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], Yongfang Li[1,2], Xiaowei Wang[3,4], Xilin Jia[1,2], Xin Gao[5], Hui-Lin Yen[6]*, Peter 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. C.) or xuhuihuang@ust.hk (X. H.)

Abstract

COVID-19 has 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. 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 RdRp exerts high drug efficacy *in vitro* and *in vivo*. However, the underlying molecular mechanisms for remdesivir to inhibit nucleotide addition and cleavage remain elusive. Here, we performed all-atom molecular dynamics (MD) simulations with an accumulated simulation time of 12 microseconds to elucidate molecular mechanisms for the inhibitory effects of remdesivir in nucleotide addition (RdRp complex: nsp12, nsp7, and nsp8) and cleavage (ExoN complex: nsp14 and nsp10). We found that the 1'-cyano group of remdesivir possesses the dual function of inhibiting both nucleotide addition by nsp12 and proofreading by nsp14. For nucleotide addition, we showed that incorporation of one remdesivir is not sufficient to terminate RNA synthesis. Instead, the presence of the 1'-cyano group of remdesivir at an upstream site induces electrostatic repulsion with the negatively charged Asp865. These repulsive interactions render translocation unfavourable and eventually lead to the delayed chain termination for RNA synthesis. For proofreading, remdesivir can inhibit proofreading due to the steric clash between its 1'-cyano group and Asn104. To further elucidate the role of 1'-cyano substitution in remdesivir's inhibitory effect, we also simulated three other NTP analogues without 1'-cyano substitution but containing various other modifications on the base and ribose: favipiravir, vidarabine and fludarabine. Our MD simulations suggest that they are more prone to the ExoN cleavage. Our computational findings were consolidated by an *in vitro* assay using live SARS-CoV-2. The dose-response curves suggest that among these NTP analogs only remdesivir exerts substantial inhibition effects on the live virus growth. Our work has provided plausible mechanisms on how remdesivir inhibits viral RNA replication at the molecular level. The predictions from our simulations also warrant future structural and biochemical studies and guide rational drug design for new treatment of COVID-19.

Commented [XG1]: Give full name?

Introduction

The 2019 novel coronavirus (COVID-19 coronavirus or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)) has spread rapidly to cause serious outbreaks around the world. As of April 2020, 2019-nCoV has been reported in 210 countries and over 2 million cases have been confirmed, resulting in more than 150,000 deaths with an estimated mortality risk of 6.85%[1]. The new coronavirus 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 the 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, and nucleic acid-based therapies targeting virus gene expression, as well as various types of vaccines. Hence, clinically approved or investigative drugs for treating other diseases should be repurposed to treat COVID-19 patients. Coronavirus (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 viral receptor have been under investigation for treating SARS-CoV-2, including Chloroquine[5], REGN3051 and REGN3048 monoclonal antibodies[6], and Griffithsin[7]. In addition, new vaccines targeting the viral surface antigens are under intense research and 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. Indeed, inhibitors targeting polymerase of SARS-CoV-2 is currently at late stages of clinical trials. Other RdRp inhibitors are also considered as treatment options for COVID-19. They include favipiravir and Xofluza that have recently been approved as treatment options for influenza virus in the clinic[9].

Compared with other viral RdRp such as influenza virus and rhinovirus, SARS-CoV-2 RdRp is much more challenging for drug development due to its intrinsic proofreading exoribonuclease (ExoN) function. This is attributed to the dual functions of non-structural protein (nsp) nsp12 (nucleotide addition) and nsp14 (proofreading). Nsp12, along with cofactors nsp8 with nsp7 is involved in nascent RNA priming and nucleotide addition[10-16]. NTP analogues that inhibits nucleotide addition in nsp12 supposedly should inhibit RNA replication. However, in SARS-CoV-2, ExoN in 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. 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 hospitalized with COVID-19[22], including those suffering from pneumonia[20]. Subsequently, clinical trials across Asia and several other countries around the world ~~has~~ 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. 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 dynamics of the nucleotide addition cycle of either NTP or NTP analogues still remain elusive. Biochemistry assays have shown in Ebola and MERS CoV[24] and recently in SARS-CoV-2[15] that remedivir exhibits "delayed chain termination" of RNA synthesis, wherein several rounds of nucleotide additions can still proceed downstream after remdesivir incorporation into the nascent strand before complete termination occurs. However, how remdesivir inhibits nucleotide addition in nsp12 of SARS-CoV-2 remains unknown. Furthermore, it is not clear how remdesivir can inhibit the proofreading activity of ExoN to prevent itself from being cleaved in order to exert its effects on RNA synthesis.

