Supporting information

The 2’- and 3’-Ribose Modifications of Nucleotide Analogs
Establish the Structural Basis to Inhibit the Viral Replication of SARS-CoV-2

Yongfang Li\textsuperscript{a,b}, Dong Zhang\textsuperscript{a,b}, Xin Gao\textsuperscript{c}, Xiaowei Wang\textsuperscript{d} and Lu Zhang\textsuperscript{a,b,*}

\textsuperscript{a} State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, 350002, Fuzhou, Fujian, China
\textsuperscript{b} University of Chinese Academy of Sciences, 100864, Beijing, China
\textsuperscript{c} Computational Bioscience Research Center, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Saudi Arabia
\textsuperscript{d} Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Kowloon 999077, Hong Kong

*Correspondence can be addressed to luzhang@fjirsm.ac.cn
1. Molecular dynamics simulations for the SARS-CoV-2 RdRp with nucleotide analogs.

The previously constructed simulation model\textsuperscript{1} with ATP at the closed active site of SARS-CoV-2 RdRp was adopted as the structural basis for constructing our simulations model of SARS-CoV-2 RdRp with the nucleotide analogs at the closed active site ($i$ site) or at the $i+1$ site. This model has been validated against several experimental structures\textsuperscript{2,3} and demonstrated high structural similarity\textsuperscript{1}. Therefore, this model, which has been well equilibrated after 100ns molecular dynamics (MD) simulation, serves as an ideal starting point for us to model the SARS-CoV-2 RdRp with nucleotide analogs.

To investigate the binding stability and incorporation capability of nucleotide analogs, ATP was replaced with the corresponding nucleotide analogs in the triphosphate form at the active site. To examine the inhibitory effect on the nucleotide addition at the active site, the nucleotide analogs that retain the 3’-hydroxyl group were also placed at the 3’-terminal of the product strand. For each system, the complex was solvated with 68,001 TIP3P water molecules\textsuperscript{4} in a 146Å×146Å×146Å dodecahedron box. Sufficient sodium ions were added to ensure the whole system is neutral. Two rounds of 10,000 steps energy minimization were conducted with and without position restraints on the heavy atoms of dsRNA, respectively. We then performed a 500 ps NVT equilibration simulation at T=300K with position restraints on all the heavy atoms with a force constant of 10kJ×mol$^{-1}$×Å$^{-2}$. Afterwards, another 500 ps equilibration simulation with position restraints on the heavy atoms under the NPT ensemble (P=1bar, T=300K) was performed. The last configuration of the NPT equilibration simulation was used to seed MD production simulations under the NVT ensemble (T=300K). For each system, 20×50ns trajectories were generated and the snapshot was saved every 20 ps. For Clofarabine, Didanosine, Fludarabine and Vidarabine, simulations were extended to 100ns and the convergence of the simulation data has been evaluated accordingly (Figure S1). The integration time step was 2 fs and the neighbor list was updated every 10 steps. The cut off of the Van der Waals interactions is 12Å. The short range electrostatic interactions were cut off at 12Å and the Particle-Mesh Ewald method\textsuperscript{5} is used to treat the long-range electrostatic interaction. All bonds were constrained using LINC algorithms\textsuperscript{6}. The V-rescale thermostat was used to control the temperature at 300K with coupling time of 0.1 ps\textsuperscript{7}.

The amber99sb-ildn force field\textsuperscript{8} was used to simulate the protein and nucleotides. We followed a similar scheme as used by amber force field to obtain the force field parameters for nucleotide analogs. Specifically, we used the neutral form of nucleotide analogs for parameterization. The structure was optimized using HF/6-31g* in Gaussian 16 package\textsuperscript{9} followed by a single point calculation. The electrostatic potential generate by the single point calculations was then used to obtain the partial charges of nucleotide analogs by Restricted Electrostatic Potential (RESP) approach\textsuperscript{10}. Distance constraints were added between Zn$^{2+}$ and surrounding histidine residues (HIE295, HIE642) and cysteine residues (CYX301, CYX306, CYX310, CYX487, CYX645,
CYX646) to maintain their coordination. The histidine residues adopt the similar protonation states as previous model\(^1\). Specifically, some histidine residues in nsp12 (residue IDs: 295, 309, 642, 872 and 892) and histidine residues in nsp7 (residue IDs: 36) are protonated at \(\text{N}_\epsilon\) while the remaining histidine residues are protonated at \(\text{N}_\delta\). All the MD simulations were performed using Gromacs 5.0 package\(^1\).

