Biochemistry

Article

Structural Insights into Binding of Remdesivir Triphosphate within the Replication−Transcription Complex of SARS-CoV-2

Jimin Wang,* Yuanjun Shi, Krystle Reiss, Federica Maschietto, Elias Lolis, William H. Konigsberg, George P. Lisi, and Victor S. Batista*

Cite This: Biochemistry 2022, 61, 1966−1973

Read Online

ACCESS | 📊 Metrics & More | 📖 Article Recommendations | 🆂🅸 Supporting Information

ABSTRACT: Remdesivir is an adenosine analogue that has a cyano substitution in the C1′ position of the ribosyl moiety and a modified base structure to stabilize the linkage of the base to the C1′ atom with its strong electron-withdrawing cyano group. Within the replication−transcription complex (RTC) of SARS-CoV-2, the RNA-dependent RNA polymerase nsp12 selects remdesivir monophosphate (RMP) over adenosine monophosphate (AMP) for nucleotide incorporation but noticeably slows primer extension after the added RMP of the RNA duplex product is translocated by three base pairs. Cryo-EM structures have been determined for the RTC with RMP at the nucleotide-insertion (i) site or at the $i + 1$, $i + 2$, or $i + 3$ sites after product translocation to provide a structural basis for a delayed-inhibition mechanism by remdesivir. In this study, we applied molecular dynamics (MD) simulations to extend the resolution of structures to the measurable maximum that is intrinsically limited by MD properties of these complexes. Our MD simulations provide (i) a structural basis for nucleotide selectivity of the incoming substrates of remdesivir triphosphate over adenosine triphosphate and of ribonucleotide over deoxyribonucleotide, (ii) new detailed information on hydrogen atoms involved in H-bonding interactions between the enzyme and remdesivir, and (iii) direct information on the catalytically active complex that is not easily captured by experimental methods. Our improved resolution of interatomic interactions at the nucleotide-binding pocket between remdesivir and the polymerase could help to design a new class of anti-SARS-CoV-2 inhibitors.
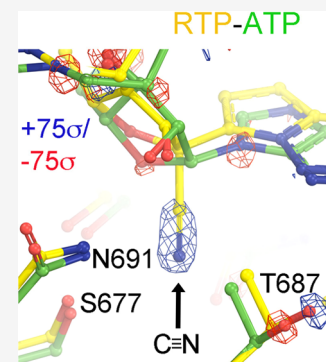
INTRODUCTION

Remdesivir (RDV) is an adenosine analogue with a cyano-group substitution at the C1′ position of the nucleotide base.[1,2] It has been used as a viral suppressor for treatment of SARS-CoV-2 infections although it shows only modest efficacy.[3−5] Nonetheless, there is great interest in understanding its binding mode and reaction mechanism. RDV-triphosphate (RTP) is a substrate of the RNA polymerase (pol) but not an inhibitor because it can be efficiently added to the growing RNA chain by the nsp12 RNA polymerase of SARS-CoV-2. Incorporation of RDV into the RNA occurs only after the prodrug RDV has been converted by the host cell into RTP.[6−10] RDV does not act as a chain terminator because it has a 3′-hydroxyl group, thereby allowing the primer to be efficiently extended by RNA pol after RDV-monophosphate (RMP) incorporation.[6−10] The rate of *in vitro* primer extension after RMP incorporation depends on the location of the added RMP and the concentration of the next incoming nucleotide triphosphate (NTP). When the added RMP in the RNA duplex product was translocated at the $i + 3$ site, an accumulation of this product was observed in the presence of a low concentration of incoming NTP but not at their physiological concentrations.[6] Thus, the reduced primer extension observed *in vitro* could also imply that the reduced viral replication and transcription *in vivo* is a mechanism known as delayed inhibition.[6] It could

also permit exoribonuclease to remove some of the added RMPs so that AMPs can be reinserted. This extra step of RMP insertion−excision could effectively slow down viral RNA synthesis.

Evidence exists that an exoribonuclease associated with both SARS-CoV and SARS-CoV-2 viruses could reduce a modest fraction of RDV inhibition,[11,12] suggesting that a significant fraction of added RDV nucleotides was not removed and thus has become part of the viral mRNAs. Nevertheless, RMP-containing viral RNAs are likely nonfunctional and thus are partially responsible for the suppression of viral reproduction.[13] In fact, the incorporation of RDV and its effects on viral RNA production in the SARS-CoV-2 infected cell lines is highly strand dependent.[12] Therefore, a new class of anti-SARS-CoV-2 inhibitors could be rationally designed, including the formation of dead-end inhibitory ternary complexes at the pol active site, translocation-inhibited complexes, or exoribonuclease-inhibited complexes. These designs all require higher-

resolution structural information than are available from the current cryo-EM structures.

In this study, we supplement high-resolution structural information using molecular dynamics (MD) simulations,[14] starting with medium- and low-resolution cryo-EM structures. Particularly, high-resolution structural information on the binding site of RTP can facilitate the rational design of effective drugs for targeting the viral RNA polymerase. Our structural analysis can also address many fundamental questions, including whether substitution of the RDV 3′-OH group by 3′-NH$_2$ or other functional groups could make the inhibitor a nonextendable chain terminator. If so, the modified RDV could terminate the RNA-synthesis complex after being incorporated. While MD simulations provide specific information about both kinetics and thermodynamics of enzymes in solution, we limit our analysis in this study to thermodynamics of pol complexes, i.e., the equilibrated structures in the MD simulation trajectories. Here, we focus on the analysis of the catalytically active complex of SARS-CoV-2 nsp12 pol in the presence of essential Mg^{2+} ions and the incoming RTP, or ATP base-paired with a template U nucleotide. These simulations provide unique and valuable structural information about the active complex that would be challenging to obtain from X-ray diffraction or cryo-EM methods. Our structural analysis suggests a specific mechanism for the incorporation of RMP or AMP by the nsp12 RNA pol of SARS-CoV-2. Our approach differs from recent MD simulations on remdesivir in different positions within the RNA duplex product that relied on free energy calculation in which the proposed remdesivir destabilizes the enzyme−RNA complex.[15] How those results correlate with experimental data remains unclear.