To determine the mechanisms of remdesivir inhibition on RdRp and ExoN, we conducted an aggregation of 12-microseconds all-atom molecular dynamics (MD) simulations based on the modelled SARS-CoV-2 nsp12-nsp7-nsp8 and nsp14-nsp10 complexes. Our work provides mechanistic insights of how remedivir impacts nucleotide addition and cleavage in the active sites of nsp12 and nsp14, respectively. Comparing with a series of other NTP analogues, we elucidated that the unique chemical feature of remdesivir lies in its 1'-cyano group that can inhibit both nucleotide addition and proofreading. Firstly, we 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 result in a delayed chain termination. Secondly, we found that remdesivir can effectively inhibit the exonuclease activity of SARS-CoV-2 nsp14-nsp10. Specifically, the relatively bulky 1'-cyano substitution on the ribose of remdesivir introduces steric interactions and effectively disrupts the stability of the catalytic cleavage site for the ExoN reaction. MD simulations of other NTP analogues including favipiravir, vidarabine and fludarabine show that they are all prone to be cleaved by ExoN. Our computational results are supported by *in vitro* assay to examine the efficiency of NTP analogues in inhibiting SARS-CoV-2 live virus growth. Only remdesivir shows inhibiting effect, consistent with our MD simulations. Our work provides novel mechanistic insights into the inhibition of SARS-CoV-2 replication by remdesivir in nucleotide addition and cleavage. This can serve as a platform for future antiviral drug development to interfere with SARS-CoV-2 transcription and replication.

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]) (Fig. 1A). Template and nascent strands have been modelled by the structural alignment to the norovirus RdRp (PDBID: 3H5Y[25]) (see Methods and SI 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*+1 site) is occupied by adenosine triphosphate (ATP) (Fig. 1A and 1C). To validate this model (named

as "wildtype-RNA"), we first confirmed that ATP is indeed stable in the active site. Specifically, the stability of the active site has been assessed by calculating the distance between the α phosphate (Pα) of ATP and O3' atom of the 3'-terminal nucleotide (Fig. 2A). Such distance required for the phosphodiester bond formation was found to be 3.4 Å from crystal structures of other RNA polymerases[25] and 3.5~4.0 Å from previous computational studies[26,27]. Consistent with our model, we found that the histogram 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[28]. 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 for the base pairing in the active site of crystal structures of polymerases of other viruses[25,29] (Fig. 2D). In addition, we investigated base stacking stability as the criteria for nucleotide addition[28]. We found that the distance between the base of NTP and the base of 3'-terminal nucleotide is populated at ~4.5 Å (Fig. 2G), consistent with the distance of 4.2 Å as observed in norovirus[25] and poliovirus[29] RdRps. We further validated our structural model against three recently solved cryo-EM structures of RdRp[14,15] (see Fig. 3 and SI for details). For the two apo SARS-CoV-2 structures, the root mean square deviation (RMSD) of Cα atoms was used as comparison to our model. We found that the averaged RMSD are 2.36 Å (PDBID: 6M71)[14] and 2.30 Å (PDBID: 7BV1[15]), respectively, indicating that the overall architecture of our nsp12 model is similar to that of apo cryo-EM structures of SARS-CoV-2. For the holo structure, the RMSD of both Cα atoms and nucleotides between our holo structural model and the holo cryo-EM structure (PDBID: 7BV2[15]) were used, and they were 2.12 Å and 2.23 Å, respectively. This further consolidated our wildtype-RNA model. Finally, we also validated the interactions between upstream RNA and proteins, and most (73 %) of the interactions observed in the holo cryo-EM structure are well maintained in our simulations (Fig. S1). The differences observed in the remaining distances between the cryo-EM structure and our models were only less than 0.5 Å. Therefore, our model is consistent with the available cryo-EM structures and could serve as structural basis to investigate the inhibitory mechanism of remdesivir.