2. Structural analysis

We used the 20×50ns MD simulations for analyzing and comparing the analogs’ stability and incorporation efficiency in the active site (Figure 2-3 in the maintext). When examining the analog’s effect on the next nucleotide addition (Figure 4 in the maintext), 20×100ns MD simulations were used for Clofarabine, Fludarabine and Vidarabine while 20×50ns MD simulations were used for 2’-NH\(_2\)-dA. When analyzing the data for Cordycepin (Figure 5 in the maintext), we still used the 20×100ns MD simulations. For comparison, the MD conformations with ATP at the active site or with adenosine at the 3’-terminal of nascent strand from the previous work\(^1\) were utilized. In all analysis, the first 20ns in each trajectory was deleted before performing the structural analysis. Unless stated otherwise, bootstrap algorithm was adopted to calculate the mean value and the standard deviations in this work. Specifically, for each system, we generate 20 bootstrap samples for the calculation. For each bootstrap sample, 20 trajectories were chosen randomly from the conformational ensemble. When demonstrating the results of structural analysis, the nucleotide analogs (Clofarabine, Didanosine, Fludarabine, Vidarabine, Cordycepin, 2’-NH\(_2\)-dA and 2’3’-Didehydro-2’3’-dideoxyadenosine) are abbreviated as XXP (COP, DI3, FL3, VD3, CR3, BN3 and ST3) when they are at i site, while as XX3(CO3, DI3, FL3, VD3, CR3, BN3 and ST3) when they are at i+1 site.

To achieve a more comprehensive understanding about the structural properties of nucleotide analogs, additional structural analysis has been performed. Specifically, we have calculated the hydrogen bonding environment around the ribose ring, the salt-bridge interactions around the triphosphate moiety, as well as some global movement of RdRp and allosteric response of the motifs upon binding of different nucleotide analogs (Figure S3-S5). For all the analogs as well as natural substrate ATP, conformations from the 20×50ns simulations were used (the first 20ns in each trajectory was removed before performing the analysis). Our results have shown that only the hydrogen bonding environment surrounding the ribose is distinguishable for different analogs, while other properties are similar. This also justifies our choice of Asp623, which could form hydrogen bond with 2’-hydroxyl group in ATP, as the metric to estimate the structural stability of nucleotide analogs in the active site of SARS-CoV-2 RdRp.

First, we have identified key amino acids around the active site in hydrogen bonding or salt-bridge interactions with the analog (Figure S3A). In particular, we first explored the hydrogen bonds around the ribose and found out the 2’-hydroxyl group could interact with Asp623 (Figure S3B) to affect the orientation of ribose. Specifically, no hydrogen bond is found for Clofarabine-TP (COP), Didanosine-TP (DIP) and 2’3’-Didehydro-2’3’-dideoxyadenosine-TP (STP) as the 2’-hydroxyl group is replaced with
hydrogen atom. For the analogs (Cordycepin-TP (CRP), 2’-NH₂-dATP (BNP)) that possess the 2’-hydroxyl group or 2’-amino group, we have observed a similar hydrogen bonding probability as ATP. Fludarabine-TP (FLP) and Vidarabine-TP (VDP) have demonstrated even higher hydrogen bonding probability than that of ATP, which could be attributed to that the 2’-hydroxyl group is located on the opposite side of the ribose relative to that in ATP (Figure 1E and 1F in the maintext) and thus the hydrogen bond is formed along a direction different from that in ATP. Such a discrepancy has also lead to that FLP and VDP adopt a distinct orientation relative to Asp623 (Figure 3 and Figure S6) and thus perturbed the overall structural stability. Next, we examined the salt-bridge interactions between the triphosphate moiety and the protein environment, and found out there exist several pairs of interactions (Figure S3A and Figure S3C-S3E). Specifically, the Pₐ group is interacting with Arg555, with the mean distance ~3.3 Å in all the nucleotide analogs as well as natural substrate ATP (Figure S3C). Both the Pₜ group and Pₕ group form salt bridges with Arg553 in all the systems, with the mean distance of ~3.3 Å and ~2.9 Å, respectively (Figure S3D and S3E). Therefore, the above analysis has suggested that the 2’-ribose modifications on nucleotide analogs have led to the different hydrogen bonding pattern surrounding the ribose while the salt-bridge interactions between triphosphate moiety and protein environment are maintained as the natural ATP.

