Role of 1'-Ribose Cyano Substitution for Remdesivir to Effectively Inhibit both Nucleotide Addition and Proofreading in SARS-CoV-2 Viral RNA Replication

Lu Zhang[1,2]*, Dong Zhang[1,2], Congmin Yuan[3,4], Xiaowei Wang[3,4], Yongfang Li[1,2], Xilin Jia[1,2], Xin Gao[5], Hui-Ling Yen[6]*, Peter Pak-Hang Cheung[3,4]*, Xuhui Huang[3,4]*

[1]*State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, Fujian, China*

[2]*University of Chinese Academy of Sciences, Beijing, China*

[3]*Department of Chemistry, Centre of Systems Biology and Human Health, State Key Laboratory of Molecular Neuroscience, The Hong Kong University of Science and Technology, Kowloon, Hong Kong*

[4]*The Hong Kong University of Science and Technology-Shenzhen Research Institute, Hi-Tech Park, Nanshan, Shenzhen 518057, China*

[5]*Computational Bioscience Research Center, Computer, Electrical and Mathematical Sciences Engineering Division, King Abdullah University of Science and Technology, Saudi Arabia*

[6]*School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong*

*correspondence could be addressed to: luzhang@fjirsm.ac.cn (L. Z.) or hyen@hku.hk (H.-L. Y.) or ppcheung@ust.hk (P. P.-H. C.) or xuhuihuang@ust.hk (X. H.)

Abstract

COVID-19 has recently caused a global health crisis and an effective interventional therapy is urgently needed. SARS-CoV-2 RNA-dependent RNA polymerase (RdRp) provides a promising but challenging drug target due to its intrinsic proofreading exoribonuclease (ExoN) function. Nucleoside triphosphate (NTP) analogues added to the growing RNA chain should supposedly terminate viral RNA replication, but ExoN can cleave the incorporated compounds and counteract their efficacy. Remdesivir targeting SARS-CoV-2 RdRp exerts high drug efficacy *in vitro* and *in vivo*. However, its underlying inhibitory mechanisms remain elusive. Here, we performed all-atom molecular dynamics (MD) simulations with an accumulated simulation time of 12.6 microseconds to elucidate the molecular mechanisms underlying the inhibitory effects of remdesivir in nucleotide addition (RdRp complex: nsp12-nsp7-nsp8) and proofreading (ExoN complex: nsp14-nsp10). We found that the 1'-cyano group of remdesivir possesses the dual role of inhibiting both nucleotide addition and proofreading. For nucleotide addition, we showed that incorporation of one remdesivir is not sufficient to terminate RNA synthesis. Instead, the presence of the polar 1'-cyano group of remdesivir at an upstream site causes instability via its electrostatic interactions with a salt bridge formed by Asp865 and Lys593, rendering translocation unfavourable. This may eventually lead to a delayed chain termination of RNA extension by three nucleotides. For proofreading, remdesivir can inhibit cleavage via the steric clash between the 1'-cyano group and Asn104. To further examine the role of 1'-cyano group in remdesivir's inhibitory effects, we studied three additional NTP analogues with other types of modifications: favipiravir, vidarabine, and fludarabine. Our simulations suggest that all three of them are prone to ExoN cleavage. Our computational findings were further supported by an *in vitro* assay in Vero E6 cells using live SARS-CoV-2. The dose-response curves suggest that among tested NTP analogues, only remdesivir exerts significant inhibitory effects on viral replication. Our work provides plausible mechanisms at molecular level on how remdesivir inhibits viral RNA replication, and our findings may guide rational design for new treatments of COVID-19 targeting viral replication.

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 pandemic around the world. As of April 2020, COVID-19 has been reported in 210 countries and almost 3 million cases have been confirmed, resulting in more than 200,000 deaths with an estimated mortality risk of 6.91%[1]. The new CoV has been declared a global emergency by the World Health Organization, as the outbreak continues to spread globally and no clinically approved interventional therapy is currently available to curb this health crisis.

Antiviral agents are urgently needed to treat COVID-19 patients. It could take up years to develop new interventions, including therapeutic antibodies, cytokines, nucleic acid-based therapies, and vaccines. Hence, repurposing clinically approved or investigative drugs for other diseases provides a promising approach to develop COVID-19 treatments. CoV has a capsid that envelops the single-stranded RNA genome[2]. Three structural surface proteins are shown to be associated with the capsid: Viral membrane, envelope, and the spike protein[3]. Because SARS-CoV-2 and SARS-CoV share 89% nucleotide identity[4], chemical compounds and monoclonal antibodies that target SARS-CoV surface proteins or interrupt its binding to host viral receptor have been under investigation for treating SARS-CoV-2[5-7]. In addition, new vaccines targeting the viral surface antigens are under intense development for the prevention of COVID-19. Unfortunately, drugs and vaccines with inhibitory mechanisms targeting surface receptors may not be effective due to the constant evolution of surface receptors to acquire drug resistance and evade host immune response[8]. In contrast, viral RNA-dependent-RNA polymerase (RdRp) is a protein that is deeply buried inside the viral capsule and is functionally conserved for viral replication, rendering it resistant to the propagation of drug-resistant virus. Thus, RdRp serves as a promising drug target for virus infections[9]. Indeed, inhibitors targeting polymerase of SARS-CoV-2 is currently at late stages of clinical trials.

Compared with other viral RdRp such as influenza virus and rhinovirus, SARS-CoV-2 RdRp is more challenging for drug development due to its intrinsic proofreading exoribonuclease (ExoN) function. Non-structural protein (nsp) 12, along with cofactors nsp8 with nsp7, is involved in nascent RNA priming and nucleotide addition[10-16]. NTP analogues that inhibit nucleotide addition in nsp12 supposedly should inhibit RNA replication. However, in SARS-CoV-2, ExoN of nsp14 has been shown to play a pivotal role in proofreading[17]. To counteract the efficacy of the NTP analogues, nsp14 excises the incorporated NTP analogues[18-19]. For example, the NTP analogue ribavirin is being incorporated in the nascent RNA to inhibit nucleotide addition of SARS-CoV RdRp. However, ribavirin is readily excised from nascent RNA by nsp14, which may explain its limited efficacy *in vivo*[16].

Remdesivir is a promising drug candidate to treat COVID-19 infection[20-22]. It has been shown to be effective as compassionate-use basis to patients hospitalised with COVID-19[22], including those suffering from pneumonia[20]. Subsequently, clinical trials on using remdesivir to treat COVID-19 have been conducted[1]. Remdesivir has shown to be a potent inhibitor of CoV replication including the MERS-CoV, SARS-CoV, and circulating human CoVs[19, 23]. Even though remdesivir is an NTP analogue incorporated by RdRp, it exerts superior antiviral activity over other NTP analogues because the rate of incorporation of remdesivir to nascent RNA by nsp12 is higher than that of its cleavage by nsp14 ExoN[19]. Recently, enormous amount of efforts has been placed in understanding the molecular basis of remdesivir's inhibitory mechanisms on RNA synthesis[14-15, 24-27]. Due to the high sequence similarity between SARS and SARS-2, the nsp12-nsp7-nsp8 and nsp14-nsp10 of SARS-CoV serve as reliable models to study the mechanisms of RNA replication of SARS-CoV-2. Furthermore, the cryo-EM structures of SARS-CoV-2 nsp12-nsp7-nsp8 have been recently solved[14-15], with the catalytic

domain of nsp12 showing very high structural similarity to that of SARS-CoV[12, 16]. However, only one of these SARS-CoV polymerase structures contains the nascent and template strands, which is in the pre-translocation state with pyrophosphate ion (PPi) still bound at the active site[15]. Hence, molecular mechanisms, particularly the dynamics of the nucleotide addition cycle of either NTP or NTP analogues remain largely unknown. Biochemical assays have shown in MERS CoV[24] and recently in SARS-CoV-2[15, 25] that remedivir causes "delayed chain termination" in RNA synthesis, wherein several rounds of nucleotide additions can still proceed after remdesivir incorporation prior to completed termination. However, the molecular mechanisms underlying this delayed termination induced by remdesivir remain largely elusive. Furthermore, it is also not clear how remdesivir can evade the proofreading activity of SARS-CoV-2 to prevent itself from being cleaved by ExoN.

To elucidate the mechanisms of remdesivir inhibition on RdRp and ExoN, we conducted an aggregation of 12.6 μs all-atom molecular dynamics (MD) simulations of the nsp12-nsp7-nsp8 and nsp14-nsp10 complexes in SARS-CoV-2. For RdRp, our MD simulations show that remdesivir at the 3'-terminal can reduce the efficiency of nucleotide addition in nsp12. More interestingly, we found that remdesivir at an upstream site (*i+3* site) can block translocation via interactions formed by its 1'-cyano group and thus results in delayed chain termination. For ExoN, we found that the relatively bulky 1'-cyano group on the ribose of remdesivir introduces steric interactions and effectively disrupts the stability of the catalytic active conformation, and thus effectively inhibit the exonuclease activity of ExoN. Furthermore, our MD simulations of three additional NTP analogues containing other base and ribose modifications (favipiravir, vidarabine and fludarabine) suggest that they are all prone to be cleaved by ExoN, underlining the important role of the 1'-cyano substitution in remdesivir. Our simulation results are supported by an *in vitro* assay to examine the efficiency of NTP analogues in inhibiting SARS-CoV-2 live virus growth. We found that only remdesivir shows inhibitory effect, consistent with our MD simulations.