We next performed 20 replicas of 30ns MD simulations with remdesivir-TP replacing ATP at the active site (*i*+1 site). We found that remdesivir-TP can satisfy the critical structural criteria for nucleotide incorporation. In particular, the critical distance between Pα of remdesivir-TP and O3' atom of the 3'-terminal nucleotide for phosphodiester bond formation (~3.9 Å) is larger and its distribution is broader when compared to wildtype ATP (see the orange curve in Fig. 2A). These observations suggest that the incorporation efficiency for remdesivir-TP could be lower than the wildtype ATP. However, no obvious negative effect on base pairing and base stacking stability is induced by remdesivir-TP at *i*+1 site (Fig. 2D and 2G). These observations suggest that remdesivir-TP can maintain the catalytical~~ly~~ active site conformation and thus can be added to the 3'-end of the growing RNA chain, which is a pre-requisite condition for remdesivir 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 next nucleotide addition at the active site. In this case, 20 replicas of 30 ns MD simulations were performed with remdesivir at 3'-terminal (*i*-1 site) to examine ATP incorporation at the active site. 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). The base pairing stability at the active site is seldom~~ly~~ affected by remdesivir at the *i*-1 site, with the histogram of the corresponding distance almost identical to that of wildtype-RNA (Fig. 1E). In contrast, the base stacking stability is weakened, with more MD conformations showing larger base-to-base

distance than wildtype-RNA (Fig. 1H). The above results suggest that remdesivir at the *i*-1 site can reduce the efficiency of the next NTP addition at the *i*+1 site. To determine if remdesivir at further upstream positions of nascent strand can interfere with NTP incorporation, we performed twenty 30-ns MD simulations for each of the systems with remdesivir at the *i*-2, *i*-3 and *i*-4 sites. 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 the *i*-2, *i*-3 or *i*-4 sites has negligible effects on the nucleotide addition at the active site. Therefore, remdesivir-TP satisfies the criteria of structural features to be incorporated into the nascent strand, and subsequently will reduce the efficiency of addition of the next incoming NTP. However, we expect that this reduction may not be sufficient to completely abolish RNA synthesis, suggesting other mechanisms existing for remdesivir to inhibit the 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[30], Nipah virus[31], and MERS-CoV[24]. Delayed chain termination occurs when natural NTP incorporation is not immediately prohibited after remdesivir incorporation into the nascent strand. Instead, several rounds of nucleotide addition can still proceed before the complete abolishment of RNA synthesis. Recent biochemical assay for SARS-CoV-2 also demonstrated delayed chain termination by remdesivir-TP at physiological concentrations, although nucleotide addition may be terminated earlier only when excess amount of remdesivir-TP is present. Remarkably, for MERS-CoV[24] and SARS-CoV-2[15] 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 (e.g. 3-nucleotide delay for MERS-CoV[24]). The molecular mechanisms underlining this delayed chain termination remains largely elusive.