MATERIALS AND COMPUTATIONAL METHODS

The structure of the nsp12 RNA pol (7bv2) was retrieved from the Protein Data Bank (PDB) as a starting point of our MD simulations.[7] The bound RMP (F86) and pyrophosphate were removed and the 3′ end of the primer chain was truncated so that the active site was left open. Triphosphate versions of adenosine, remdesivir, and galidesivir (GLD) were placed at the active site with the triphosphate tail coordinated to two Mg^{2+} ions via nonbonding interactions. Mg^{2+} ions were also coordinated with formal bonds to D618, Y619, D760, and D761 using the Metal Center Parameters Builder (MCPB) from AmberTools.[16,17] A weak restraint (10 kcal/mol/Å^2 centered at 2.1 Å) was added to constrain the distance between Mg^{2+} ions and the deprotonated primer-terminal O3′ (ptO3′), accounting for the bond between those two atoms as MCPB does not currently support metal bonds to nucleic acid chains. MCPB was also used to create parameters for the two Zn^{2+} ions bound to three cysteines and a histidine. Parameters for the triphosphate molecules were generated with Antechamber.[17]

The system was immersed in a water box and neutralized with Na$^+$ ions. All MD simulations were performed using the Amber force field and NAMD.[17,18] Equilibration was performed in three steps, starting with equilibration of the solvent, then both solvent and sidechains, and finally the complete system. Each equilibration step consisted of minimization, gradual heating to 310 K, and thermalization. Once the system was stabilized at 310 K at the end of the thermalization step, a production run of 100 ns was performed. For the ATP-containing MD simulations, the coordinates were sampled at intervals of 20 ps per frame. For the RTP- and

GLD-TP-containing MD trajectories, coordinates were sampled at 40 ps per frame.

Individual frames of MD trajectories were aligned using the program VMD and saved as individual PDB files.[19] Electron scattering factors for neutral atoms were assigned for the calculation of ESP maps and X-ray scattering factors for the calculations of ED maps. All maps were combined using the CCP4 suite (the programs *sfall*, *fft*, and *mapsig* for adding one map per frame at a time).[20] The unit cell box for the ATP-containing structure was a cube with an edge length of 120 Å. The grid number was set to 360 (the grid size is 0.333 Å), and the initial ΔB = 8 Å^2 was used to minimize the Fourier series termination effect at the nominal resolution of 1.0 Å for calculations of both structure factors and maps. The preequilibrated 7bv2 model was used as a starting model to fit the MD-ESP (or MD-ED) maps.[7] The models were rebuilt using the graphics program Coot, following residual maps and refined using both Phenix and Refmac5 with standard iterations.[21−23] All figures were designed using PyMol.[24]

RESULTS AND DISCUSSION

MD-Derived Electrostatic Potential Maps Versus Cryo-EM Maps. Our analysis is based on the thermal distribution of atomic configurations sampled by MD simulations, providing (i) electron density (ED) maps computed with X-ray atomic scattering factors, (ii) electrostatic potential (ESP) maps obtained with electron atomic scattering factors (Figure 1), and (iii) probability density
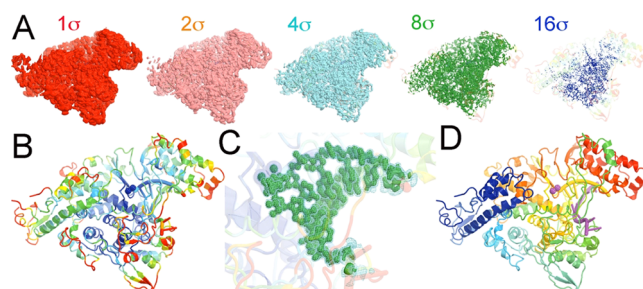

Figure 1. Overall comparison of MD-ESP maps with the experimental cryo-EM maps. (A) MD-ESP map with added blurring ΔB = 8 Å^2 contoured at +1 (red), +2 (salmon), +4 (cyan), +8 (green), or +16σ (blue) superimposed on the cartoon drawing of the refined atomic model colored by a rainbow spectrum of atomic B-factors refined in transparency. (B) Refined atomic model colored by the rainbow spectrum of atomic B-factors (blue−white−red for small−medium−large B-factors, respectively). (C) Close-up view of the MD-ESP map for the primer/template RNA duplex plus ATP contoured at +4σ (cyan mesh) and +8σ (green isosurface). (D) Refined atomic model colored by rainbow spectrum of residue sequence numbers.

functions (PDFs) of atomic coordinates. Therefore, the MD trajectories provide both the variance of atomic distributions (i.e., atomic B-factors) and the average atomic positions at thermal equilibrium (i.e., equilibrated structures). When there is a single conformation for a protein residue or nucleic acid base, the individual atomic coordinate distributions can be described to a good approximation by three-dimensional Gaussian functions. When there are multiple conformations, the distributions are described by a superposition of multiple overlapping Gaussian functions. In this study, electron scattering factors of neutral atoms were used. Partial charges of ionizable atoms, highlighting differences between exper-
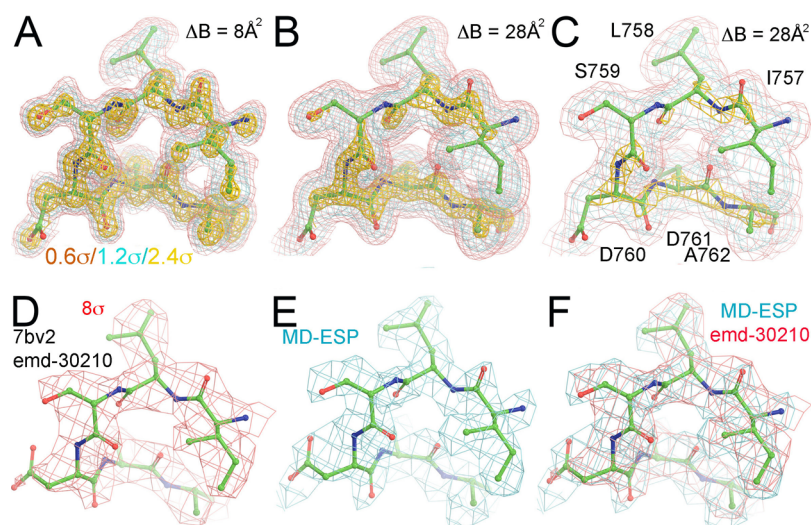
Figure 2. Detailed comparison of MD-ESP maps with the experimental cryo-EM maps. (A) Close-up view of the MD-ESP map for a catalytic loop ($\Delta B = 8$ Å^2) contoured at $+0.6\sigma$ (salmon), $+1.2\sigma$ (cyan), and $+2.4\sigma$ (gold). (B) Same catalytic loop of the MD-ESP map as panel (A) but with $\Delta B = 28$ Å^2. (C) Same as panel (B) but with coarse grid numbers corresponding to about 2.0 Å resolution. (D) Experimental emd-30120 map contoured $+8\sigma$. (E) Resampled MD-ESP map on the experimental map grids. (F) Superimposition of MD-ESP maps with the emd-30210 map and reported atomic model.