Results and Discussions

Remdesivir can reduce the efficiency of nucleotide addition in RdRp

Our structural model of SARS-CoV-2 RdRp (nsp12-nsp7-nsp8) has been constructed based on the cryo-EM structure of SARS-CoV RdRp (PDBID: 6NUR[12], see Fig. 1A). Template and nascent RNA strands have been constructed by the structural alignment of modelled SARS-CoV-2 RdRp to the norovirus RdRp (PDBID: 3H5Y[28], see Methods and SI Section 1.1 for details). Twenty replicas of 30-ns MD simulations were performed for the RdRp in the post-translocation state, in which the active site (*i* site) is occupied by adenosine triphosphate (ATP) (Fig. 1A and 1C, see SI Section 2.2.1 for details). To validate this model (named as "wildtype-RNA"), we first examined the stability of the active site by assessing the distance between the α phosphate (Pα) of ATP and O3' atom of the 3'-terminal nucleotide (Fig. 2A). Typically, such distance is required to be below 4 Å to allow efficient catalysis to form the phosphodiester bond (e.g. 3.4 Å in the structure of norovirus RdRp[28] and 3.5~4.0 Å suggested by previous computational studies[29-30]). Consistent with this distance requirement, we found that the distribution of the O3'-Pα distance peaks at ~3.5 Å (see the gray curve in Fig. 2A). Furthermore, the base-to-base distance between ATP and its base-paired template nucleotide was also measured as an estimation of base-pairing stability[31]. We found that such base-to-base distance is most populated at ~7.5 Å, similar to the distance of 7.4~7.6 Å as observed in structures of other viral polymerases[28, 32] (Fig. 2D). In addition, we investigated base stacking stability as another criteria for the propensity of nucleotide addition[31]. We found that the distance between the base of NTP and the base of 3'-terminal nucleotide is most populated at ~4.5 Å (Fig. 2G), consistent with the distance of 4.2 Å as observed in norovirus[28] and poliovirus[32] RdRps.

In addition to assessing the active site stability, we further validated our structural model against three recently solved cryo-EM structures of SARS-CoV-2 RdRp[14-15] (see Fig. 3 and SI Section 4.2 for details). For the two apo structures, the root-mean-square-deviation (RMSD) of Cα atoms was used as comparison to our model. We found that the averaged RMSDs are 2.36 Å (PDBID: 6M71)[14] and 2.30 Å (PDBID: 7BV1[15]), indicating that the overall architecture of our nsp12 model is similar to that of the apo cryo-EM structures of SARS-CoV-2. For the holo structure, the RMSDs of Cα atoms and nucleotides between our structural model and the holo cryo-EM structure (PDBID: 7BV2[15]) were 2.12 Å and 2.23 Å, respectively. This further validated our wildtype-RNA model. Finally, we also examined the interactions between upstream RNA and proteins, and found that most of the interactions (73%) observed in the holo cryo-EM structure are well maintained in our MD simulations (Fig. S1).

After validating our structural model, we next performed twenty 30-ns MD simulations with remdesivir-TP replacing ATP at the active site (*i* site) (see SI Section 2.2.2 for details). Interestingly, we found that the distribution of the critical O3'-Pα distance in the remdesivir-TP system is still largely below 4 Å, even though it is substantially broader compared to ATP (Fig. 2A). These results suggest that the incorporation of remdesivir-TP is still allowed but with a reduced efficiency compared to ATP. Furthermore, no obvious negative effect on base pairing and base stacking stability is induced by remdesivir-TP at *i* site (Fig. 2D and 2G). Altogether, our observations suggest that remdesivir-TP can largely maintain a catalytically active conformation, and thus can be added to the 3'-end of the nascent RNA chain even at a reduced efficiency. Therefore, remdesivir satisfies the pre-requisite condition to act as a chain-termination inhibitor.

After remdesivir is added to the RNA chain and translocated to the adjacent upstream site (*i+1* site), we found that remdesivir can reduce the efficiency of addition for the next nucleotide at the active site. In this case, we also performed twenty 30-ns MD simulations with remdesivir at 3'-terminal (*i+1* site) to examine ATP incorporation at the active site (see SI Section 2.2.2 for details). We found that the incoming ATP becomes less stable when remdesivir is at the *i+1* site. In particular, the Pα-O3' distance becomes larger and shifts to ~3.7 Å on average (Fig. 1B). Moreover, the base stacking stability is substantially weakened, with more MD conformations showing larger base-to-base distance than wildtype-RNA (Fig. 1H). These observations suggest that remdesivir at *i+1* site can reduce the efficiency of the next NTP addition at *i* site.

To examine if remdesivir at further upstream positions of the nascent RNA strand can interfere with NTP incorporation, we performed twenty 30-ns MD simulations for each of the systems with remdesivir at *i+2*, *i+3* and *i+4* sites (see SI Section 2.2.2 for details). The histogram of critical distances of NTP incorporation shows no obvious discrepancy from that of wildtype-RNA (Figs. 1C, 1F, 1I and S2). This suggests that remdesivir at *i+2*, *i+3* or *i+4* site has negligible impact on the nucleotide addition at the active site. Altogether, we expect that remdesivir-TP can be incorporated into the nascent strand, and subsequently would reduce the efficiency of addition of the next incoming NTP. However, we expect that this reduction may not be sufficient to completely abolish the RNA synthesis, suggesting that other mechanisms may exist for remdesivir to inhibit RdRp function.

Remdesivir can abolish nucleotide addition in RdRp by hindering translocation

Remdesivir has been proposed as a delayed chain terminator for the RdRps of respiratory syncytial virus[33], Nipah virus[34], and MERS-CoV[24]. Recent biochemical experiments for SARS-CoV-2 RdRp also demonstrated delayed chain termination by remdesivir-TP at physiological concentrations[15, 25], although nucleotide addition may be terminated earlier when excess amount of remdesivir-TP is present[15]. For MERS-CoV[24] and SARS-CoV-2[15, 25] RNA

replication, remdesivir induced delayed chain termination wherein the addition efficiency of the first few nucleotides is maintained. However, nucleotide addition is abruptly abolished at a specific upstream site with a 3-nucleotide delay. The molecular mechanisms underlining this delayed chain termination remain largely elusive.

To elucidate mechanisms of the delayed termination, we performed MD simulations for RdRp with remdesivir added to the nascent RNA strand and at different sites (*i* to *i+4* site) during the translocation process. To obtain statistically meaningful results, we performed twenty 30-ns MD simulations for each of these systems (see SI Section 2.2.2 for details). To investigate the propensity of translocation, we first inspect how remdesivir may impact stability of the RdRp complex by examining probability of hydrogen bonds formed between remdesivir and its template uracil. As shown in Fig. 4A and Fig. S3, the hydrogen bonds of remdesivir:U are nearly intact when remdesivir is at *i*, *i*+1, and *i*+2 sites in both pre- (pre-T) and post-translocation (post-T) states, suggesting that translocation of remdesivir is allowed in these systems. It is worth mentioning that interactions between remdesivir (at *i* site) and surrounding residues in our model are consistent with those in the recently solved SARS-CoV-2 RdRp holo structure[15] (Fig. S4). Our observations are also consistent with a recent structural modelling study, which suggests that remdesivir at the active site (*i* site) cannot hinder translocation[26].

In sharp contrast, we found that remdesivir at *i+3* site greatly weakens the hydrogen bond probability of base pairing with its template uracil (Fig. 4A). Further investigations show that the presence of the highly polar 1'-cyano group on remdesivir will exert strong electrostatic attractions (~16.1 KJ/mol) with the salt bridge formed by Lys593 and Asp865 of RdRp (Fig. 4B). These electrostatic attractions will further pull remdesivir away from its canonical conformation and cause the destabilization of its base pairing (Fig. 4B). Indeed, the hydrogen bonds for the base pairing of remdesivir:U are subsequently disrupted, rendering the corresponding RdRp complex unstable (Fig. 4A). Recently solved Cryo-EM structure[15] also shows that Lys593 (d=4.6 Å) and Asp865 (d=4.3 Å) are in spatial proximity with 1'-cyano group when remdesivir was modelled at *i+3* site. For translocation of RNA polymerases, a post-T state with the empty active site often has comparable stability with its corresponding pre-T state, and subsequent binding of NTP can greatly stabilize post-T state and favor forward translocation[35]. We thus anticipate that translocation of remdesivir into *i+3* site (post-T) will still be allowed due to the stabilization of NTP binding, but after NTP incorporation, the unstable complex with remdesivir at *i*+3 site may hinder the RNA extension. Next, we investigate if this instability induced by remdesivir at *i+3* site can affect the dynamics of subsequent translocation to *i+4* site by generating a translocation pathway using the Climber algorithm[35-36] (see Methods and SI Section 3 for details). Interestingly, we found that the 1'-cyano group need to first move closer to the negatively charged Asp865 before it translocates away to *i+4* site, while its distance with Lys593 is noticeably larger than that with Asp865 during translocation (Fig. 4C). As the electron-withdrawing 1'-cyano moiety exhibits δ^--anion repulsion with Asp865, the intermediate conformation with reduced distance to Asp865 will experience unfavorable repulsions (~10.5 kJ/mol raised energy, see Fig. 4D), and we expect that this may hinder forward translocation (Fig. 4C). Altogether, the above observations suggest a delayed termination of three nucleotides as observed in SARS-CoV-2[25] and MERS-CoV[24].

Gordon *et al.*[25] proposed an alternative explanation where the steric clash between Ser861 and remdesivir would inhibit translocation from *i+3* site to *i+4* site. In our translocation simulation, we found that the 1'-cyano group of remdesivir constantly approaches Ser861 with their distance monotonically decreased from 7 Å to around 3 Å (Fig. 4C). However, we did not observe strong steric clashes between Ser861 and remdesivir. Nevertheless, due to its spatial

proximity, we anticipate that Ser861 may still play an auxiliary role to the inhibition of translocation. Based on homology modelling, Shannon *et al.*[26] proposed another mechanism where translocation is inhibited due to the steric clash between Arg858 and remdesivir's ribose at *i+4* site. However, our MD simulations with remdesivir at *i+4* site show a large distance (~11.4 Å) between Arg858 and 1'-cyano group of remdesivir (Fig. S5A). Consistent with our simulations, a large separation of 9.5 Å between Arg858 and the 1'-cyano group of remdesivir was also observed when we modelled remdesivir at *i+4* site using the recent cryo-EM structure (PDBID: 7BV2[15], see Fig. S5B). To obtain the translocation pathway, we adopted the Climber method[35-36], in which the RdRp complex is progressively driven from pre-T to post-T state by external energy. To rigorously study dynamics of large-scale conformational changes during translocation, we need to apply more advanced methods such as Markov State Model[35, 37-42], which is beyond the scope of our current study. Finally, comparison of the sequences between SARS-CoV-2 and other coronaviruses indicates that the $Asp865^{-}$-$Lys593^{+}$ salt bridge is conserved except that Lys was replaced by another positively charged Arg in two human CoVs (Fig. S6), implying that our proposed mechanism may be employed by other CoV RdRp.