To provide mechanistic insights into "delayed" termination, we performed MD simulations for RdRp in both the pre- and post-translocation states (pre-T and post-T, respectively) with remdesivir at different sites of the nascent RNA strand. For the pre-T state (with remdesivir at i+1, i-1, or i-2 site) and post-T state (with remdesivir at i-1, i-2, or i-3 site) states, 30 replicas of 20 ns simulations were performed. We first investigated the stability of pre-T and post-T during the translocation of 3'-terminal remdesivir from the *i*+1 site to the *i*-1 site (Fig. 4D). It is worth mentioning that the interactions between remdesivir and the surrounding residues in our model of the pre-T state with remdesivir at the *i*+1 site are consistent with those in the recently solved SARS-CoV-2 RdRp holo structure[15] (Fig. S3), validating our model for the subsequent investigation of translocation of nascent and template strands. To investigate the propensity of translocation, we compared the stability of pairs of the pre-T and post-T states (Fig. 4A, 4C and 4D) by examining the hydrogen bond probability of base pairs in the *i*+1, *i*-1 up to the *i*-4 sites. We found that the pre-T and post-T states have similar base pairing stability, enabling the forward translocation of remdesivir from the *i*+1 to the *i*-1 site to make the active site available for the next cycle of nucleotide addition. This is consistent with the recent finding based on a structural model of nsp12, which also suggests that remdesivir has enough space to allocate its 1'-cyano group at the *i*+1 site to facilitate translocation[32]. We then investigated translocation with the movement of remdesivir from the *i*-1 to the *i*-2 site and also found no significant discrepancy in the base pairing stability of the pre-T and post-T states, hence allowing translocation to occur (Fig. 3C). Strikingly, we found that forward translocation is thermodynamically unfavored only when remdesivir moves from the *i*-2 site to the *i*-3 site (Fig. 3A). In particular, the hydrogen bond probability for the pre-T state (with remdesivir at the *i*-2 site) is apparently higher than that for the post-T state (with remdesivir at the *i*-3 site), suggesting that the pre-T state is thermodynamically more stable than the post-T state. The

discrepancy of the base pairing stability is mainly observed when remdesivir is at the *i*-3 site of the post-T state. Therefore, remdesivir most likely accounts for such instability.

To elucidate the origin of the observed base pair instability with remdesivir at the *i*-3 site, we performed additional analysis based on the ensemble of MD conformations. We found that it is the electrostatic repulsion caused by the 1'-cyano group of remdesivir that is responsible for the instability. When remdesivir is at the *i*-3 site, its partially negatively charged 1'-cyano moiety exhibits anion-δ^- repulsion with the negatively charged Asp865. As remdesivir is confined by its upstream and downstream nucleotides in the nascent RNA strand, this unfavored anion-δ^- repulsion energy can only be released by destabilizing the conformation of remdesivir. Indeed, the hydrogen bonds for the base pairing of remdesivir:U are subsequently disrupted, rendering the corresponding post-T state thermodynamically unfavorable and further inhibiting translocation (Fig. 3A). Cryo-EM structure also shows that Asp865 is indeed in spatial proximity with the *i*-3 nucleotide, with even a closer distance when we modelled remdesivir at the *i*-3 site (see Fig. S4A and S4B). Comparison of the sequences between SARS-CoV-2 and other coronaviruses indicates that Asp865 is highly conserved (Fig. S4C), implying that similar mechanism may be employed by other CoV RdRp. This further supports our observations that the 1'-cyano group of remdesivir plays an important role in inhibiting forward translocation and subsequent nucleotide addition.

Even though we proposed a model where translocation from the *i*-2 to the *i*-3 site is hindered, our model does not exclude the possibility that the inhibition of translocation may also occur when remdesivir moves from the *i*-3 to the *i*-4 site, which may result in a delayed termination of three nucleotides as observed in MERS-CoV[24]. Under this condition, anion (Asp865)-δ^- repulsion can not only destabilizes the post-T state with remdesivir at the *i*-3 site, but also destabilizes the transition state for the translocation with remdesivir from the *i*-3 site to the *i*-4 site. However, this possibility is beyond the scope of our current study and methods such as Markov State Model[33-38] (MSM) may help explore this possibility in the future. We also note that a recent study based on the homology structural model of SARS-CoV nsp12 provided an alternative mechanistic explanation for the delayed termination by remdesivir[32]. They proposed that RdRp translocation is inhibited due to the steric clash between Arg858 and remdesivir's ribose at the *i*-4 site. However, our simulations with remdesivir at the *i*-4 site show a large distance (~ 11.4 Å) between the amino groups of Arg858 and the nitrogen atom of 1'-ribose cyano group of remdesivir (Fig. S5B). 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 the *i*-4 site using the recent cryo-EM structure (PDBID: 7BV2[15], see Fig. S5C). Finally, our simulations also show a negligible impact of remdesivir (at the *i*-4 site) on the stability of base pairing from the *i*+1 up to the *i*-4 sites (Fig. S5A). Altogether, we anticipate that the steric clash between Arg858 and remdesivir's ribose at the *i*-4 site may be unlikely to occur.