imental cryo-EM maps and MD-simulated ESP maps, properly account for the actual partial atomic charges.

The RNA pol active site represents the most rigid part of this enzyme complex, with the smallest distribution variance of atomic coordinates as revealed by MD-derived ESP maps. Figure 1 shows the overall ESP map derived from 5000 frames sampled from 100 ns MD trajectories (MD-ESP) and its comparison to the closely related experimental emd-30120 map.[7] This map was generated after filtering the MD-ESP map to comparable resolution and resampling onto the experimental grids.[7] The MD-ESP map was averaged over 20 ps per frame and superimposed on the equilibrium positions of the atomic model that was refined against the MD-ESP map.

The ensemble of configurations of the RNA pol complex sampled by MD simulations represents the highest achievable resolution of an error-free image reconstruction. The variance of the distribution resulting from thermal atomic motion determines the atomic B-factor and defines the local resolution. The larger the variance (and the B-factors), the broader the atomic ESP peak distribution, and the lower the amplitude of the atomic ESP. When the MD-ESP map is contoured at $+1$, $+2$, $+4$, $+8$, or $+16\sigma$ in the isosurface ESP representation, the region that is visible in the highest contour level of $+16\sigma$ represents the highest local resolution, while the regions that are only visible in the lowest contour level of $+1\sigma$ represent the lowest local resolution (Figure 1A). The pol active site is the region with highest resolution, while the peripheral loops on the surface of the complex correspond to the regions with lowest local resolution (Figure 1B). Clearly, the distribution of local resolution in MD-ESP maps is fully consistent with the distribution of refined atomic B-factors in the complex at thermal equilibrium.

We filtered the resolution of the MD-ESP map by adding blurring B-factors ($\Delta B = 28$ Å^2) to compare the MD-derived ESP maps with cryo-EM maps, and we resampled the MD-ESP onto the grids of the cryo-EM emd-30210 map (to ~2.5 Å) as reported for the 7bv2 structure (Figure 1).[7] The comparison shows that the two maps are remarkably similar, except for the orientation of the carboxylates of D760 and D761 (Figure 2).

However, modeling the orientation of carboxylate sidechains at medium resolution can be challenging, particularly when the carboxylate groups are bound to divalent metal ions.[25−28] Furthermore, we note that the experimental map corresponds to the enzyme−RNA product complex with incorporated RMP and bound pyrophosphate whereas our MD-ESP map corresponds to the ATP substrate complex. Nevertheless, these differences do not appreciably affect the structure of the nsp12 RNA pol except for the triphosphate moiety-binding pocket.[7]

Experimentally, it is difficult to capture the catalytically active complex that represents a minor population in the reaction mixture without using specific methods to enrich its population. To experimentally probe the catalytically relevant complex(es), it would be necessary to stall the RNA synthesis using, for example, transition-state analogue inhibitors, a 3′-deoxy nonextendable primer, a nonhydrolyzable nucleotide analogue (such as α,β-amide-substituted ATP or AMPNPP), or catalytically inactive divalent metal ions (such as Ca^{2+}), as reported for RB69 DNA pol and human pol β (and to a lesser extent, for other DNA pols).[29−38] Another method is reliant on certain crystal lattice constraints to trap the pyrophosphate product and a third divalent metal ion.[39,40] However, one must be aware of the fact that any of these methodologies could have artifacts and are prone to misinterpretation.[41,42]

