bennett's acceptance ratio (MBAR) method [57]. In comparison, in current study, we alchemically calculated the relative binding free energetics among RTP analogue, cognate ATP, and non-cognate dATP/GTP, for both the initial binding and insertion states. We modeled the systems using the high-resolution complex structures of SARS-CoV-2 RdRp resolved in both the open (as from apo state [23]) and closed (as from the product state [24]) forms. Our study indicates that RTP and ATP are similarly stabilized upon initial binding, according to the relative binding free energetics obtained from the alchemical calculations. Interestingly, in the insertion state, while RDV-TP or RTP remains less stabilized (~2 kcal/mol) than ATP in terms of binding free energetics, energetically it appears more favored than ATP (i.e., from the restricted geometries with respect to the template and from the MMPBSA energetic calculations), displaying notable rigidity. In contrast, ATP displays sufficient flexibility from initial binding to the more stabilized insertion state. Hence, our current results and analyses suggest that the ATP insertion to the active site of the viral RdRp is entropically more favored than the RTP insertion complex, prechemically.

Conclusions

In this alchemical all-atom simulations characterizing relative binding free energetics between nucleotide analogue RTP and natural cognate substrate ATP upon initial binding and pre-catalytic insertion to the active site of SARS-CoV-2 RdRp, we indicated that upon initial binding, RTP and ATP show similar binding free energies to the active site while in the insertion state, ATP is more stabilized (~ -2.4 kcal/mol) than RTP in free energetics. Additional analyses show that the RTP is more stabilized in binding energetics than ATP, in both the insertion and initial binding states, with RTP more stabilized due to electrostatic energy in the insertion state and due to vdW energy in initial binding state. Therefore, it appears that natural cognate ATP excels at association stability with the RdRp active site due to that ATP maintains sufficient flexibilities e.g., in base pairing with the template, which exemplifies entropic contribution to the cognate substrate stabilization.

Author Contributions

J Yu conceived and designed the computational work. C Long and M.E Romero performed the computation. C Long, L Dai, and M.E Romero analyzed the data. C Long and J Yu wrote the paper.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

CL is supported by National Natural Science Foundation of China grant #12005029 and Natural Science Foundation of Chongqing Grant #cstc2020jcyj-msxmX0811 and the Start-up Founding of Chongqing

University of Posts and Telecommunication (A2020-029). Part of this work was supported by National Science Foundation (NSF) Award #2028935. JY has also been supported by the CMCF of UCI via NSFDMS 1763272 and the Simons Foundation grant #594598 and start-up funding from UCI.

References

1. WHO. http://covid19.who.int/. 2022.

2. Lythgoe MP, Middleton P. Ongoing clinical trials for the management of the COVID-19 pandemic. Trends in pharmacological sciences. 2020;41(6):363-82.

3. Sarisky RT. Non-nucleoside inhibitors of the HCV polymerase. Journal of Antimicrobial Chemotherapy. 2004;54(1):14-6.

4. Jordheim LP, Durantel D, Zoulim F, Dumontet C. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. Nature reviews Drug discovery. 2013;12(6):447-64.

5. Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, et al. Remdesivir for the treatment of Covid-19. New England Journal of Medicine. 2020;383(19):1813-26.

6. Warren TK, Jordan R, Lo MK, Ray AS, Mackman RL, Soloveva V, et al. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. Nature. 2016;531(7594):381-5.

7. Sheahan TP, Sims AC, Graham RL, Menachery VD, Gralinski LE, Case JB, et al. Broad-spectrum antiviral GS-5734 inhibits both epidemic and zoonotic coronaviruses. Science translational medicine. 2017;9(396):eaal3653.

8. Agostini ML, Andres EL, Sims AC, Graham RL, Sheahan TP, Lu X, et al. Coronavirus susceptibility to the antiviral remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease. MBio. 2018;9(2).

9. Shannon A, Le NT-T, Selisko B, Eydoux C, Alvarez K, Guillemot J-C, et al. Remdesivir and SARS-CoV-2: Structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites. Antiviral research. 2020;178:104793.