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 to remove incorporated NTP analogues in order to counteract the efficacy of the inhibitory compounds. In this regard, an additional criterion to be an effective inhibitor against SARS-CoV-2 RdRp is to escape from the cleavage by ExoN. After confirming that remdesivir can inhibit RdRp, we next examine how efficient remdesivir is being cleaved by ExoN.

We constructed the structural model for nsp14 bound with its activator nsp10 based on the crystal structure of nsp14-nsp10 of SARS-CoV (PDBID: 5C8S[39]), the ExoN domain of which resembles that in a separate structure[16]. As three base pairs are proposed to be melted in the cleavage site of ExoN[40], we modelled a single-stranded RNA containing three nucleotides based on structural alignment to DNA polymerase I Klenow Fragment[40] and DNA Polymerase III ε subunit[41] (see Method and SI for details). Twenty 30-ns MD simulations were performed for the wildtype-RNA, in which the 3'-terminal is occupied by adenine. To assess the stability of cleavage site, we determined the distance between MgA and the O3' atom of the second last nucleotide at the 3'-terminal as it is essential for cleavage[42-44] (Fig 5A). This distance is well maintained in most of the simulations (Fig. S6), and the histogram shows only one peak in the region of 2.0~2.5 Å (Fig. 5A) suitable for cleavage[42-44]. Mg-Mg distance is also critical for cleavage[42-45] and our calculations show a sharp peak at ~ 3.6 Å (Fig. 5B), a distance permissible for cleavage[42-45]. By sharp contrast, 20 replicas of 30ns MD simulations with remdesivir at 3'-terminal show destabilization of the cleavage site in ExoN, with the coordination between MgA and O3' atoms more frequently disrupted (Fig. S7). Consistently, a second peak appears at ~3.7Å in the histogram 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 Mg-Mg distance with a second peak clearly appearing at 4.5~5.2 Å (Fig. 5B). Examination of the base stepping twisting angle (Fig. S8) 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 show 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 of 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). Two representative conformations clearly demonstrate this steric repulsion (Fig. 5F). On one hand, if remdesivir adopts the similar conformation as wildtype-RNA at the 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 the cleavage site to the steric repulsion between 1'-cyano group of remdesivir and Asn104. This observation from MD ensemble 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 [32]. This destabilization would reduce the possibility of remdesivir to be excised by Exon, enabling RVD to be an effective inhibitor for SARS-CoV-2.

Elucidations of important chemical structural 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 another three 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). Fludarabine has modifications in both the base and ribose that are similar to remdesivir. 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 (the *i*+1 site) 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 greatly raised to ~ 5.0 Å for fludarabine-TP, implicating the difficulty of fludarabine to be incorporated into the nascent strand. 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 2'-oxygen atom of the template uracil, 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 extended by < 0.25 Å for all the NTP analogues compared to wildtype-RNA (Fig. S9A). Similarly, the mean base stacking distances for the NTP analogues are equal to or even shorter than that for ATP. These observations suggest that all NTP analogues except fludarabine-TP satisfy the structural feature criteria to be incorporated into the nascent strand.