Binding of ATP, RTP, and Galidesivir-Triphosphate. Figure 3 shows the MD-derived electron density (MD-ED) maps for the pol active site of SARS-CoV-2, containing the nascent base pair between a template nucleotide U and each of the incoming substrates: ATP, RTP, or Galidesivir (GLD)-triphosphate (GLD-TP). RTP has a cyano substitution at the C1′ position while GLD-TP has an amine substitution at O4′ (Figure 3B,C). Upon RTP binding, the cyano-binding pocket expands only slightly as seen in the outward displacement of the sidechain of T687 (Figure 3G). The equilibrium conformation of RTP differs from the binding mode of ATP and GLD-TP in that it is missing an intramolecular hydrogen bond (HB) between its O3′ and the β-phosphate group (Figure 3A−C). In the ATP-bound complex, a cluster of
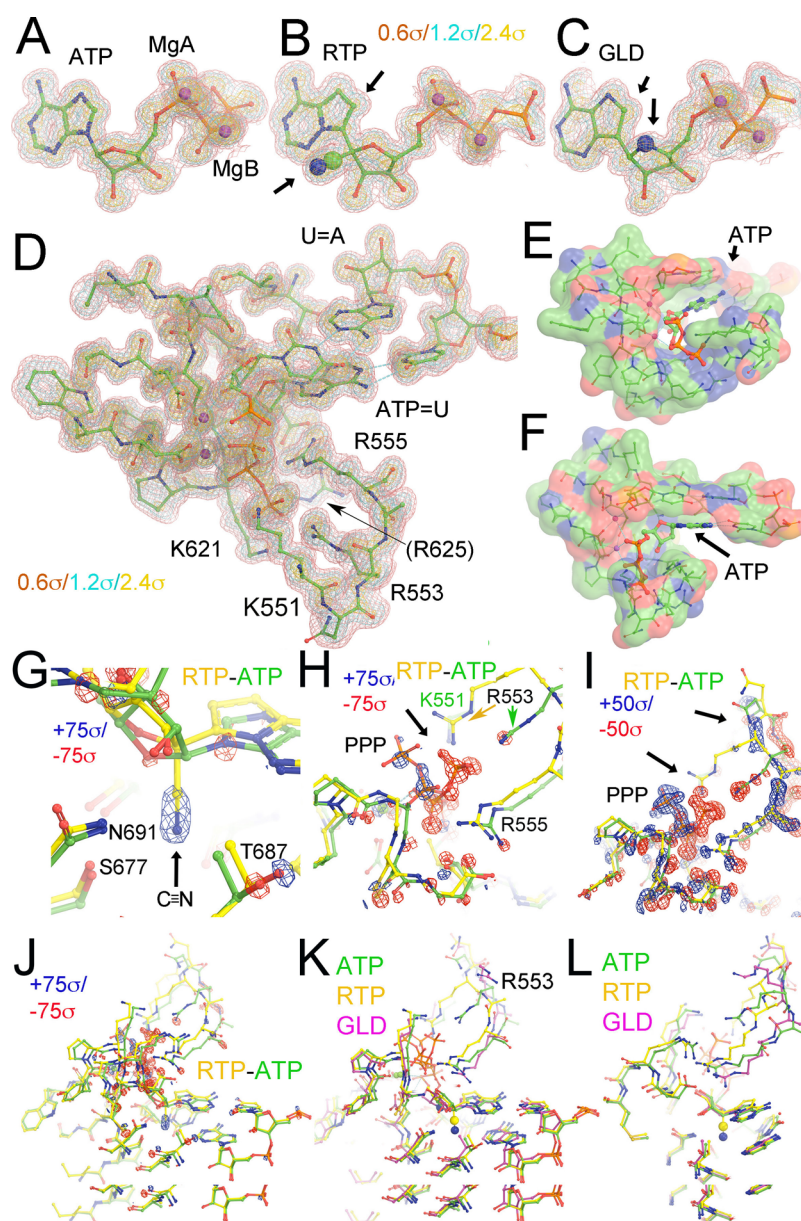
Figure 3. MD-ED maps for ATP-, RTP-, or GLD-TP-containing complex and their differences. Maps for (A) ATP, (B) RTP, and (C) GLD-TP were contoured at +0.6σ (salmon), +1.2σ (cyan), and +2.4σ (gold). Substituted groups are in large CPK model as indicated by black arrows. (D) Pol active site including the nascent base pair and primer-terminal base pair. (E, F) Surface representation of the ATP-binding pocket colored by charges (negative, red; positive, blue; and neutral, green). (G−J) Four different views of the RTP minus ATP difference map at the highest contouring levels (+75σ, blue/−75σ, red; G, H, and J) show the added cyano substitution in RTP relative to ATP (G, it should be noted that only relative values are meaningful because MD-ED maps are contoured relative to standard deviation of the entire maps and the absolute value of their standard deviation is highly dependent on the size of an arbitrary box used for MD simulations and also on sampling resolution, unlike crystallographic ED maps), (H) displaced triphosphate moiety, and at the second highest contouring levels (I, +50σ, blue/−50σ, red), and (J) zoom-out view. Arrows indicate displacements of the triphosphate tail and its interacting loop. (K, L) Two views of comparison of ATP (green), RTP (gold), and GLD-TP (magenta) bound structures.

positively charged amino acid residues form a well-defined binding pocket for the triphosphate moiety of ATP and exhibit well-ordered MD-ED features (Figure 3D−F), of which four directly interact with the triphosphate moiety (Figure 4). The binding mode observed suggests that ATP adopts a single conformation at the binding site. The same positive residues also bind the triphosphate moiety of RTP and GLD-TP. However, the triphosphate groups of RTP and GLD-TP exhibit multiple conformations so their MD-ED features are only approximately defined when compared to the ATP complex. Here, we describe only the dominant conformations

even though RTP and GLD-TP exhibit multiple conformations. In the ATP-binding pocket, ordered water molecules occupy the pocket where the substituted cyano group of RTP occupies it in the RTP complex (Figure 5).

The RTP-*minus*-ATP MD-ED difference map, when comparing RTP to ATP binding, shows that the largest positive feature corresponds to the cyano substitution of RTP at the C1′ position (Figure 3G). The largest negative difference corresponds to the triphosphate moiety, due to the different conformation of the nucleotide triphosphate in the two structures (Figure 3G). Associated with these
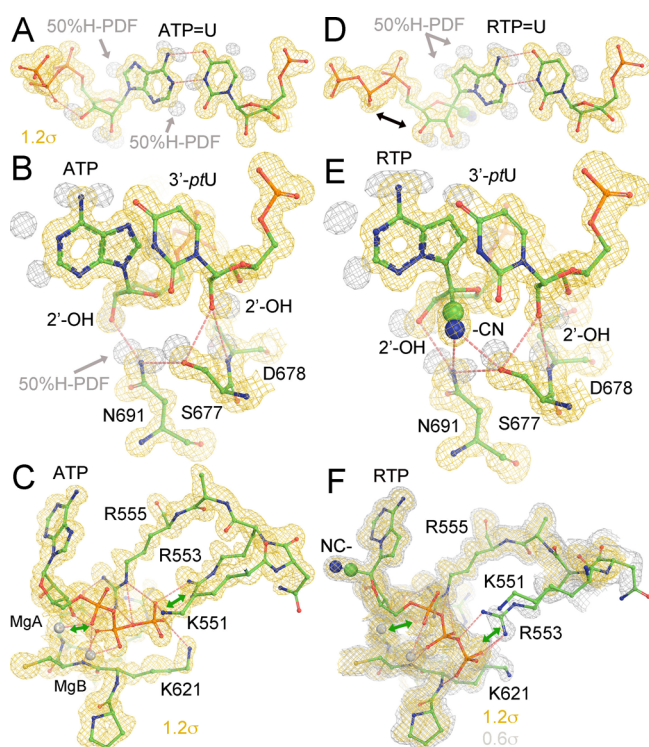
Figure 4. HB patterns for the recognition of 2′-OH and cyano groups in (A−C) ATP and (D−F) RTP complexes, respectively. MD-ED maps are contoured at low (+0.6σ, silver, in panel (F)) and high levels (+1.2σ, gold, all panels). H atom probability density function (PDF) maps are contoured at about 50% probability.

structural differences, there are many other MD-ED difference features due to the repositioning of five positively charged sidechains of the key amino acid residues and the

corresponding main chain backbone, suggesting correlated motions and different conformations of the triphosphate moiety for the three incoming NTPs. Therefore, the MD-ED difference Fourier analysis is very sensitive to any subtle structural change.