Remdesivir can inhibit cleavage in ExoN

One main challenge for drug design targeting SARS-CoV-2 RdRp is due to the existence of an ExoN domain in nsp14, which binds to nsp12 and performs proofreading function that can remove incorporated NTP analogues and counteract the efficacy of the inhibitory compounds. Therefore, an additional criterion to be an effective inhibitor SARS-CoV-2 viral replication is the ability to escape from the cleavage by ExoN. Next, we examine how efficient remdesivir is being cleaved by ExoN.

We built the structural model for nsp14 bound with its activator nsp10 based on the crystal structure of nsp14-nsp10 of SARS-CoV (PDBID: 5C8S[43]). As previous work suggests that three nucleotide base pairs are melted in the active site of ExoN[44], we modelled a single-stranded RNA containing three nucleotides based on structural alignment to DNA polymerase I Klenow Fragment[44] and DNA Polymerase III ε subunit[45] (see Method and SI Section 1.2 for details). Twenty 30-ns MD simulations were performed for the wildtype-RNA, in which the 3'-terminal is occupied by adenine (see SI Section 2.3.1 for details). To assess the stability of cleavage site, we examined the distance between MgA and the O3' atom of the second last nucleotide at the 3'-terminal as it is essential for cleavage[46-48] (Fig 5A). This distance is well maintained in most of the simulations (Fig. S7), and the histogram shows only one peak in the region of 2.0~2.5 Å (Fig. 5A) suitable for cleavage[46-48]. MgA-MgB distance is also critical for cleavage[46-49] and our calculations show a sharp peak at ~ 3.6 Å (Fig. 5B), a distance permissible for cleavage[46-49]. By sharp contrast, MD simulations with remdesivir at 3'-terminal show destabilization of cleavage site in ExoN, with the coordination between MgA and O3' atoms more frequently disrupted (Fig. S8). Consistently, a second peak appears at ~3.7Å in the distribution of O3'-MgA distance, indicating that remdesivir causes an increase in the population of cleavage-inhibited configurations (Fig. 5A). This increase was also observed in the histogram of MgA-MgB distance with a second peak clearly appearing at ~4.9 Å (Fig. 5B). Further examination of the base stepping twisting angle (Fig. S9) shows that the base of remdesivir twists more significantly than that of adenine (Fig. 5C), which also suggests that remdesivir is experiencing unfavorable interactions at the 3'-terminal. All these observations indicate that remdesivir is not stable in the cleavage site and hence not prone to be cleaved.

Structural analysis based on MD conformations with remdesivir at the 3'-terminal indicates that the bulky 1'-cyano group plays a key role in protecting remdesivir from ExoN cleavage. The cleavage site is on the surface of nsp14 protein and the base of nucleotides are exposed to the solvent. In contrast, the ribose ring is more buried within the protein environment, thus it

is more susceptible to any instability incurring at the cleavage site. The wildtype nucleotide does not have a bulky group attached at the ribose ring. Hence, no specific steric effect is observed in the cleavage site as shown by a representative conformation of wildtype-RNA (Fig. 5E). Conversely, the 1'-cyano group of remdesivir causes steric clash with the side chain of protein residue Asn104. This steric repulsion separates Asn104 and remdesivir apart from each other, as consolidated by the larger distance between the base of Asn104 and nitrogen atom in the amide group of Asn104 (Fig. 5D). This steric interaction is clearly demonstrated in representative conformations as shown in Fig. 5F. On one hand, if remdesivir adopts a similar conformation as wildtype-RNA at 3'-terminal, the cyano group will push down the side chain of Asn104. On the other hand, if the side chain of Asn104 maintains its configuration as in the wildtype-RNA, the steric interaction will push the ribose along with the base of remdesivir rotating away (Fig. 5F). Therefore, we attribute the destabilization of remdesivir at cleavage site to the steric repulsion between 1'-cyano group of remdesivir and Asn104. This observation from MD simulations is consistent with a recent study based on static structural model of nsp14 of SARS-CoV-2, which also proposes that the 1'-cyano group has steric clash with the cleavage site[26]. This destabilization would reduce the possibility of remdesivir to be excised by ExoN, enabling remdesivir to be an effective inhibitor for SARS-CoV-2.

Elucidations of important chemical features of NTP analogues for the inhibition of SARS-CoV-2 RNA replication

To further understand important chemical features of remdesivir allowing it to inhibit nucleotide addition and cleavage in the viral replication of SARS-CoV-2, we explored the inhibitory effects of three additional NTP analogues containing various chemical structure features: favipiravir, vidarabine and fludarabine. All these three compounds are FDA approved drugs and display the same hydrogen bonding pattern as the corresponding natural adenine nucleotide (Fig. 6A). However, they are structurally different from adenine in either the base (favipiravir) or the ribose (vidarabine). Like remdesivir, fludarabine has modifications in both the base and ribose. Therefore, examination of these NTP analogues can provide insights on the types of modifications that may benefit inhibitor design.

We first performed MD simulations by placing the triphosphate form of the NTP analogue in the active site (*i* site) (see SI Section 2.2.2 for details) to evaluate if these NTP analogues could satisfy the criteria of being incorporated into the nascent strand in RdRp as potential chain terminators. Favipiravir-TP and vidarabine-TP were shown to possess similar mean Pα-O3' distance (~4.0 Å) compared with remdesivir (Fig. 6B). This distance is significantly raised to ~5.0 Å for fludarabine-TP, implicating the difficulty of fludarabine to be incorporated into the nascent RNA strand. Further inspection of the MD conformations revealed that the instability of fludarabine-TP in the active site could be attributed to the fluoro-substitution on the base. Fluorine atom is electronegative, and causes a δ^--δ^- repulsion with the oxygen atom in the base of the template uracil (Fig. S10), thus destabilizing its configuration at the active site. The base-to-base distance between the NTP analogue and the template nucleotide in the active site only marginally increased by < 0.25 Å for all the NTP analogues compared to wildtype-RNA (Fig. S11A). Similarly, the mean base stacking distances for the NTP analogues are equal to or even shorter than that for ATP (Fig. S11C). These observations suggest that all NTP analogues except fludarabine-TP satisfy the structural feature criteria to be incorporated into the nascent strand.

We next examined if these NTP analogues can inhibit nucleotide addition by placing them at 3'-terminal (*i+1* site) with ATP in the active site (Twenty 30-ns MD simulations were performed for each of these systems, see SI Section 2.2.2 for details). For favipiravir, we found that the averaged O3'-Pα distance is ~3.6 Å when favipiravir is at *i+1* site, similar to that when

remdesivir is at *i+1* site (Fig. 6C), demonstrating favipiravir's weak capability to reduce the efficiency for the next NTP incorporation. Vidarabine and fludarabine were shown to have even shorter distance (~3.3 Å) than wildtype-RNA, due to an extra hydrogen bond formed between their 2'-hydroxyl group of the NTP analogue and 4'-oxygen of NTP (hydrogen bond probability of 89% and 94% for vidarabine and fludarabine, respectively). Base-to-base distance between NTP and the template nucleotides are not perturbed by the presence of the NTP analogue at *i+1* site (Fig. S11B). Favipiravir also shows similar base stacking distance as wildtype-RNA, while this distance is slightly increased by around 0.3~0.4 Å for vidarabine and fludarabine (Fig. S11D). In addition, as none of these three NTP analogues contains a bulky substitution group on the ribose like 1'-cyano on remdesivir (Fig. 6A), we anticipate that it will be unlikely for them to inhibit RNA synthesis via a delayed termination mechanism. Altogether, we expect that favipiravir can reduce NTP incorporation to a similar extent as remdesivir, while vidarabine and fludarabine at *i+1* do not reduce NTP incorporation due to the shortening of the O3'-Pα distance. Therefore, based on the comprehensive analysis of three NTP analogues' performance in RdRp, favipiravir is the most possible candidate.

Next, we examined if these NTP analogues could inhibit proofreading of ExoN by placing them at the 3'-terminal of the nascent RNA strand (see SI Section 2.3.2 for details). For all three NTP analogues, the O3'-MgA and MgA-MgB distances were comparable to that when adenine is at the 3'-terminal, implicating that they are more prone to be excised by ExoN (Figs. 6D and S12A). Furthermore, these three NTP analogues adopt similar configuration as adenine in the cleavage site (Fig. S13) and do not experience any steric clashes with Asn104 (Fig. 6E). We anticipate that the stability of the three compounds at the cleavage site can be attributed to the lack of bulky substitutions at 1'-ribose. This is consistent with our proposal that the bulky 1'-cyano group causes steric effect, contributing to the instability of the cleavage site. Altogether, our results suggest that these three NTP analogues are more prone to be excised than remdesivir, implicating that they are not effective to inhibit RNA replication of SARS-CoV-2.

To consolidate the above-mentioned simulation results, we performed *in vitro* experiments in Vero E6 cells and show that NTP analogues favipiravir, vidarabine and fludarabine indeed fail to inhibit SARS-CoV-2 replication (Fig. 6F). In particular, we performed *in vitro* tissue culture assays using live SARS-CoV-2 virus to examine the antiviral effect of remdesivir, favipiravir, vidarabine, and fludarabine against SARS-CoV-2 virus. Among the four NTP analogues tested, only remdesivir was found to inhibit SARS-CoV-2 replication in Vero E6 cells with EC50 under 100 μM (Fig. 6F). This is consistent with our previous work[50], which determined the EC50 of vidarabine to be 23.15 μM and 26.90 μM using virus titer and real-time PCR, respectively. Other NTP analogues (vidarabine, fludarabine and favipiravir) did not inhibit viral replication under 100 μM (Fig. 6F). However, it is worthy to note that NTP analogues require conversions into their triphosphate forms by host nucleoside kinases, whose activity may differ among cell types. Hence, further evaluation of the effect of these NTP analogues in primary cells such as primary human airway epithelial cells that more clinically relevant is needed.