Secondly, we examined if these NTP analogues can inhibit nucleotide addition in RdRp by placing the NTP analogue at 3'-terminal (the *i*-1 site) with ATP in the active site (we performed twenty 30-ns MD simulations for each system). We found that the averaged O3'-Pα distance is ~ 3.6 Å when favipiravir is at the *i*-1 site, similar to that when remdesivir is at the *i*-1 site (Fig. 6C), demonstrating its 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. Base-to-base distance between NTP and the template nucleotides are not perturbed by the presence of the NTP analogue at the *i*-1 site (Fig. S9B), as even the most significant difference is only ~ 0.05 Å observed for fludarabine. Favipiravir shows similar base stacking distance as wildtype-RNA, while this distance is marginally increased by around 0.3~0.4 Å for vidarabine and fludarabine. We also noticed that none of these three NTP analogues has a substitution similar to the 1'-cyano group of remdesivir (Fig. 6A), thus they are not expected to inhibit translocation. Overall, favipiravir can reduce NTP incorporation to a similar extent as remdesivir, while vidarabine and fludarabine at the *i*-1 site 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. For each of these NTP analogues, we calculated the O3'-MgA distances, which were comparable to that when wildtype adenine is at the 3'-terminal (Fig. 6D). Furthermore, the mean distance is shorter than that of 3'-remdesivir, implicating that they are more prone to be excised by ExoN. These three NTP analogues adopt similar configuration as adenine in the cleavage site (Fig. S10) and ~~don't~~ do not experience any steric clash~~es~~ with Asn104 (Fig. 6E). The stability of the three compounds at the cleavage site can be attributed to the lack of a bulky group at the ribose. This is consistent with our proposal that the bulky 1'-cyano group causes steric effect, contributing to the

instability of the cleavage site. The Mg-Mg distance is consistent with our conclusion (Fig. S11A), wherein the three NTP analogues show significant base-to-base twisting angle compared with adenine (Fig. S11B). Altogether, these results suggest that all 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 further performed *in vitro* experiments 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 4 NTP analogues tested, only remdesivir was found to inhibit SARS-CoV-2 replication in Vero E6 cells with EC50 under 100 μM (Figure 6J). This is consistent with our previous work[46], 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 (Figure 6F).

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~~ Fludarabine has the lowest tendency to be incorporated into the nascent strand, while vidarabine could not reduce the efficiency of incoming NTP incorporation. ~~favipiravir~~ 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 ExoN.

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 the *i*-1 site can reduce nucleotide addition at the *i*+1 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 stopped a few nucleotides after remdesivir incorporation. We proposed two mechanisms explaining the delayed chain termination. Firstly, because the incorporation of one remdesivir at the *i*-1 site could only reduce NTP incorporation rate to a limited extent, multiple rounds of remdesivir incorporations may be required to fully terminate RNA chain extension. A similar phenomenon has been observed in Ebola virus[24]. On the other hand, the electrostatic repulsion between the partially negatively charged 1'-cyano group of remdesivir at the *i*-3 site and the nearby negatively charged Asp865 could lead to the instability of the RdRP structure at the post-T state, which will prohibit translocation and terminate subsequent nucleotide addition. Furthermore, 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 cleavage site, rendering it resistant to ExoN cleavage. Further investigation using three other NTP analogues show 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 inhibitor 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[25]) were used as the structural basis to construct the nCoV RdRp containing dsRNA strands with ATP bound at the active site (see SI for details). The protonation states of histidines were predicted using propka3.0 module[47] in the pdbpqr2.2.1[48] 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[49], 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[39]) serves as the structural basis to construct the nsp10-nsp14 of SARS-CoV-2 (see SI for details). Missing residues in nsp14 ~~was~~ were modelled by modeller9.21[50]. 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[40]) and the ε-subunit of DNA polymerase III (PDBID: 1J53[41]) due to the similar architecture of the cleavage site. The protonation states of histidines were estimated by propka3.0 module[47] of the pdbpqr2.2.1[48] package. The complex was solvated with TIP3P water in a dodecahedron box, the edge of which was at least 12Å 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[51] to simulate protein and nucleotides. The force field parameters for NTP analogues were derived based on the existing amber99sb-ildn force field parameters or the general amber force field[52,53] (see SI for details). Partial charges of NTP analogues were generated following the Restricted Electrostatic Potential scheme[54,55] (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.*[56]. The system was gradually relaxed before the production simulation under NVT ensemble (T=300K) (see SI for details). V-rescale thermostat[57] 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.[58] 12 Å was used as the cut-off for Lennard-Jones interactions. The neighbo~~u~~r list was updated every 10 steps. An integration time step of 2.0 fs was used and the LINCS algorithm[59] was applied to constrain all the bonds. We saved the snapshots every 20 ps. All simulations were performed with Gromacs 5.0[60]. The equilibrated system containing wildtype RNA serves as the structural basis for modelling NTP analogue at the corresponding site to investigate their inhibitory effects (see SI for details).