The conformation of the triphosphate moiety of GLD-TP resembles that of ATP. However, the MD-ED features of the triphosphate group of RTP and GLD-TP are less well-defined than those of ATP (Figures 3 and 4). In all three cases, the nucleobase and α-phosphate group exhibit distinct MD-ED features (Figure 3) as well as distinct dynamics in response to the different incoming nucleoside triphosphates. In addition, we have repeated all of our MD simulations, starting with the 6xez coordinates,[43] which largely reproduced the overall structural features reported in this study. However, these replicas also showed that there were more than two alternate conformations of the triphosphate-binding loops for substrates described here.

H-Bonding Interactions of RTP within the Polymerase Complex. MD simulations allow us to analyze detailed H-bond (HB) networks that are difficult to determine by simple inspection of medium- and low-resolution experimental maps. HB analysis provides valuable insights into the structural and functional roles of the HB network and shows that it involves the sidechain of S677, the 2′ and 3′ hydroxyl groups, and pol backbone. The 2′-OH groups of both ATP and the 3′ primer-terminal nucleotide U (*pt*U) are recognized by the pol via HB interactions (Figure 4). The N691 sidechain and the backbone amide of D678 provide selectivity against 2′-deoxy nucleotides, ensuring that this pol only accommodates RNA and discriminates against DNA (Figure 4B,E). We note that the HB between N691 and S667 can also keep these two nucleotides adjacent to each other, which is essential for the RNA polymerization reaction.
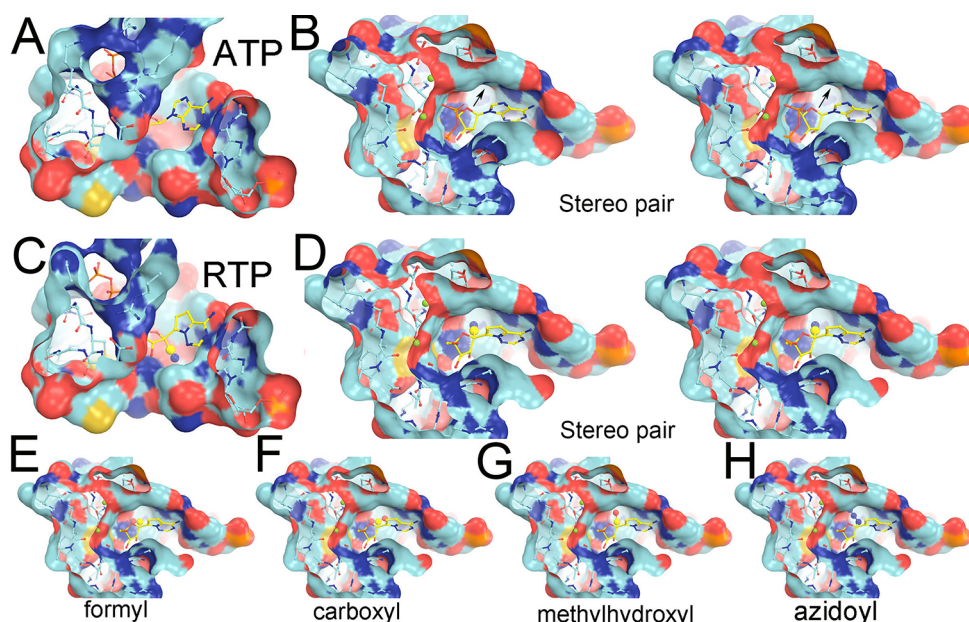

Figure 5. ATP-binding pocket and the preexisting binding pocket for the cyano group of RTP and hypothetical models for new drug design. (A, B) Two views of the ATP-bound complex show the preexisting pocket underneath the C1′ atom where the cyano group is substituted in the RTP complex. (B) In stereodiagram. Arrows indicate a preexisting pocket in the pol. (C, D) Same two views of the RTP-bound complex. (D) In stereodiagram. (E−H) With hypothetical formal, carboxyl, methylhydroxyl, or azidoyl substitution at C1′. Solvent-accessible surface is colored by negative charge (red), positive charge (blue), uncharged (cyan) atoms, and sulfur atom (gold).

ATP and RTP form two nearly identical base-pairing HBs with the template nucleotide U (Figure 4). Recognition of the 2′-OH group of RTP and *pt*U in the RTP complex is nearly identical to that of the ATP complex and is not affected by differences in the MD-ED maps due to the orientations of the 2′-OH and 3′-OH groups in RTP and ATP. The cyano group exhibits an HB interaction with the sidechain of S677 and a polar interaction with the sidechain of N691 (Figure 4E). These interactions increase the binding affinity of RTP in the pol-substrate complex. The same base-pairing interaction between GLD-TP and the template is observed between ATP and the template U.

DNA pols use a "sugar-gate" residue to stereochemically prevent rNTPs with 2′-OH groups from binding to the dNTP-binding pocket.[44,45] RNA pols have a specific interacting residue(s) to establish the HB that recognizes the 2′-OH group, enabling them to select rNTPs over dNTPs.[46−48] However, the detailed HB pattern responsible for introducing selectivity for the substrate has remained unknown due to the limited resolution of cryo-EM maps of the RTC of SARS-CoV-2.[7,49−51] Here, we find evidence that the sidechain of N691 recognizes the 2′-OH group of the incoming ATP (and other NTP substrates). The orientation of the N691 sidechain is constrained by a HB between the sidechains of N691 and S677. In addition, the preexisting pocket underneath the C1′ atom of ATP is filled by ordered water molecules, providing a strong cooperative effect for recognition of the RTP substrate as observed in our MD simulations. The HB network also assists in the recognition of the 2′-OH of the 3′ primer-terminal nucleotide by the backbone amide of D678 (Figure 4B,E).