Second, we have examined the global movement of SARS-CoV-2 RdRp by performing the principle component analysis (PCA) (Figure S4). In particular, we have extracted the Cₐ atoms of all the systems and aligned all the MD conformations based on the Cₐ atoms of nsp12. The PCA analysis was then performed and we found out each of the top two components has contributed more than 10% of the overall movement (PC1: 25.0% and PC2: 10.5%, Figure S4A) while the remaining PCs has individual contribution < 5%. Both PC1 and PC2 correspond to the movement of nsp8b although the direction of the movement described by the two PCs are different (Figure S4B). We have found that such a movement of nsp8b is observed in all the SARS-CoV-2 RdRps containing different nucleotide analogs. Specifically, the MD conformations of each nucleotide system were separately projected onto the top two PCs and we found out they have sampled a similar distribution (Figure S4C), suggesting that the binding of different nucleotide analogs demonstrates the similar global movement of RdRp as the binding of natural ATP.

Third, we have also investigated if the binding of nucleotide analogs would affect the flexibility of different motifs (Figure S5), as an inference to the allosteric effect exerted by nucleotide analogs relative to ATP. To achieve this, we computed the RMSF of each individual residue in the A-G motifs for all the systems. We found out all the systems with bound nucleotide analogs have demonstrated a similar trend of RMSF as the system with ATP binding at the active site (Figure S5). Notably, the motifs D, E, F and G have shown higher structural flexibility than the other motifs, which are consistent with a previous computational study of viral RdRps. We also noticed that in our computational results, the flexibility of motif A is small while the previous work of viral RdRps has observed a higher flexibility in motif A. The difference could be attributed to that the structures used in the previous work contain an open active site
and in the absence of dsRNA, while our simulation models have ATP or nucleotide analog binding at a closed active site, which would exert extra stabilization to the motifs (including motif A) surrounding the active site.

Overall, the above additional structural analysis has indicated that the hydrogen bonding environment is affected by the 2'-ribose modifications (Figure S3B), while other structural features (including the salt-bridge interactions (Figure S3C-S3E), the global movement of RdRp (Figure S4) and allosteric response of motifs (Figure S5) are similar between nucleotide analogs and natural substrate ATP. We also noticed that our simulations have 50ns in length per trajectory, and thus may not be sufficiently long to capture the global movement induced by the nucleotide analogs, which could happen at a timescale of tens of microseconds or milliseconds. Even so, the convergence analysis based on the structural features around the active site (Figure S1) has suggested the 20x50ns simulations could serve the purpose to investigate the incorporation of nucleotide analog and the immediate termination, which take place at a more local region around the active site.

3. Calculation of relative binding free energy

The relative binding free energy ($\Delta\Delta G_{\text{binding}}$) was computed for each nucleotide analog relative to the ATP binding at the active site of SARS-CoV-2 RdRp. The thermodynamic cycle (Figure S2) was designed and free energy perturbation method was utilized to estimate the $\Delta\Delta G_{\text{binding}}$ based on explicit solvent MD simulations.