In vitro **assays with live SARS-CoV-2 virus for the determination of EC50 of NTP analgoues** 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 were obtained from MedChemExpress and the stocks were prepared with DMSO (50 mM remdesivir, 100 mM favipiravir, 10 mM fludarabine phosphate, 10 mM

vidarabine). 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 innculation by SARS-CoV-2 virus at Multiplication 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. 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.*, 2020).
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.* **3**, 237-261, (2016).
4. Yuen, K.-S. *et al.* SARS-CoV-2 and COVID-19: The most important research questions. *Cell Biosci.* **10**, 40, (2020).
5. Vincent, M. J. *et al.* Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virol. J.* **2**, 69, (2005).
6. Pascal, K. E. *et al.* 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.* **112**, 8738, (2015).
7. Keefe, B. R. *et al.* Broad-Spectrum In Vitro Activity and In Vivo Efficacy of the Antiviral Protein Griffithsin against Emerging Viruses of the Family <em>Coronaviridae</em&gt. *J. Virol.* **84**, 2511, (2010).
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.* **115**, 12878-12886, (2018).
9. Lythgoe, M. P. & Middleton, P. Ongoing Clinical Trials for the Management of the COVID-19 Pandemic. *Trends Pharmacol. Sci.*, (2020).
10. Imbert, I. *et al.* A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* **25**, 4933-4942, (2006).
11. Subissi, L. *et al.* One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E3900, (2014).
12. Kirchdoerfer, R. N. & Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* **10**, 2342, (2019).
13. Zhai, Y. *et al.* Insights into SARS-CoV transcription and replication from the structure of the nsp7–nsp8 hexadecamer. *Nat. Struct. Mol. Biol.* **12**, 980-986, (2005).
14. Gao, Y. *et al.* Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*, eabb7498, (2020).
15. Yin, W. *et al.* Structural Basis for the Inhibition of the RNA-Dependent RNA Polymerase from SARS-CoV-2 by Remdesivir. *bioRxiv*, 2020.2004.2008.032763, (2020).
16. Ferron, F. *et al.* Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E162, (2018).
17. Eckerle, L. D. *et al.* Infidelity of SARS-CoV Nsp14-Exonuclease Mutant Virus Replication Is Revealed by Complete Genome Sequencing. *PLoS Pathog.* **6**, e1000896, (2010).
18. Smith, E. C. *et al.* Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. *PLoS Pathog.* **9**, e1003565, (2013).
19. Agostini, M. L. *et al.* Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio* **9**, e00221-00218, (2018).
20. Holshue, M. L. *et al.* First Case of 2019 Novel Coronavirus in the United States. *N. Engl. J. Med.* **382**, 929-936, (2020).
21. Wang, M. *et al.* Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **30**, 269-271, (2020).
22. Grein, J. *et al.* Compassionate Use of Remdesivir for Patients with Severe Covid-19. *N. Engl. J. Med.*, (2020).
23. Sheahan, T. P. *et al.* Broad-spectrum antiviral GS-5734 inhibits both epidemic and zoonotic coronaviruses. *Sci. Transl. Med.* **9**, eaal3653, (2017).