Considering the inhibition effect of all three NTP analogues in RdRp and ExoN, we found that none of them exhibits superior inhibitory activity than remdesivir. Our work elucidates the molecular mechanism for their lack of inhibitory effects on RNA synthesis. Fludarabine has the lowest tendency to be incorporated into the nascent strand, while vidarabine could not reduce the efficiency of incoming NTP incorporation in SARS-CoV-2 RdRp. Favipiravir demonstrates similar performance as remdesivir in RdRp; however, favipiravir is highly likely

to be cleaved in ExoN. Hence, only remdesivir can inhibit RNA synthesis in RdRp and simultaneously avoid its excision by the ExoN of SARS-CoV-2 RdRp.

Conclusions

In this study, we have elucidated the important role of 1'-cyano group on the ribose in remdesivir's inhibition effect targeted RNA replication (Fig. 7). Extensive MD simulations were performed based on the structural model of RdRp and ExoN of SARS-CoV-2. Remdesivir at *i+1* site can reduce nucleotide addition at the active site but cannot fully inhibit nucleotide addition. Instead, remdesivir mainly inhibits nucleotide addition via a delayed chain terminating mechanism, where the growing RNA chain will be terminated once a few nucleotides are added after remdesivir incorporation. We propose a model where the polar 1'-cyano group of remdesivir at the upstream *i+3* site will cause instability of the RdRp complex, because the salt bridge ($Asp865^{-}$-$Lys593^{+}$) will pull remdesivir away from its canonical conformation to break its base pairing. Furthermore, the dynamic transition of translocation from *i+3* site to *i+4* site will be further hindered due to unfavorable anion-δ^{-} repulsion between Asp865 and the 1'-cyano group of remdesivir. As a result, remdesivir leads to a delayed termination of three nucleotides. In addition, remdesivir was found to be capable of inhibiting the proofreading function of ExoN. The 1'-cyano group of remdesivir exhibits steric clash with a nearby residue Asn104, which destabilizes remdesivir at the active site, rendering it resistant to ExoN cleavage. Further investigation using three other NTP analogues showed different inhibiting effects on nucleotide addition in RdRp, albeit all of them are more prone to be cleaved than remdesivir. Comparison among the structural modifications of remdesivir with other three NTP analogues suggests that the modification on the ribose, possibility by using a bulky group substitution may be a better strategy than on the base to inhibit the RNA synthesis of SARS-CoV-2. Our work provides mechanistic insights at atomistic details on how remdesivir can effectively inhibit SARS-CoV-2 RNA synthesis. We also revealed that the chemical modification at the ribose of remdesivir that carries the dual inhibitory function to terminate both nucleotide addition and proofreading could be a promising general strategy for designing prospective nucleoside-analog inhibitors to treat COVID-19.

Methodology

Structural model of nsp12-nsp7-nsp8 complex The cryo-EM structure of nsp12-nsp7-nsp8 complex (PDBID: 6NUR)[12] and the norovirus holo-RdRp (PDBID: 3H5Y[28]) were used as the structural basis to construct the nCoV RdRp containing dsRNA strands with ATP bound at the active site (see SI Section 1.1 for details). The protonation states of histidines were predicted using propka3.0 module[51] in the pdbpqr2.2.1[52] package, followed by manual investigation to ensure that the coordination with the zinc ion can be formed. The whole complex was placed in a dodecahedron box with the box edges at least 12Å away from the complex surface. The box was filled with TIP3P water molecules[53], and enough counter ions were added to make the whole system neutral.

Structural model of nsp14-nsp10 complex The crystal structure of nsp10-nsp14 complex of SARS-CoV (PDBID: 5C8S[43]) serves as the structural basis to construct the nsp10-nsp14 of SARS-CoV-2 (see SI Section 1.2 for details). Missing residues in nsp14 were modelled by modeller9.21[54]. Single-strand RNA containing three nucleotides were modelled by structural alignment to the cleavage site of the proofreading ExoN domain of DNA polymerase I Klenow fragment (PDBID: 1KLN[44]) and the ε-subunit of DNA polymerase III (PDBID: 1J53[45]) due to the similar architecture of the cleavage site. The protonation states of histidines were estimated by propka3.0 module[51] of the pdbpqr2.2.1[52] package. The complex was solvated with TIP3P

water in a dodecahedron box, the edge of which was at least 12 Å away from the complex surface. Sufficient counter ions were added to neutralize the system.

MD setup and parameters We used the amber99sb-ildn force field[55] to simulate protein and nucleotides. The force field parameters for NTP analogues were derived based on the existing amber99sb-ildn force field parameters or general amber force field[56-57] (see SI Section 2.1 for details). Partial charges of NTP analogues were generated following the Restricted Electrostatic Potential scheme[58-59] (see SI for details). For ATP or NTP analogues in the triphosphate form, parameters for the triphosphate tail were taken from those developed by Meagher *et al.*[60]. The system was gradually relaxed before the production simulation under NVT ensemble (T=300 K). V-rescale thermostat[61] was applied with the coupling time constant of 0.1 ps. The long-range electrostatic interactions beyond the cut-off at 12 Å were treated with the Particle-Mesh Ewald method.[62] Lennard-Jones interactions were smoothly switched off from 10 Å to 12 Å. The neighbor list was updated every 10 steps. An integration time step of 2.0 fs was used and the LINCS algorithm[63] was applied to constrain all the bonds. We saved the snapshots every 20 ps. All simulations were performed with Gromacs 5.0[64]. The equilibrated system containing wildtype RNA serves as the structural basis for modelling NTP analogue at the corresponding site to investigate their inhibitory effects. The details of MD simulations are in SI Section 2.2 and 2.3.

***Translocation pathway of remdesivir from i+3* to *i+4* site** We used Climber algorithm[35-36] to generate a pathway for the translocation of remdesivir from *i+3* to *i+4* site (see SI Section 3 for details). The pre-T and post-T conformations were constructed based on the pre-T model with remdesivir at *i+3* site and post-T model with remdesivir at *i+4* site, respectively. We performed a 500-step Climber simulation, where double stranded RNA was gradually driven from the pre-T state to the post-T state by external energy.

***In vitro* assays with live SARS-CoV-2 virus for the determination of EC50 of NTP analogues** SARS-CoV-2 virus, named BetaCoV/Hong Kong/VM20001061/2020, was isolated from the nasopharynx aspirate and throat swab of a confirmed COVID-19 patient in Hong Kong. Stock virus (107.25 TCID50/mL) was prepared after 3 passages in using Vero E6 cells (ATCC: CRL-1586) in infection media (DMEM supplemented with 4.5 g/L D-glucose, 100 mg/L sodium pyruvate, 2% FBS, 100,000 U/L Penicillin-Streptomycin, and 25 mM HEPES). Compounds (MedChemExpress) and the stocks were prepared with DMSO. To evaluate the effect of compounds *in vitro*, Vero E6 cells were pre-treated with compounds diluted in infection media for 1 h prior to inoculation by SARS-CoV-2 virus at Multiplicity of Infection at 0.02. Antiviral compounds were present with the virus inoculum during the 2-h incubation period, after which the inoculum was removed. Then, Vero E6 cells were overlaid with infection media containing the respective diluted compounds. After 48 h incubation at 37 °C, supernatants were collected to quantify viral loads by TCID50 assay or quantitative real-time RT-PCR (TaqMan™ Fast Virus 1-Step Master Mix, Thermo Scientific) based on a previous protocol[65]. Logistic regression was used to fit the dose-response curves to determine the 50% effective concentrations (EC50) of the compounds that inhibit live viral replication. Cytotoxicty of selected compounds was evaluated in Vero E6 cells using the CellTiter-Glo Luminescent Cell Viability Assay (Promega).

Acknowledgement
L. Z. was supported by the National Natural Science Foundation of China Grants (21733007 and 21803071). X.H. was supported by Shenzhen Science and Technology Innovation Committee (JCYJ20170413173837121), Hong Kong Research Grant Council (16303919, 16307718, 16318816, AoE/P-705/16, AoE/M-09/12, and T13-605/18-W), and Innovation and Technology Commission (ITCPD/17-9 and ITC-CNERC14SC01). X. G. was supported by the

King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No. FCC/1/1976-18, FCC/1/1976-23, FCC/1/1976-26, URF/1/4098-01-01, and REI/1/0018-01-01. This research made use of the computing resources of the Supercomputing Laboratory at King Abdullah University of Science and Technology.

References

(1) WHO, Coronavirus disease 2019 (COVID-19): Situation Report – 23, Covid-19 Situatioanal Reports.

(2) Samuel, B., Medical microbiology. *The University of Texas Medical Branch at Galveston–Tx, USA* **1996**.

(3) Li, F., Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* **2016,** *3* (1), 237-261.

(4) Yuen, K.-S.; Ye, Z. W.; Fung, S.-Y.; Chan, C.-P.; Jin, D.-Y., SARS-CoV-2 and COVID-19: The most important research questions. *Cell Biosci.* **2020,** *10* (1), 40.

(5) Vincent, M. J.; Bergeron, E.; Benjannet, S.; Erickson, B. R.; Rollin, P. E.; Ksiazek, T. G.; Seidah, N. G.; Nichol, S. T., Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virol. J.* **2005,** *2* (1), 69.

(6) Pascal, K. E.; Coleman, C. M.; Mujica, A. O.; Kamat, V.; Badithe, A.; Fairhurst, J.; Hunt, C.; Strein, J.; Berrebi, A.; Sisk, J. M.; Matthews, K. L.; Babb, R.; Chen, G.; Lai, K.-M. V.; Huang, T. T.; Olson, W.; Yancopoulos, G. D.; Stahl, N.; Frieman, M. B.; Kyratsous, C. A., Pre- and postexposure efficacy of fully human antibodies against Spike protein in a novel humanized mouse model of MERS-CoV infection. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112* (28), 8738.

(7) Keefe, B. R.; Giomarelli, B.; Barnard, D. L.; Shenoy, S. R.; Chan, P. K. S.; McMahon, J. B.; Palmer, K. E.; Barnett, B. W.; Meyerholz, D. K.; Wohlford-Lenane, C. L.; McCray, P. B., Broad-Spectrum In Vitro Activity and In Vivo Efficacy of the Antiviral Protein Griffithsin against Emerging Viruses of the Family Coronaviridae. *J. Virol.* **2010,** *84* (5), 2511.