The preexisting cyano-binding pocket is also observed in the ATP complex although filled with water molecules and replaced by the hydrophilic cyano group of remdesivir in the RTP complex (Figure 5A−D). The size and hydrophilic nature of the pocket also make it suitable for other types of polar moieties such as formyl, carboxyl, methylhydroxyl, keto, or amide groups that could substitute for the cyano group of C1′ (Figure 5E−H). Also, with modest puckering of the nucleotide sugar, the pocket could also fit an azidoyl group (Figure 5H). Additional MD simulations with these new compounds are ongoing.

Our analysis provides new insights on specific interactions responsible for binding of nucleotide analogues that could mimic ATP and outperform remdesivir. Therefore, our structural insight is expected to stimulate further studies of modified nucleotides with new functional groups designed for high-affinity binding at the nascent base-pair-binding pocket.

Relationship between MD Simulations and Biophysical Studies. MD simulations and simulated annealing have been the standard procedures of model refinement for both X-ray crystallography and NMR for many decades because they can rapidly find the equilibrated structure present in crystals or in solution starting with only moderately accurate coordinates.[52,53] Slow-cooling procedures are commonly used for finding a single equilibrated structure. We provide an alternative in this study. The equilibrated structures derived from MD-ESP or MD-ED maps can also be used for interpretation of corresponding medium- and low-resolution cryo-EM ESP maps and crystallographic ED maps. We present evidence that MD simulations can complement high-resolution structural information for these structures because they provide MD-ESP and MD-ED maps that are highly

comparable with the experimental data. Figure 1 shows a striking similarity between the MD-ESP and cryo-EM map when they are compared locally. However, during a global comparison, we discovered that the magnification used for image reconstruction was often only approximate and required additional correction for voxel rescaling between a global comparison of MD-ESP and cryo-EM maps to be made reliably.

MD simulations are powerful theoretical single-molecule techniques and should be comparable to experimental single-molecule spectroscopy. Because single-molecule kinetic behaviors are fully reproduced by the kinetic behaviors of bulk solution assembly complexes (i.e., Michaelis−Menten kinetics for complex enzymatic reactions),[54] one expects that MD simulations should reproduce nearly all macroscopically measurable kinetic and thermodynamic properties of any enzyme, even though this study focuses on only the equilibrated structures for the purpose of drug discovery. At the single-molecule level, enzymes fluctuate stochastically, and this information should be present in the MD simulation trajectories alongside information for the kinetic rates of conversion between any pair of specific functional states.[55−58] With assistance of MD simulations between two closely functional states, one may be able to deconstruct such states from experimental maps when they are inadvertently merged into a single state. Therefore, this study is just an example but it opens up a new avenue for powerful MD simulation-based rationale for drug design.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.2c00341.

Refined PDB coordinates of the equilibrated ATP-bound complex against MD-derived ESP maps (PDB)

■ AUTHOR INFORMATION

Corresponding Authors

Jimin Wang − *Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114, United States;* ⓞ orcid.org/0000-0002-4504-8038; Email: jimin.wang@yale.edu

Victor S. Batista − *Department of Chemistry, Yale University, New Haven, Connecticut 06520-8499, United States;* ⓞ orcid.org/0000-0002-3262-1237; Email: victor.batista@yale.edu

Authors

Yuanjun Shi − *Department of Chemistry, Yale University, New Haven, Connecticut 06520-8499, United States*

Krystle Reiss − *Department of Chemistry, Yale University, New Haven, Connecticut 06520-8499, United States*

Federica Maschietto − *Department of Chemistry, Yale University, New Haven, Connecticut 06520-8499, United States;* ⓞ orcid.org/0000-0002-5995-2765

Elias Lolis − *Department of Pharmacology, Yale University, New Haven, Connecticut 06520-8066, United States;* ⓞ orcid.org/0000-0002-7902-7868

William H. Konigsberg − *Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114, United States*

George P. Lisi − *Department of Molecular and Cell Biology and Biochemistry, Brown University, Providence, Rhode Island 02912, United States;* ● orcid.org/0000-0001-8878-5655

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.biochem.2c00341

Author Contributions

J.W. and V.S.B. designed and supervised the research; Y.S. and K.R. performed MD simulations; Y.S., K.R., F.M., and J.W. analyzed the data; J.W. wrote the draft paper with input from W.H.K., G.P.L., V.S.B., K.R., F.M., Y.S., E.L.; and J.W., W.H.K., and V.S.B. wrote the final version of the manuscript.

Funding

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

RdRp, RNA-dependent RNA polymerase; pol, polymerase; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; P/T, primer/template; RDV, remdesivior; RMP, remdesivir monophosphate; RTP, remdesivir triphosphate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; GLD, galidesivir; GLD-TP, galidesivir triphosphate; MD, molecular dynamics; MD-ESP, MD-derived electrostatic potential maps; MD-ED, MD-derived electron density maps; PDF, probability density function of atomic distribution; HB, hydrogen bond or hydrogen bonding

■ REFERENCES

(1) Seley-Radtke, K. L.; Yates, M. K. The evolution of nucleoside analogue antivirals: A review for chemists and non-chemists. Part 1: Early structural modifications to the nucleoside scaffold. *Antiviral Res.* **2018**, *154*, 66−86.

(2) Yates, M. K.; Seley-Radtke, K. L. The evolution of antiviral nucleoside analogues: A review for chemists and non-chemists. Part II: Complex modifications to the nucleoside scaffold. *Antiviral Res.* **2019**, *162*, 5−21.

(3) Wang, Y.; Zhang, D.; Du, G. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial (vol 395, 1569, 2020). *Lancet* **2020**, *395*, 1694.