10. Joyce CM, Benkovic SJ. DNA polymerase fidelity: kinetics, structure, and checkpoints. Biochemistry. 2004;43(45):14317-24.

11. Sydow JF, Cramer P. RNA polymerase fidelity and transcriptional proofreading. Current opinion in structural biology. 2009;19(6):732-9.

12. Cameron C, Moustafa I, Arnold J. Fidelity of nucleotide incorporation by the RNA-dependent RNA polymerase from poliovirus. The Enzymes. 2016;39:293-323.

13. Long C, Romero ME, La Rocco D, Yu J. Dissecting nucleotide selectivity in viral RNA polymerases. Computational and Structural Biotechnology Journal. 2021;19:3339-48.

14. Yuzenkova Y, Bochkareva A, Tadigotla VR, Roghanian M, Zorov S, Severinov K, et al. Stepwise mechanism for transcription fidelity. BMC biology. 2010;8(1):1-15.

15. Yu J. Efficient fidelity control by stepwise nucleotide selection in polymerase elongation. Abstract: Polymerases select nucleotides. Computational and Mathematical Biophysics. 2014;2(1):141-60.

16. Johnson KA. Conformational coupling in DNA polymerase fidelity. Annual review of biochemistry. 1993;62(1):685-713.

17. Anand VS, Patel SS. Transient state kinetics of transcription elongation by T7 RNA polymerase. Journal of Biological Chemistry. 2006;281(47):35677-85.

18. Long C, E C, Da L-T, Yu J. Determining selection free energetics from nucleotide pre-insertion to insertion in viral T7 RNA polymerase transcription fidelity control. Nucleic acids research. 2019;47(9):4721-35.

19. Long C, Chao E, Da L-T, Yu J. A viral t7 RNA polymerase ratcheting along DNA with fidelity control. Computational and structural biotechnology journal. 2019;17:638-44.

20. Long C, Yu J. Balancing Non-Equilibrium driving with nucleotide selectivity at kinetic checkpoints in polymerase fidelity control. Entropy. 2018;20(4):306.

21. Temiakov D, Patlan V, Anikin M, McAllister WT, Yokoyama S, Vassylyev DG. Structural basis for substrate selection by T7 RNA polymerase. Cell. 2004;116(3):381-91.

22. Yin YW, Steitz TA. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. Cell. 2004;116(3):393-404.

23. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, Cao L, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science. 2020;368(6492):779-82.

24. Yin W, Mao C, Luan X, Shen D-D, Shen Q, Su H, et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science. 2020;368(6498):1499-504.

25. Gong P, Peersen OB. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. Proceedings of the National Academy of Sciences. 2010;107(52):22505-10.

26. Wang Q, Wu J, Wang H, Gao Y, Liu Q, Mu A, et al. Structural basis for RNA replication by the SARS-CoV-2 polymerase. Cell. 2020;182(2):417-28. e13.

27. Kokic G, Hillen HS, Tegunov D, Dienemann C, Seitz F, Schmitzova J, et al. Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. Nature communications. 2021;12(1):1-7.

28. Romero ME, Long C, La Rocco D, Keerthi AM, Xu D, Yu J. Probing remdesivir nucleotide analogue insertion to SARS-CoV-2 RNA dependent RNA polymerase in viral replication. Molecular Systems Design & Engineering. 2021;6(11):888-902.

29. Shirts MR, Pitera JW, Swope WC, Pande VS. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. The Journal of chemical physics. 2003;119(11):5740-61.

30. Mobley DL, Chodera JD, Dill KA. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. The Journal of chemical physics. 2006;125(8):084902.

31. Villa A, Mark AE. Calculation of the free energy of solvation for neutral analogs of amino acid side chains. Journal of computational chemistry. 2002;23(5):548-53.

32. Christ CD, Mark AE, Van Gunsteren WF. Basic ingredients of free energy calculations: a review. Journal of computational chemistry. 2010;31(8):1569-82.

33. Zwanzig RW. High-temperature equation of state by a perturbation method. I. Nonpolar gases. The Journal of Chemical Physics. 1954;22(8):1420-6.

34. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1:19-25.

35. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. Journal of chemical theory and computation. 2015;11(8):3696-713.