(8) Kennedy, D. A.; Read, A. F., Why the evolution of vaccine resistance is less of a concern than the evolution of drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* **2018,** *115* (51), 12878-12886.

(9) Lythgoe, M. P.; Middleton, P., Ongoing Clinical Trials for the Management of the COVID-19 Pandemic. *Trends Pharmacol. Sci.* **2020**.

(10) Imbert, I.; Guillemot, J.-C.; Bourhis, J.-M.; Bussetta, C.; Coutard, B.; Egloff, M.-P.; Ferron, F.; Gorbalenya, A. E.; Canard, B., A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* **2006,** *25* (20), 4933-4942.

(11) Subissi, L.; Posthuma, C. C.; Collet, A.; Zevenhoven-Dobbe, J. C.; Gorbalenya, A. E.; Decroly, E.; Snijder, E. J.; Canard, B.; Imbert, I., One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111* (37), E3900.

(12) Kirchdoerfer, R. N.; Ward, A. B., Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* **2019,** *10* (1), 2342.

(13) Zhai, Y.; Sun, F.; Li, X.; Pang, H.; Xu, X.; Bartlam, M.; Rao, Z., Insights into SARS-CoV transcription and replication from the structure of the nsp7–nsp8 hexadecamer. *Nat. Struct. Mol. Biol.* **2005,** *12* (11), 980-986.

(14) 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**, eabb7498.

(15) 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 the Inhibition of the RNA-Dependent RNA Polymerase from SARS-CoV-2 by Remdesivir. *bioRxiv* **2020**, 2020.2004.2008.032763.

(16) Ferron, F.; Subissi, L.; Silveira De Morais, A. T.; Le, N. T. T.; Sevajol, M.; Gluais, L.; Decroly, E.; Vonrhein, C.; Bricogne, G.; Canard, B.; Imbert, I., Structural and molecular basis

of mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl. Acad. Sci. U. S. A.* **2018,** *115* (2), E162.

(17) Eckerle, L. D.; Becker, M. M.; Halpin, R. A.; Li, K.; Venter, E.; Lu, X.; Scherbakova, S.; Graham, R. L.; Baric, R. S.; Stockwell, T. B.; Spiro, D. J.; Denison, M. R., Infidelity of SARS-CoV Nsp14-Exonuclease Mutant Virus Replication Is Revealed by Complete Genome Sequencing. *PLoS Pathog.* **2010,** *6* (5), e1000896.

(18) Smith, E. C.; Blanc, H.; Vignuzzi, M.; Denison, M. R., Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. *PLoS Pathog.* **2013,** *9* (8), e1003565.

(19) 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* (2), e00221-00218.

(20) Holshue, M. L.; DeBolt, C.; Lindquist, S.; Lofy, K. H.; Wiesman, J.; Bruce, H.; Spitters, C.; Ericson, K.; Wilkerson, S.; Tural, A.; Diaz, G.; Cohn, A.; Fox, L.; Patel, A.; Gerber, S. I.; Kim, L.; Tong, S.; Lu, X.; Lindstrom, S.; Pallansch, M. A.; Weldon, W. C.; Biggs, H. M.; Uyeki, T. M.; Pillai, S. K., First Case of 2019 Novel Coronavirus in the United States. *N. Engl. J. Med.* **2020,** *382* (10), 929-936.

(21) Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G., Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **2020,** *30* (3), 269-271.

(22) 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**.

(23) Sheahan, T. P.; Sims, A. C.; Graham, R. L.; Menachery, V. D.; Gralinski, L. E.; Case, J. B.; Leist, S. R.; Pyrc, K.; Feng, J. Y.; Trantcheva, I.; Bannister, R.; Park, Y.; Babusis, D.; Clarke, M. O.; Mackman, R. L.; Spahn, J. E.; Palmiotti, C. A.; Siegel, D.; Ray, A. S.; Cihlar, T.; Jordan, R.; Denison, M. R.; Baric, R. S., Broad-spectrum antiviral GS-5734 inhibits both epidemic and zoonotic coronaviruses. *Sci. Transl. Med.* **2017,** *9* (396), eaal3653.

(24) Gordon, C. J.; Tchesnokov, E. P.; Feng, J. Y.; Porter, D. P.; Götte, M., The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J. Biol. Chem.* **2020,** *295* (15), 4773-4779.

(25) 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. *Journal of Biological Chemistry* **2020**.

(26) Shannon, A.; Le, N. T.-T.; Selisko, B.; Eydoux, C.; Alvarez, K.; Guillemot, J.-C.; Decroly, E.; Peersen, O.; Ferron, F.; Canard, B., Remdesivir and SARS-CoV-2: Structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites. *Antiviral Res.* **2020,** *178*, 104793.

(27) Zhang, L. L.; Zhou, R. H., Binding Mechanism of Remdesivir to SARS-CoV-2 RNA Dependent RNA Polymerase. *Preprints* **2020,** *2020030267*, (doi: 10.20944/preprints202003.200267.v202001).

(28) Zamyatkin, D. F.; Parra, F.; Machín, Á.; Grochulski, P.; Ng, K. K. S., Binding of 2′-Amino-2′-Deoxycytidine-5′-Triphosphate to Norovirus Polymerase Induces Rearrangement of the Active Site. *J. Mol. Biol.* **2009,** *390* (1), 10-16.

(29) Sgrignani, J.; Magistrato, A., The Structural Role of Mg2+ Ions in a Class I RNA Polymerase Ribozyme: A Molecular Simulation Study. *J. Phys. Chem. B* **2012,** *116* (7), 2259-2268.

(30) Radhakrishnan, R., Coupling of fast and slow modes in the reaction pathway of the minimal hammerhead ribozyme cleavage. *Biophys. J.* **2007,** *93* (7), 2391-2399.

(31) Huang, X.; Wang, D.; Weiss, D. R.; Bushnell, D. A.; Kornberg, R. D.; Levitt, M., RNA polymerase II trigger loop residues stabilize and position the incoming nucleotide triphosphate in transcription. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (36), 15745.

(32) Gong, P.; Peersen, O. B., Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (52), 22505.

(33) Tchesnokov, E. P.; Feng, J. Y.; Porter, D. P.; Götte, M., Mechanism of Inhibition of Ebola Virus RNA-Dependent RNA Polymerase by Remdesivir. *Viruses* **2019,** *11* (4), 326.

(34) Jordan, P. C.; Liu, C.; Raynaud, P.; Lo, M. K.; Spiropoulou, C. F.; Symons, J. A.; Beigelman, L.; Deval, J., Initiation, extension, and termination of RNA synthesis by a paramyxovirus polymerase. *PLoS Pathog.* **2018,** *14* (2), e1006889.

(35) Silva, D. A.; Weiss, D. R.; Avila, F. P.; Da, L. T.; Levitt, M.; Wang, D.; Huang, X. H., Millisecond dynamics of RNA polymerase II translocation at atomic resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111* (21), 7665-7670.

(36) Weiss, D. R.; Levitt, M., Can morphing methods predict intermediate structures? *J Mol Biol* **2009,** *385* (2), 665-674.

(37) Klippenstein, S. J.; Pande, V. S.; Truhlar, D. G., Chemical Kinetics and Mechanisms of Complex Systems: A Perspective on Recent Theoretical Advances. *J. Am. Chem. Soc.* **2014,** *136* (2), 528-546.

(38) Chodera, J. D.; Noe, F., Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **2014,** *25*, 135-144.

(39) Bowman, G. R.; Huang, X. H.; Pande, V. S., Using generalized ensemble simulations and Markov state models to identify conformational states. *Methods* **2009,** *49* (2), 197-201.

(40) Zhang, L.; Pardo-Avila, F.; Unarta, I. C.; Cheung, P. P. H.; Wang, G.; Wang, D.; Huang, X. H., Elucidation of the Dynamics of Transcription Elongation by RNA Polymerase II using Kinetic Network Models. *Acc. Chem. Res.* **2016,** *49* (4), 687-694.

(41) Wang, W.; Cao, S. Q.; Zhu, L. Z.; Huang, X. H., Constructing Markov State Models to elucidate the functional conformational changes of complex biomolecules. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018,** *8* (1).

(42) Da, L.-T.; Pardo-Avila, F.; Xu, L.; Silva, D.-A.; Zhang, L.; Gao, X.; Wang, D.; Huang, X., Bridge helix bending promotes RNA polymerase II backtracking through a critical and conserved threonine residue. *Nat. Commun.* **2016,** *7* (1), 11244.

(43) Ma, Y.; Wu, L.; Shaw, N.; Gao, Y.; Wang, J.; Sun, Y.; Lou, Z.; Yan, L.; Zhang, R.; Rao, Z., Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112* (30), 9436.

(44) Beese, L. S.; Derbyshire, V.; Steitz, T. A., Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* **1993,** *260* (5106), 352.

(45) Hamdan, S.; Carr, P. D.; Brown, S. E.; Ollis, D. L.; Dixon, N. E., Structural Basis for Proofreading during Replication of the Escherichia coli Chromosome. *Structure* **2002,** *10* (4), 535-546.

(46) Lan, P.; Tan, M.; Zhang, Y.; Niu, S.; Chen, J.; Shi, S.; Qiu, S.; Wang, X.; Peng, X.; Cai, G.; Cheng, H.; Wu, J.; Li, G.; Lei, M., Structural insight into precursor tRNA processing by yeast ribonuclease P. *Science* **2018,** *362* (6415), eaat6678.

(47) Lei, J.; Sheng, G.; Cheung, P. P.-H.; Wang, S.; Li, Y.; Gao, X.; Zhang, Y.; Wang, Y.; Huang, X., Two symmetric arginine residues play distinct roles in Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl. Acad. Sci. U. S. A.* **2019,** *116* (3), 845.

(48) Tse, C. K. M.; Xu, J.; Xu, L.; Sheong, F. K.; Wang, S.; Chow, H. Y.; Gao, X.; Li, X.; Cheung, P. P.-H.; Wang, D.; Zhang, Y.; Huang, X., Intrinsic cleavage of RNA polymerase II adopts a nucleobase-independent mechanism assisted by transcript phosphate. *Nat. Catal.* **2019,** *2* (3), 228-235.