Accordingly, the difference in the binding free energy ($\Delta G = \Delta G_{\text{Analog}}^{\text{binding}} - \Delta G_{\text{ATP}}^{\text{binding}}$) can be obtained by calculating the free energy changes between $\Delta G_{\text{ATP} \rightarrow \text{Analog}}^{\text{bound}} - \Delta G_{\text{ATP} \rightarrow \text{Analog}}^{\text{unbound}}$ (Figure S2). The system Hamiltonian is coupled to a parameter $\lambda$, which gradually drives the system from ATP ($\lambda=0$) to the analog ($\lambda=1$). Specifically, simulations were performed with 27 $\lambda$ values ($\lambda=0, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.42, 0.44, 0.46, 0.48, 0.50, 0.52, 0.54, 0.56, 0.58, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1) For each $\lambda$, the system was first energy minimized by 10,000 steps. Subsequently, 200ps position restraint simulations (force constant of 10 kJ×mol$^{-1}$×Å$^{-2}$ on all the heavy atoms) were performed under NVT (T=300 K) ensemble, followed by another 500ps position restraint simulations under NPT (T=300 K and P=1 bar) ensemble. Afterwards, one 20ns simulations were performed under NPT ensemble (T=300 K and P=1bar) to obtain the relative binding free energy using the Multistate Bennett Acceptance Ratio method.

4. Investigation into the performance of Remdesivir in SARS-CoV-2 RdRp

Remdesivir is a well-characterized chain terminator for SARS-CoV-2 RdRp and should be included as a control to validate our computational model and methodology. Remdesivir is an adenosine analog prodrug that has the 1’-ribose modification (Figure S10). In our previous work, we have performed extensive MD simulations to examine the “delayed chain termination” mechanism of Remdesivir in SARS-CoV-2 RdRp. Our previous results are consistent with the experimental observations that Remdesivir can
be efficiently incorporated into the nascent strand while the translocation of Remdesivir from \(i+3\) to \(i+4\) site is hindered and thus the termination is delayed\(^{17, 18, 19}\). Therefore, the MD conformations of Remdesivir in SARS-CoV-2 RdRp from our previous work can serve as a good control to compare with the performance of the nucleotide analogs investigated in the current manuscript. Moreover, we have also performed extra calculations to estimate the relative binding free energy of Remdesivir in the triphosphate form (RDV-TP) as another control to validate our computational protocol and compared it with that of the testing nucleotide analogs. Below are the details about the results of Remdesivir as well as the respective comparison and discussions.

First, we have compared the capability of RDV-TP in maintaining the catalytically active conformation in the active site with that of Didanosine-TP (DIP) and Cordycepin-TP (CRP), which are the two nucleotide analogs we have proposed to be effective in inhibiting the RNA synthesis in SARS-CoV-2 RdRp in the current work (Figure S10). Our calculations based on 20x100ns MD simulations for each analog have shown when RDV-TP is in the active site, the population of MD conformations that can maintain the \(P_a-O3'\) distance within 4 Å (57.8±5.9%) is similar to that when DIP or CRP is in the active site (50.7±8.8% and 53.7±7.8% for DIP and CRP, respectively) (Figure S10B). Moreover, our calculations have shown that the base pairing is well formed between RDV-TP:U pair (89.3±2.8%) at the active site, resembling that of DIP/CRP:U pair with a hydrogen bonding probability of 87.7±2.5% and 83.4±3.5%, respectively (Figure S10C). Besides, RDV-TP has demonstrated similar structural flexibility in the active site and adopted similar orientation relative to Asp623 as DIP and CRP (Figure S10D and S10E). All these observations have suggested that RDV-TP has similar capability to maintain the catalytically active conformation as DIP and CRP in the active site. As the experimental data has supported that RDV-TP can be efficiently incorporated into the nascent strand\(^{17, 18}\), the above comparison using RDV-TP as a control has further consolidated our proposal that DIP and CRP can be incorporated into the nascent strand.