36. Ivani I, Dans PD, Noy A, Pérez A, Faustino I, Hospital A, et al. Parmbsc1: a refined force field for DNA simulations. Nature methods. 2016;13(1):55-8.

37. Meagher KL, Redman LT, Carlson HA. Development of polyphosphate parameters for use with the AMBER force field. Journal of computational chemistry. 2003;24(9):1016-25.

38. Price DJ, Brooks III CL. A modified TIP3P water potential for simulation with Ewald summation. The Journal of chemical physics. 2004;121(20):10096-103.

39. Darden T, York D, Pedersen L. Particle mesh Ewald: An N · log (N) method for Ewald sums in large systems. The Journal of chemical physics. 1993;98(12):10089-92.

40. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. The Journal of chemical physics. 1995;103(19):8577-93.

41. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied physics. 1981;52(12):7182-90.

42. Nosé S, Klein M. Constant pressure molecular dynamics for molecular systems. Molecular Physics. 1983;50(5):1055-76.

43. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. The Journal of chemical physics. 2007;126(1):014101.

44. Kumar R, Schmidt J, Skinner J. Hydrogen bonding definitions and dynamics in liquid water. The Journal of chemical physics. 2007;126(20):05B611.

45. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the APBS biomolecular

View Article Online
DOI: 10.1039/D2CP05883A

solvation software suite. Protein Science. 2018;27(1):112-28.

46. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. Journal of molecular graphics. 1996;14(1):33-8.

47. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proceedings of the National Academy of Sciences. 2001;98(18):10037-41.

48. Kumari R, Kumar R, Consortium OSDD, Lynn A. g_mmpbsa A GROMACS tool for high-throughput MM-PBSA calculations. Journal of chemical information and modeling. 2014;54(7):1951-62.

49. Crick FH. Codon—anticodon pairing: the wobble hypothesis. Journal of molecular biology. 1966;19(2):548-55.

50. Varani G, McClain WH. The G· U wobble base pair. EMBO reports. 2000;1(1):18-23.

51. Vendeix FA, Munoz AM, Agris PF. Free energy calculation of modified base-pair formation in explicit solvent: A predictive model. Rna. 2009;15(12):2278-87.

52. Zhang L, Zhang D, Wang X, Yuan C, Li Y, Jia X, et al. 1'-Ribose cyano substitution allows Remdesivir to effectively inhibit nucleotide addition and proofreading during SARS-CoV-2 viral RNA replication. Physical Chemistry Chemical Physics. 2021;23(10):5852-63.

53. Wakchaure PD, Ghosh S, Ganguly B. Revealing the inhibition mechanism of RNA-dependent RNA polymerase (RdRp) of SARS-CoV-2 by remdesivir and nucleotide analogues: A molecular dynamics simulation study. The Journal of Physical Chemistry B. 2020;124(47):10641-52.

54. Arba M, Paradis N, Wahyudi ST, Brunt DJ, Hausman KR, Lakernick PM, et al. Unraveling the binding mechanism of the active form of Remdesivir to RdRp of SARS-CoV-2 and designing new potential analogues: Insights from molecular dynamics simulations. Chemical physics letters. 2022;799:139638.

55. Arba M, Wahyudi ST, Brunt DJ, Paradis N, Wu C. Mechanistic insight on the remdesivir binding to RNA-Dependent RNA polymerase (RdRp) of SARS-cov-2. Computers in biology and medicine. 2021;129:104156.

56. Zhang L, Zhou R. Structural basis of the potential binding mechanism of remdesivir to SARS-CoV-2 RNA-dependent RNA polymerase. The Journal of Physical Chemistry B. 2020;124(32):6955-62.

57. Byléhn F, Menéndez CA, Perez-Lemus GR, Alvarado W, De Pablo JJ. Modeling the binding mechanism of remdesivir, favilavir, and ribavirin to SARS-CoV-2 RNA-dependent RNA polymerase. ACS central science. 2021;7(1):164-74.

58. Appleby TC, Perry JK, Murakami E, Barauskas O, Feng J, Cho A, et al. Structural basis for RNA replication by the hepatitis C virus polymerase. Science. 2015;347(6223):771-5.