(49) Yang, W.; Lee, J. Y.; Nowotny, M., Making and breaking nucleic acids: Two-Mg2+-ion catalysis and substrate specificity. *Mol. cell* **2006,** *22* (1), 5-13.

(50) Choy, K.-T.; Wong, A. Y.-L.; Kaewpreedee, P.; Sia, S. F.; Chen, D.; Hui, K. P. Y.; Chu, D. K. W.; Chan, M. C. W.; Cheung, P. P.-H.; Huang, X.; Peiris, M.; Yen, H.-L., Remdesivir, lopinavir, emetine, and homoharringtonine inhibit SARS-CoV-2 replication in vitro. *Antiviral Res.* **2020,** *178*, 104786.

(51) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011,** *7* (2), 525-537.

(52) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A., PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Research* **2004,** *32* (suppl_2), W665-W667.

(53) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **1983,** *79* (2), 926-935.

(54) Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinf.* **2016,** *54* (1), 5.6.1-5.6.37.

(55) 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: Struct., Funct., Bioinf.* **2010,** *78* (8), 1950-1958.

(56) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004,** *25* (9), 1157-1174.

(57) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006,** *25* (2), 247-260.

(58) Woods, R. J.; Chappelle, R., Restrained electrostatic potential atomic partial charges for condensed-phase simulations of carbohydrates. *Theochem* **2000,** *527* (1-3), 149-156.

(59) Wang, J.; Cieplak, P.; Kollman, P. A., How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000,** *21* (12), 1049-1074.

(60) Meagher, K. L.; Redman, L. T.; Carlson, H. A., Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.* **2003,** *24* (9), 1016-1025.

(61) Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007,** *126* (1), 014101.

(62) 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* (19), 8577-8593.

(63) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18* (12), 1463-1472.

(64) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015,** *1*, 19-25.

(65) Chu, D. K. W.; Pan, Y.; Cheng, S. M. S.; Hui, K. P. Y.; Krishnan, P.; Liu, Y.; Ng, D. Y. M.; Wan, C. K. C.; Yang, P.; Wang, Q.; Peiris, M.; Poon, L. L. M., Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. *Clin. Chem.* **2020,** *66* (4), 549-555.

Figures:

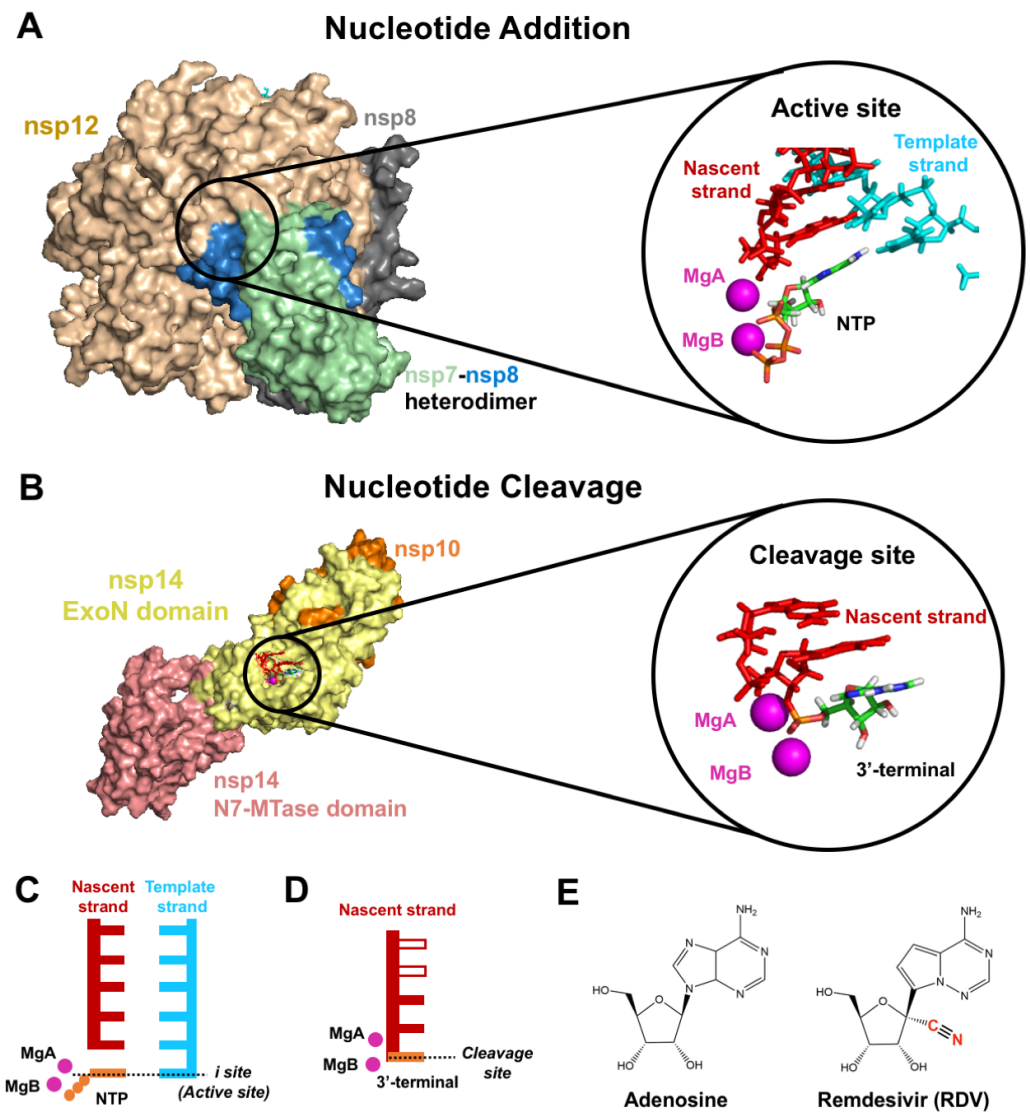

Fig. 1 Structural model of SARS-CoV-2 RdRp and ExoN for investigating the inhibitory mechanism of remdesivir (RDV). (A) Model of nsp12-nsp7-nsp8 complex for nucleotide addition. The active site is circled and amplified on the right panel. The nascent and template strands are colored in red and cyan, respectively. NTP (in orange) is bound at the active site and two Mg^{2+} ions are shown in magenta spheres. (B) Model of nsp14-nsp10 complex, including ExoN domain for nucleotide cleavage. The cleavage site is circled and amplified on the right. Three nucleotides are modelled, including the 3'-terminal site used for modelling RDV or other NTP analogues. Magenta spheres represent two Mg^{2+} ions for cleavage. (C) Cartoon model of the active site in RdRp. (D) Cartoon model of the cleavage site in ExoN. The three terminal nucleotides used in our model are represented by color-filled rectangles. The ones not included in our model are represented by empty rectangles. (E) Chemical structures of adenine and RDV. The 1'-cyano group of RDV is highlighted in red.

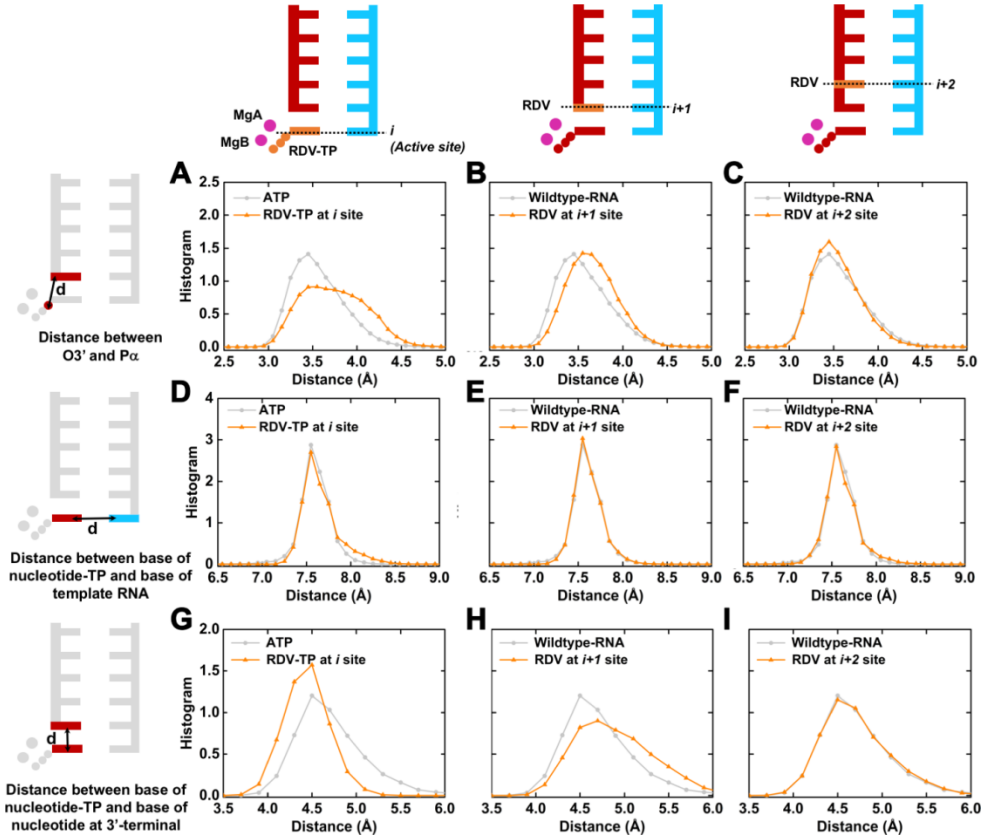

Fig. 2 Stability of active site for nucleotide addition in RdRp when RDV is at *i*, *i+1*, or *i+2* site. The three cartoons in the top panel denote the site where RDV is located (in orange). The three cartoons in the left panel describe the details about the distance for calculation. For clarification, only the molecules involved in the calculations are colored. Each of the nine histograms is calculated for the distance (cartoon in the left vertical panel) using the model with RDV at a specific site (cartoon in top horizontal panel). (A)-(C) Histogram of distance between Pα and O3' of the 3'-terminal nucleotide when RDV is at *i* (A), *i+1* (B), or *i+2* (C) site. (D)-(F) Histogram of distance between the base of ATP/RDV-TP and the base of the corresponding template nucleotide when RDV is at *i* (D), *i+1* (E) or *i+2* (F) site. (G)-(I) Histogram of distance between the base of ATP/RDV-TP and the base of the 3'-terminal nucleotide when RDV is at *i* (G), *i+1* (H) or *i+2* (I) site. For each panel, the histogram for the respective distance of wildtype-RNA with ATP in the active site is shown in light grey as a reference. See SI Section 4.1 for details about the base-to-base distance calculations.