24. Gordon, C. J. *et al.* The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J. Biol. Chem.* **295**, 4773-4779, (2020).
25. Zamyatkin, D. F. *et al.* Binding of 2′-Amino-2′-Deoxycytidine-5′-Triphosphate to Norovirus Polymerase Induces Rearrangement of the Active Site. *J. Mol. Biol.* **390**, 10-16, (2009).
26. 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* **116**, 2259-2268, (2012).
27. Radhakrishnan, R. Coupling of fast and slow modes in the reaction pathway of the minimal hammerhead ribozyme cleavage. *Biophys. J.* **93**, 2391-2399, (2007).
28. Huang, X. *et al.* RNA polymerase II trigger loop residues stabilize and position the incoming nucleotide triphosphate in transcription. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15745, (2010).
29. 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.* **107**, 22505, (2010).
30. Tchesnokov, E. P. *et al.* Mechanism of Inhibition of Ebola Virus RNA-Dependent RNA Polymerase by Remdesivir. *Viruses* **11**, 326, (2019).
31. Jordan, P. C. *et al.* Initiation, extension, and termination of RNA synthesis by a paramyxovirus polymerase. *PLoS Pathog.* **14**, e1006889, (2018).
32. Shannon, A. *et al.* Remdesivir and SARS-CoV-2: Structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites. *Antiviral Res.* **178**, 104793, (2020).
33. Klippenstein, S. J. *et al.* Chemical Kinetics and Mechanisms of Complex Systems: A Perspective on Recent Theoretical Advances. *J. Am. Chem. Soc.* **136**, 528-546, (2014).
34. Chodera, J. D. & Noe, F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **25**, 135-144, (2014).
35. Bowman, G. R. *et al.* Using generalized ensemble simulations and Markov state models to identify conformational states. *Methods* **49**, 197-201, (2009).
36. Silva, D. A. *et al.* Millisecond dynamics of RNA polymerase II translocation at atomic resolution. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7665-7670, (2014).
37. Zhang, L. *et al.* Elucidation of the Dynamics of Transcription Elongation by RNA Polymerase II using Kinetic Network Models. *Acc. Chem. Res.* **49**, 687-694, (2016).
38. Wang, W. *et al.* Constructing Markov State Models to elucidate the functional conformational changes of complex biomolecules. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**, (2018).
39. Ma, Y. *et al.* Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 9436, (2015).
40. Beese, L. S. *et al.* Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* **260**, 352, (1993).
41. Hamdan, S. *et al.* Structural Basis for Proofreading during Replication of the Escherichia coli Chromosome. *Structure* **10**, 535-546, (2002).
42. Lan, P. *et al.* Structural insight into precursor tRNA processing by yeast ribonuclease P. *Science* **362**, eaat6678, (2018).
43. Lei, J. *et al.* 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.* **116**, 845, (2019).
44. Tse, C. K. M. *et al.* Intrinsic cleavage of RNA polymerase II adopts a nucleobase-independent mechanism assisted by transcript phosphate. *Nat. Catal.* **2**, 228-235, (2019).
45. Yang, W. *et al.* Making and Breaking Nucleic Acids: Two-Mg²⁺-Ion Catalysis and Substrate Specificity. *Mol. cell* **22**, 5-13, (2006).
46. Choy, K.-T. *et al.* Remdesivir, lopinavir, emetine, and homoharringtonine inhibit SARS-CoV-2 replication in vitro. *Antiviral Res.* **178**, 104786, (2020).
47. Olsson, M. H. M. *et al.* PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **7**, 525-537, (2011).
48. Dolinsky, T. J. *et al.* PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Research* **32**, W665-W667, (2004).

49. Jorgensen, W. L. *et al.* Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **79**, 926-935, (1983).
50. Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinf.* **54**, 5.6.1-5.6.37, (2016).
51. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Bioinf.* **78**, 1950-1958, (2010).
52. Wang, J. M. *et al.* Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157-1174, (2004).
53. Wang, J. *et al.* Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **25**, 247-260, (2006).
54. Woods, R. J. & Chappelle, R. Restrained electrostatic potential atomic partial charges for condensed-phase simulations of carbohydrates. *Theochem* **527**, 149-156, (2000).
55. Wang, J. *et al.* How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **21**, 1049-1074, (2000).
56. Meagher, K. L. *et al.* Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.* **24**, 1016-1025, (2003).
57. Bussi, G. *et al.* Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101, (2007).
58. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593, (1995).
59. Hess, B. *et al.* LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463-1472, (1997).
60. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19-25, (2015).