Furthermore, as another control to validate our computational protocol, extra calculations were performed to estimate the \(\Delta\Delta G\) of RDV-TP relative to ATP in SARS-CoV-2 RdRp using free energy perturbation method. Specifically, the same protocol as used for calculating the \(\Delta\Delta G\) of other nucleotide analogs based on the thermodynamic cycle (Figure S2, see SI Section 3 for details) was applied to compute the \(\Delta\Delta G\) of RDV-TP relative to ATP. Our calculations have shown that RDV-TP has a more favorable binding than ATP, as the \(\Delta\Delta G\) equals to -6.192±0.362 kcal/mol. This is consistent with the previous computational studies\(^{20, 21}\) as well as the experimental measurement of the binding affinity\(^{18}\). This also validates our protocol for estimating the binding affinity of nucleotide analogs.

We also compared the binding affinity of RDV-TP with the two potential inhibitors (Didanosine-TP (DIP) and Cordycepin-TP (CRP)) we proposed in the current manuscript. We note that both DIP (\(\Delta\Delta G=1.603±0.500\) kcal/mol) and CRP (\(\Delta\Delta G=-0.767±0.437\) kcal/mol) have less favorable binding than RDV-TP (\(\Delta\Delta G=-6.192±0.362\) kcal/mol). Even so, the binding affinities of DIP and CRP are comparable to that of ATP.
and further structural analysis (as shown in Figure S10) have suggested that both of them can still well maintain the catalytically active configuration in the active site as RDV-TP does, and thereby could still be incorporated into the nascent strand. Interestingly, a manuscript has been just published\textsuperscript{22}, which has used SARS-CoV-2 antiviral assay and observed that Cordycepin (CRP) is effective in inhibiting SARS-CoV-2 viral replication. This experimental work has well validated our computational proposal that CRP can serve as an obligate terminator for SARS-CoV-2 RdRp. More importantly, it again underlined that the binding affinity is not deterministic to estimate the inhibitory potential of nucleotide analogs and extra structural analysis need to be performed to deliver an overall evaluation.

In summary, our investigations into the performance of RDV-TP in SARS-CoV-2 RdRp have served as solid controls and a comprehensive comparison with DIP and CRP have further confirmed that our computational model and protocol are reasonable for the evaluation of the inhibitory potential of nucleotide analogs against SARS-CoV-2 RdRp.
References

(1) Zhang, L.; Zhang, D.; Wang, X.; Yuan, C.; Li, Y.; Jia, X.; Gao, X.; Yen, H.-L.; Cheung, P. P.-H.; Huang, X. l'-Ribose Cyano Substitution Allows Remdesivir to Effectively Inhibit Nucleotide Addition and Proofreading During SARS-CoV-2 Viral RNA Replication. Phys. Chem. Chem. Phys. 2021, 23, 5852-5863.

(2) Hillen, H. S.; Kokie, G.; Farnung, L.; Dienemann, C.; Tegunov, D.; Cramer, P. Structure of Replicating SARS-CoV-2 Polymerase. Nature 2020, 584, 154-156.

(3) Wang, Q.; Wu, J.; Wang, H.; Gao, Y.; Liu, Q.; Mu, A.; Ji, W.; Yan, L.; Zhu, Y.; Zhu, C.; et al. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. Cell 2020, 182, 417-428.

(4) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. J. Chem. Phys. 1983, 79, 926-935.

(5) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. J. Chem. Phys. 1995, 103, 8577-8593.

(6) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. LINCS: A Linear Constraint Solver for Molecular Simulations. J. Comput. Chem. 1997, 18, 1463-1472.

(7) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. J. Chem. Phys. 2007, 126, 014101.

(8) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99sb Protein Force Field. Proteins 2010, 78, 1950-1958.

(9) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. A Well-Behaved Electrostatic Potential Based Method using Charge Restraints for Deriving Atomic Charges - The RESP Model. J. Phys. Chem. 1993, 97, 10269-10280.

(10) Abraham, M. J. Abraham, T. M., R. Schulz, S. Pa’l, J. C. Smith, B. Hess and E. Lindahl Gromacs. SoftwareX 2015, 1, 19-25.