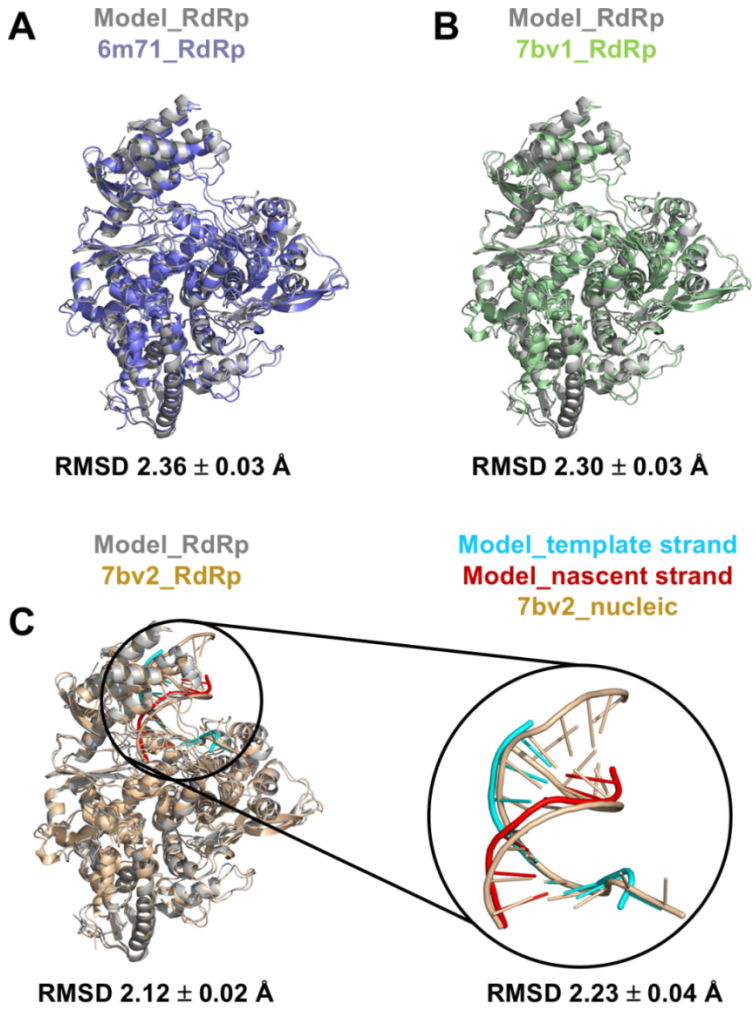

Fig. 3 Comparison of our RdRp structural model with the cryo-EM structures. (A) Comparison between our model (grey) and apo cryo-EM structure (PDBID: 6M71, in blue). (B) Comparison between our model (grey) and apo cryo-EM structure (PDBID: 7BV1, in green). (C) Comparison between our model (grey) and holo cryo-EM structure (PDBID: 7BV2, in orange). In the right panel, the template strand (in cyan) and nascent strand (in red) of our model are compared to those in the cryo-EM structure. See SI Section 4.2 for calculation details.

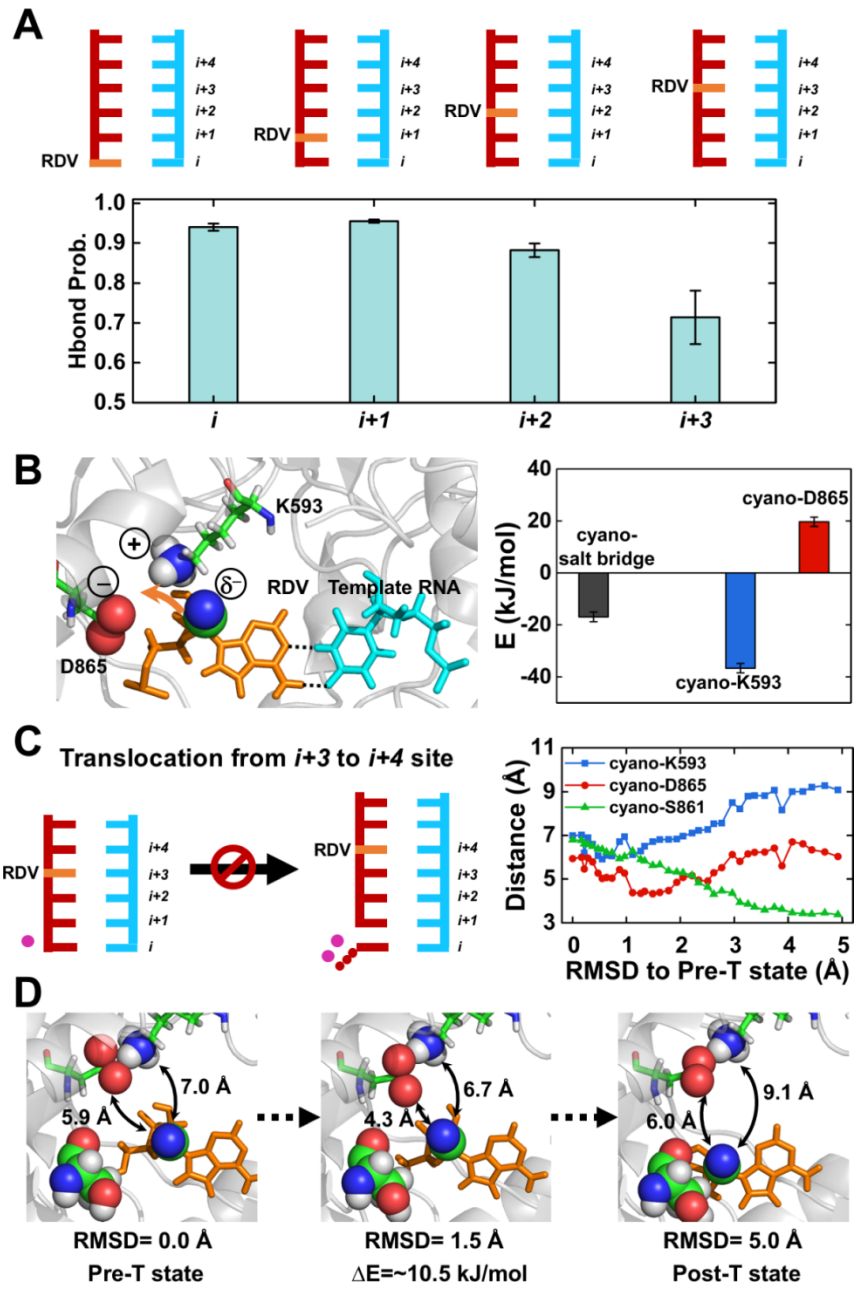

Fig. 4 Remdesivir inhibits translocation in RdRp. (A) The top panels contain cartoon models of pre-T state with RDV (shown in orange) at *i*, *i+1*, *i+2* or *i+3* site. The bottom panel corresponds to the averaged hydrogen bond probability for two hydrogen bonds of the RDV:U base pair when RDV is at *i*, *i+1*, *i+2* or *i+3* site (see SI Section 4.3 for calculation details). (B) The left panel: A representative MD conformation displaying that the D865$^-$-K593$^+$ salt bridge attracts the RDV (shown in orange) from its canonical conformation (two hydrogen bonds for RDV:U pair were shown in dashed lines). The 1'-cyano group of RDV, the two oxygen atoms in the carboxyl group of D865 and the quaternary amine group of K593 are shown in spheres. The right panel: Interaction energy between the 1'-cyano group of RDV and D865$^-$-K593$^+$ salt bridge, 1'-cyano group and D865, as well as 1'-cyano group and K593 are shown in black, red and blue, respectively (see SI Section 4.4 for calculation details). (C) Cartoon model showing the translocation with RDV from *i+3* to *i+4* site is inhibited (Left panel). The distance between 1'-cyano group and D865/K593/S861 versus the RMSD of RNA to the pre-T state during translocation is shown in the Right panel (see SI Section 4.5 and 4.6 for details). (D) Representative conformations along the translocation pathway of RDV from *i+3* to *i-4* site.

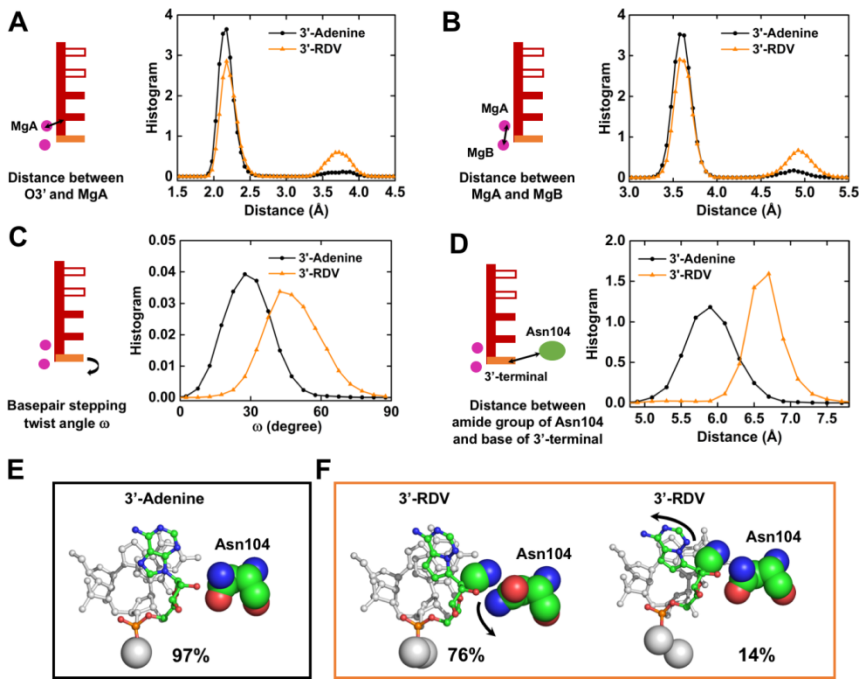

Fig. 5 Remdesivir inhibits cleavage in ExoN. (A) Histogram of distance between MgA and O3' atom of the second nucleotide at 3'-terminal. (B) Histogram of distance between two Mg^{2+} ions. (C) Histogram of the basepair stepping twist angle between the last two nucleotides at the 3'-terminal. (D) Histogram of the distance between the nitrogen atom in the amide group of Asn104 and the base of 3'-terminal (see SI Section 4.7 for details). In (A)-(D), the histogram for the scenario when RDV is at 3'-terminal and that with 3'-adenine is shown in orange and black, respectively. (E) Representative MD conformation containing wildtype RNA with 3'-adenine (F) Representative MD conformations with RDV at 3'-terminal. The 1'-cyano group of RDV and Asn104 are shown in sphere. In (E) and (F), K-center clustering was performed to divide the conformational ensemble into 10 clusters. Only the center conformations of the clusters with population > 10% are shown (see SI Section 4.8 for details).