(4) Eastman, R. T.; Roth, J. S.; Brimacombe, K. R.; Simeonov, A.; Shen, M.; Patnaik, S.; Hall, M. D. Remdesivir: A review of its discovery and development leading to emergency use authorization for treatment of COVID-19. *ACS Cent. Sci.* **2020**, *6*, 672−683.

(5) Grein, J.; Ohmagari, N.; Shin, D.; Diaz, G.; Asperges, E.; Castagna, A.; Feldt, T.; Green, G.; Green, M. L.; Lescure, F. X.; Nicastri, E.; Oda, R.; Yo, K.; Quiros-Roldan, E.; Studemeister, A.; Redinski, J.; Ahmed, S.; Bernett, J.; Chelliah, D.; Chen, D.; Chihara, S.; Cohen, S. H.; Cunningham, J.; D'Arminio Monforte, A.; Ismail, S.; Kato, H.; Lapadula, G.; L'Her, E.; Maeno, T.; Majumder, S.; Massari, M.; Mora-Rillo, M.; Mutoh, Y.; Nguyen, D.; Verweij, E.; Zoufaly, A.; Osinusi, A. O.; DeZure, A.; Zhao, Y.; Zhong, L.; Chokkalingam, A.; Elboudwarej, E.; Telep, L.; Timbs, L.; Henne, I.; Sellers, S.; Cao, H.; Tan, S. K.; Winterbourne, L.; Desai, P.; Mera, R.; Gaggar, A.; Myers, R. P.; Brainard, D. M.; Childs, R.; Flanigan, T. Compassionate use of remdesivir for patients with severe Covid-19. *N. Engl. J. Med.* **2020**, *382*, 2327−2336.

(6) 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.

(7) Yin, W.; Mao, C.; Luan, X.; Shen, D. D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; Chang, S.; Xie, Y. C.; Tian, G.; Jiang, H. W.; Tao, S. C.; Shen, J.; Jiang, Y.; Jiang, H.; Xu, Y.; Zhang, S.; Zhang, Y.; Xu, H. E. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020**, *368*, 1499−1504.

(8) Tchesnokov, E. P.; Feng, J. Y.; Porter, D. P.; Gotte, M. Mechanism of inhibition of Ebola virus RNA-dependent RNA polymerase by remdesivir. *Viruses* **2019**, *11*, No. 326.

(9) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* **2020**, *181*, 894−904.e9.

(10) Dangerfield, T. L.; Huang, N. Z.; Johnson, K. A. Remdesivir is effective in combating COVID-19 because it is a better substrate than ATP for the viral RNA-dependent RNA polymerase. *iScience* **2020**, *23*, No. 101849.

(11) Agostini, M. L.; Andres, E. L.; Sims, A. C.; Graham, R. L.; Sheahan, T. P.; Lu, X.; Smith, E. C.; Case, J. B.; Feng, J. Y.; Jordan, R.; Ray, A. S.; Cihlar, T.; Siegel, D.; Mackman, R. L.; Clarke, M. O.; Baric, R. S.; Denison, M. R. Coronavirus susceptibility to the antiviral remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease. *mBio* **2018**, *9*, No. e00221-18.

(12) Zhao, Y.; Sun, J.; Li, Y.; Li, Z.; Xie, Y.; Feng, R.; Zhao, J.; Hu, Y. The strand-biased transcription of SARS-CoV-2 and unbalanced inhibition by remdesivir. *iScience* **2021**, *24*, No. 102857.

(13) Wang, J.; Reiss, K.; Shi, Y.; Lolis, E.; Lisi, G. P.; Batista, V. S. Mechanism of inhibition of the reproduction of SARS-CoV-2 and Ebola Viruses by remdesivir. *Biochemistry* **2021**, *60*, 1869−1875.

(14) Wang, J.; Shi, Y.; Reiss, K.; Allen, B.; Maschietto, F.; Lolis, E.; Konigsberg, W. H.; Lisi, G. P.; Batista, V. S. Insights into binding of single-stranded viral RNA template to the replication-transcription complex of SARS-CoV-2 for the priming reaction from molecular dynamics simulations. *Biochemistry* **2022**, *61*, 424−432.

(15) Byléhn, F.; Menendez, C. A.; Perez-Lemus, G. R.; Alvarado, W.; de Pablo, J. J. Modeling the binding mechanism of remdesivir, favilavir, and ribavirin to SARS-CoV-2 RNA-dependent RNA polymerase. *ACS Cent. Sci.* **2021**, *7*, 164−174.

(16) Li, P.; Merz, K. M., Jr. MCPB.py: A python based metal center prameter builder. *J. Chem. Inf. Model.* **2016**, *56*, 599−604.

(17) Case, D. A.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER 2017*; University of California: San Francisco, 2017.

(18) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Henin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kale, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, No. 044130.

(19) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

(20) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235−242.

(21) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.

(22) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240−255.

(23) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213−221.

(24) Delano, W. L. *Pymol*; Schrodinger, Inc. http://pymol.org/, software downloaded on March, 2020.

(25) Wang, J.; Moore, P. B. On the interpretation of electron microscopic maps of biological macromolecules. *Protein Sci.* **2017**, *26*, 122−129.

(26) Wang, J.; Natchiar, S. K.; Moore, P. B.; Klaholz, B. P. Identification of Mg(2+) ions next to nucleotides in cryo-EM maps using electrostatic potential maps. *Acta Crystallogr., Sect. D: Struct. Biol.* **2021**, *77*, 534−539.

(27) Wang, J.; Liu, Z.; Frank, J.; Moore, P. B.; et al. Identification of ions in experimental electrostatic potential maps. *IUCrJ* **2018**, *5*, 375−381.

(28) Wang, J. On the appearance of carboxylates in electrostatic potential maps. *Protein Sci.* **2017**, *26*, 396−402.

(29) Wang, M.; Xia, S.; Blaha, G.; Steitz, T. A.; Konigsberg, W. H.; Wang, J. Insights into base selectivity from the 1.8 A resolution structure of an RB69 DNA polymerase ternary complex. *Biochemistry* **2011**, *50*, 581−590.