(11) Moustafa, I. M.; Shen, H.; Morton, B.; Colina, C. M.; Cameron, C. E. Molecular Dynamics Simulations of Viral RNA Polymerases Link Conserved and Correlated Motions of Functional Elements to Fidelity. J. Mol. Biol. 2011, 410, 159-181.

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

(13) Jorgensen, W. L.; Ravimohan, C. Monte Carlo Simulation of Differences in Free Energies of Hydration. J. Chem. Phys., 83, 3050-3054.

(14) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the Analysis of Free Energy Calculations. J. Comput.-Aided Mol. Des. 2015, 29, 397-411.

(15) Haber, M. Bayesian Estimation of Free Energies from Equilibrium Simulations. Phys. Rev. Lett. 2012, 109, 100601.

(16) Wang, Q.; Wu, J.; Wang, H.; Gao, Y.; Liu, Q.; Mu, A.; Ji, W.; Yan, L.; Zhu, Y.; Zhu, C.; et al. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. Cell 2020, 182, 417-428.
(18) Gordon, C. J.; Tchesnokov, E. P.; Woolner, E.; Perry, J. K.; Feng, J. Y.; Porter, D. P.; Gotte, M. Remdesivir is a Direct-Acting Antiviral that Inhibits RNA-Dependent RNA Polymerase from Severe Acute Respiratory Syndrome Coronavirus 2 with High Potency. *J. Biol. Chem.* **2020**, *295*, 6785-6797.

(19) Tchesnokov, E. P.; Gordon, C. J.; Woolner, E.; Kocinkova, D.; Perry, J. K.; Feng, J. Y.; Porter, D. P.; Gotte, M. Template-Dependent Inhibition of Coronavirus RNA-Dependent RNA Polymerase by Remdesivir Reveals a Second Mechanism of Action. *J. Biol. Chem.* **2020**, *295*, 16156-16165.

(20) Elfiky, A. A. Ribavirin, Remdesivir, Sofosbuvir, Galidesivir, and Tenofovir against SARS-CoV-2 RNA Dependent RNA Polymerase (RdRp): A Molecular Docking Study. *Life Sci.* **2020**, *253*, 117592.

(21) Zhang, L.; Zhou, R. Structural Basis of the Potential Binding Mechanism of Remdesivir to SARS-CoV-2 RNA-Dependent RNA Polymerase. *J. Phys. Chem. B* **2020**, *124*, 6955-6962.