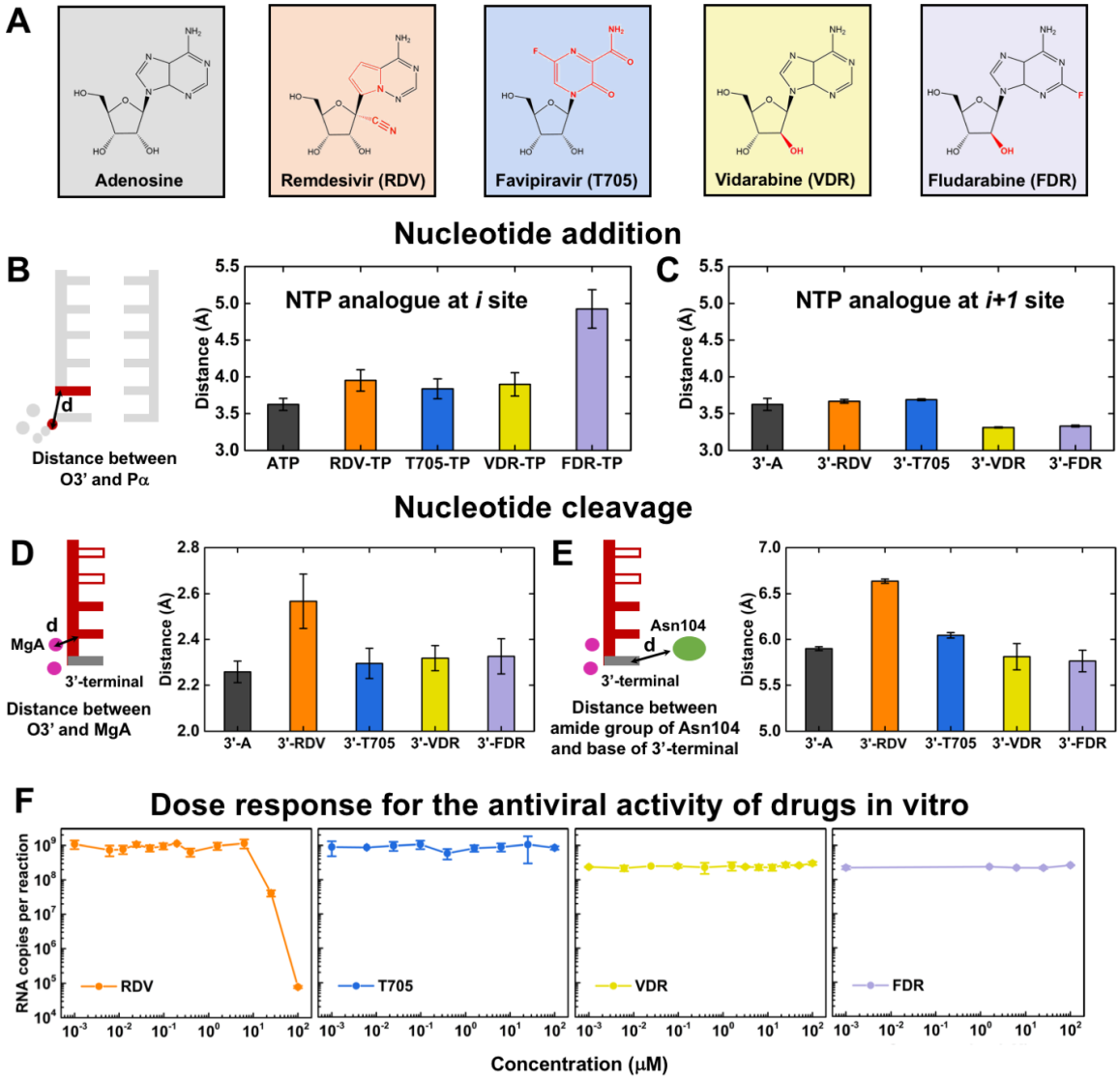

Fig. 6 Investigation of the inhibitory effect of favipiravir (T705), vidarabine (VDR) and fludarabine (FDR) in RdRp and ExoN. (A) Chemical structures of all NTP analogues in comparison with that of adenine. The structural components different from adenine are highlighted in red. (B)-(C) Distance between MgA and O3' atom of the 3'-terminal nucleotide in the RdRp model with NTP analogue at *i* site (B) or *i+1* site (C). (D) Distances between MgA and O3' atom of the second nucleotide at 3'-terminal in the ExoN model with NTP analogues at 3'-terminal. (E) Similar to (D), except they are distances between the base of 3'-terminal nucleotide and the nitrogen atom of the amide group of Asn104 (see SI Section 4.7 for details). For the distances in (B)-(E), bootstrap algorithm was used to estimate the mean values and standard deviations by using 20 samples. Each sample contains 20 trajectories randomly selected with replacement from the ensemble. (F) Experimental results for the viral RNA copy number under increasing concentration of the NTP analogues *in vitro* using live SARS-CoV-2 virus infecting Vero E6 cells.

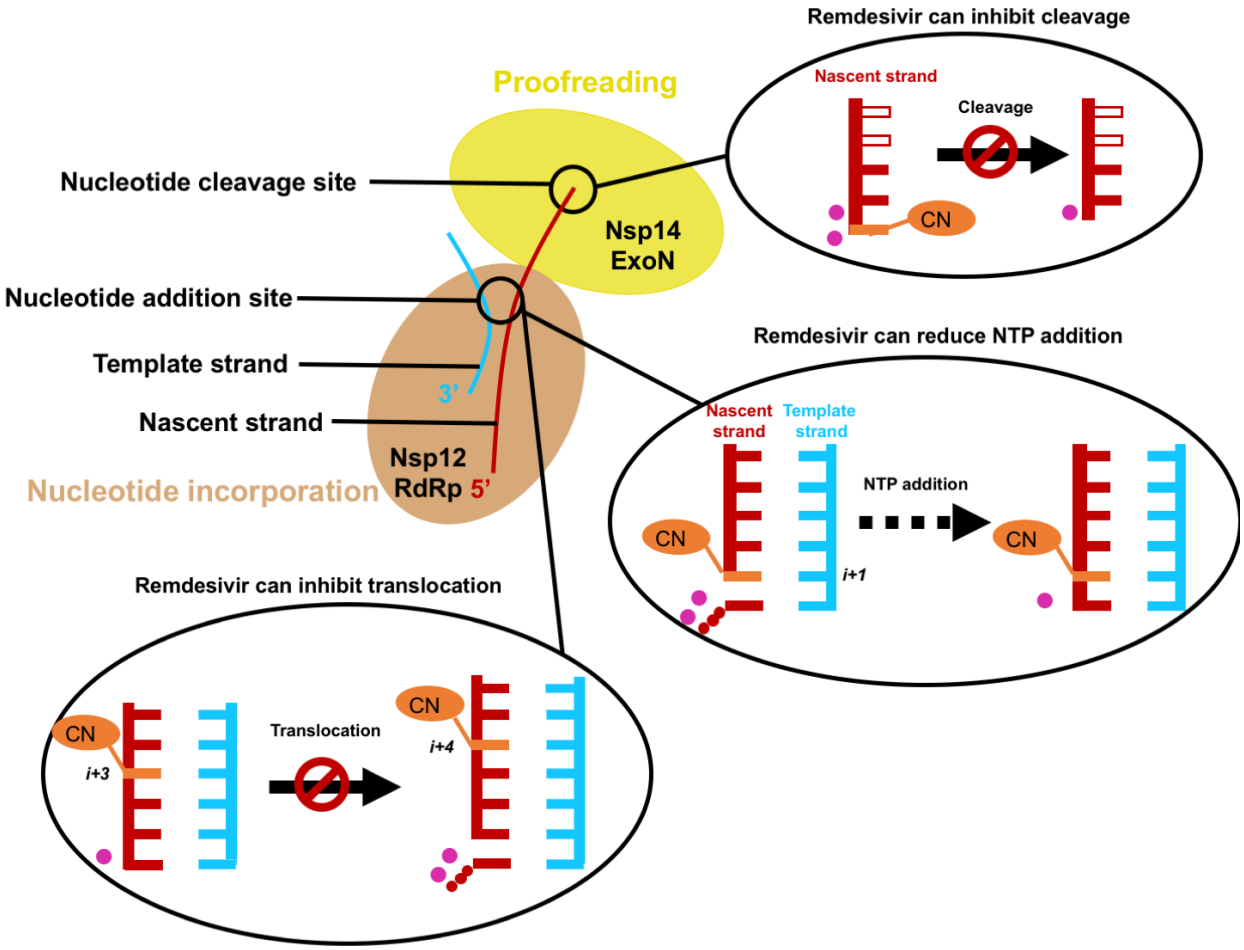

Fig. 7 Cartoon model of remdesivir's inhibition mechanisms in SARS-CoV-2 viral RNA replication.