(30) Xia, S.; Wang, M.; Blaha, G.; Konigsberg, W. H.; Wang, J. Structural insights into complete metal ion coordination from ternary complexes of B family RB69 DNA polymerase. *Biochemistry* **2011**, *50*, 9114−9124.

(31) Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure* **2006**, *14*, 757−766.

(32) Franklin, M. C.; Wang, J.; Steitz, T. A. Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* **2001**, *105*, 657−667.

(33) Sawaya, M. R.; Pelletier, H.; Kumar, A.; Wilson, S. H.; Kraut, J. Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. *Science* **1994**, *264*, 1930−1935.

(34) Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J. Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* **1994**, *264*, 1891−1903.

(35) Doublié, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. *Nature* **1998**, *391*, 251−258.

(36) Huang, H.; Chopra, R.; Verdine, G. L.; Harrison, S. C. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **1998**, *282*, 1669−1675.

(37) Li, Y.; Korolev, S.; Waksman, G. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J.* **1998**, *17*, 7514−7525.

(38) Li, Y.; Kong, Y.; Korolev, S.; Waksman, G. Crystal structures of the Klenow fragment of *Thermus aquaticus* DNA polymerase I complexed with deoxyribonucleoside triphosphates. *Protein Sci.* **1998**, *7*, 1116−1123.

(39) Zamyatkin, D. F.; Parra, F.; Alonso, J. M.; Harki, D. A.; Peterson, B. R.; Grochulski, P.; Ng, K. K. Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase. *J. Biol. Chem.* **2008**, *283*, 7705−7712.

(40) Zamyatkin, D. F.; Parra, F.; Machin, A.; Grochulski, P.; Ng, K. K. Binding of 2′-amino-2′-deoxycytidine-5′-triphosphate to norovirus polymerase induces rearrangement of the active site. *J. Mol. Biol.* **2009**, *390*, 10−16.

(41) Park, J.; Youn, H. S.; An, J. Y.; Lee, Y.; Eom, S. H.; Wang, J. Structure of New Binary and Ternary DNA Polymerase Complexes From Bacteriophage RB69. *Front. Mol. Biosci.* **2021**, *8*, No. 704813.

(42) Wang, J.; Konigsberg, W. H. Two-metal-ion catalysis: Inhibition of DNA polymerase activity by a third divalent metal ion. *Front. Mol. Biosci.* **2022**, *9*, No. 824794.

(43) Chen, J.; Malone, B.; Llewellyn, E.; Grasso, M.; Shelton, P. M. M.; Olinares, P. D. B.; Maruthi, K.; Eng, E. T.; Vatandaslar, H.; Chait, B. T.; Kapoor, T. M.; Darst, S. A.; Campbell, E. A. Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. *Cell* **2020**, *182*, 1560−1573.e13.

(44) Brown, J. A.; Suo, Z. Unlocking the sugar ″steric gate″ of DNA polymerases. *Biochemistry* **2011**, *50*, 1135−1142.

(45) Joyce, C. M.; Steitz, T. A. Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.* **1994**, *63*, 777−822.

(46) Su, Y.; Egli, M.; Guengerich, F. P. Mechanism of ribonucleotide incorporation by human DNA polymerase eta. *J. Biol. Chem.* **2016**, *291*, 3747−3756.

(47) Vaisman, A.; Woodgate, R. Ribonucleotide discrimination by translesion synthesis DNA polymerases. *Crit. Rev. Biochem. Mol. Biol.* **2018**, *53*, 382−402.

(48) Yoon, H.; Warshel, A. The control of the discrimination between dNTP and rNTP in DNA and RNA polymerase. *Proteins* **2016**, *84*, 1616−1624.

(49) Hillen, H. S.; Kokic, G.; Farnung, L.; Dienemann, C.; Tegunov, D.; Cramer, P. Structure of replicating SARS-CoV-2 polymerase. *Nature* **2020**, *584*, 154−156.

(50) Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; Ge, J.; Zheng, L.; Zhang, Y.; Wang, H.; Zhu, Y.; Zhu, C.; Hu, T.; Hua, T.; Zhang, B.; Yang, X.; Li, J.; Yang, H.; Liu, Z.; Xu, W.; Guddat, L. W.; Wang, Q.; Lou, Z.; Rao, Z. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **2020**, *368*, 779−782.

(51) Wang, Q.; Wu, J.; Wang, H.; Gao, Y.; Liu, Q.; Mu, A.; Ji, W.; Yan, L.; Zhu, Y.; Zhu, C.; Fang, X.; Yang, X.; Huang, Y.; Gao, H.; Liu, F.; Ge, J.; Sun, Q.; Yang, X.; Xu, W.; Liu, Z.; Yang, H.; Lou, Z.; Jiang, B.; Guddat, L. W.; Gong, P.; Rao, Z. Structural basis for RNA replication by the SARS-CoV-2 polymerase. *Cell* **2020**, *182*, 417−428.e13.

(52) Rice, L. M.; Brunger, A. T. Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins* **1994**, *19*, 277−290.

(53) Stein, E. G.; Rice, L. M.; Brunger, A. T. Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation. *J. Magn. Reson.* **1997**, *124*, 154−164.

(54) Xie, X. S. Biochemistry. Enzyme kinetics, past and present. *Science* **2013**, *342*, 1457−1459.

(55) English, B. P.; Min, W.; van Oijen, A. M.; Lee, K. T.; Luo, G.; Sun, H.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nat. Chem. Biol.* **2006**, *2*, 87−94.

(56) Min, W.; English, B. P.; Luo, G.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. Fluctuating enzymes: lessons from single-molecule studies. *Acc. Chem. Res.* **2005**, *38*, 923−931.

(57) Sunney Xie, X. Single-molecule approach to dispersed kinetics and dynamic disorder: Probing conformational fluctuation and enzymatic dynamics. *J. Chem. Phys.* **2002**, *117*, 11024−11032.

(58) Yang, H.; Xie, X. S. Statistical approaches for probing single-molecule dynamics photon-by-photon. *Chem. Phys.* **2002**, *284*, 423−437.