(22) Rabie, A. M. Potent Inhibitory Activities of the Adenosine Analogue Cordycepin on SARS-CoV-2 Replication. *ACS Omega* **2022**, *7*, 2960-2969.
Figure S1. Investigation of the structural stability of nucleotide analogs in the active site of SARS-CoV-2 RdRp using MD simulations of different lengths. (A) Comparison of the population of MD conformations showing the Pα-O3’ distance within 4Å; (B) Comparison of the hydrogen bonding probability for the base pair at the active site.
Figure S2. Thermodynamic cycle for estimating the binding free energy of analog relative to that of ATP ($\Delta \Delta G$).
Figure S3. Examination over the hydrogen bonds and salt-bridge interactions around the nucleotide analogs at the active site. (A) A diagram showing the hydrogen bonds and salt-bridge interactions; (B) The probability to form hydrogen bond between 2'-hydroxyl group of analogs and Asp623; (C) The distance for salt bridge between $P_\alpha$ group of analogs and Arg555; (D) The distance for salt bridge between $P_\beta$ group of analogs and Arg553; (E) The distance for salt bridge between $P_\gamma$ group of analogs and Arg553. MD conformations from 20x50ns simulations were used for the analysis.
Figure S4. PCA analysis for the global movement of SARS-CoV-2 RdRp. (A) The accumulated contribution of the top 10 PCs; (B) The porcupine plot for the top two PCs, where the nsp12, nsp7, nsp8a and nsp8b are shown in orange, blue, magenta and grey; (C) Projection of MD conformations onto the top two PCs in each individual system where the red color indicates a high population region while the blue color for a low-populated region. MD conformations from 20x50ns simulations were used for the analysis.
Figure S5. RMSF analysis of motifs A-G. On the left is a diagram showing the positions of motifs relative to the active site, with each motif displayed in a distinct color. On the right is the comparison of RMSF analysis for each system. MD conformations from 20x50ns simulations were used for the analysis.
Figure S6. Representative conformations of nucleotide analogs and Asp623 at the active site. The conformation of ATP was used as the reference and each nucleotide analog was aligned to it by the heavy atoms of Asp623 backbone. (A) Ribose of Clofarabine-TP (COP) (pink) presents an obvious angle torsion compared with ATP (gray). The sidechain of Asp623 in COP system is repelled away from the ribose. (B) Ribose of Didanosine-TP (DIP) (green) keeps the same orientation with that of ATP. (C)-(D) Ribose of Fludarabine-TP (FLP) (blue) and Vidarabine-TP (VDP) (purple) was tilted compared with ATP ribose (gray) due to the abnormal orientation of their 3'-hydroxyl group. (E)-(F) Ribose of 2'-NH$_2$-dATP (BNP) and 2',3'-Didehydro-2',3'-dideoxyadenosine-TP (STP) maintain a similar orientation as that in ATP.
Figure S7. Representative conformation of Asp623 and Lys621 when ATP, Didanosine-TP (DIP) or 2’,3’-Didehydro-2’,3’-dideoxyadenosine-TP (STP) is at the active site. (A) Conformation with Asp623 interacting with Lys621 when ATP is located at the active site. (B) The same as (A) but with DIP at the active site. (C) The same as (A) but with STP at the active site.
**Figure S8.** The base modification of Fludarabine (FLP) on the C2 atom increases the distance between base paired nucleotides. (A) The distance between C2 atom on the base of analogs and O2 atom on the base of the template nucleotide was calculated in (B)-(G). (B)-(G) The distribution of the distance defined in (A) for six analogs. The distance between the base of analogs and that of the template nucleotide increases due to the repulsion caused by their base modifications. MD conformations from 20x50ns simulations were used for the analysis.
Figure S9. Modification on 2’-ribose of Clofarabine (CO3) at i+1 site partially impairs the nucleotide addition at the active site. (A) Diagram of distance between the analog at i+1 site and ATP at i site. (B)-(E) The distribution of the distance between C2’ atom on the ribose of four analogs (Clofarabine (CO3), Fludarabine (FL3), Vidarabine (VD3) and 2’-NH₂-dA (BN3)) at i+1 site and O4’ atom on the ribose of the ATP at active site. Fluorine modification on the 2’-ribose in the CO3 (pink) disrupted this distance due to its notable electronegativity. In (B)-(D), MD conformations from 20x100ns simulations were used for the analysis. In (E), the MD conformations from 20x50ns simulations were used for the analysis. The relative shift observed in (B) could be due to the polar repulsion between the 2’-F of CO3 and O4’ of ATP, while that observed in (C)-(E) could be due to the configurational adjustment to adapt to the variation in the hydrogen bonding orientation formed between 2’-modification and O4’ of ATP.
Figure S10. Comparing the performance of Remdesivir-TP (RDV-TP) with that of ATP, Didanosine-TP (DIP) and Cordycepin-TP (CRP) in the active site of SARS-CoV-2 RdRp. (A) The chemical structures of Adenosine, Remdesivir, Didanosine and Cordycepin in the nucleoside form. (B) The population of MD conformations showing P$_\alpha$-O$_{3'}$ distance within 4 Å; (C) The hydrogen bonding probability for the base pair at the active site; (D) The RMSF of ATP and nucleotide analogs in the triphosphate form in the active site; (E) Histogram of the dihedral angle between the ribose ring of ATP/nucleoside analogs and Asp623. The results in (B)-(E) were calculated using 20x100ns MD simulations for each analog.