 doi:10.20944/preprints202003.0267.v1

Binding mechanism of remdesivir to SARS-CoV-2 RNA dependent RNA polymerase

Leili Zhang[1] and Ruhong Zhou[1,2]*

[1]*Computational Biology Center, IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA*

[2]*Department of Chemistry, Columbia University, New York, NY 10027, USA*

*All Correspondence should be addressed to: rz24@columbia.edu

Abstract

Starting from December 2019, coronavirus disease 2019 (COVID-19) has emerged as a once-in-a-century pandemic with deadly consequences, which urgently calls for new treatments, cures and supporting apparatuses. Remdesivir was reported by World Health Organization (WHO) as the most promising drug currently available for the treatment of COVID-19. Here, we use molecular dynamics simulations and free energy perturbation methods to study the inhibition mechanism of remdesivir to its target SARS-CoV-2 virus RNA-dependent RNA polymerase (RdRp). In the absence of a crystal structure of the SARS-CoV-2 RdRp, we first construct the homology model of this polymerase based on a previously available structure of SARS-CoV NSP12 RdRp (with a sequence identify of 95.8%). We then build the putative binding mode by aligning the remdesivir + RdRp complex to the ATP bound poliovirus RdRp. The putative binding structure is further optimized with molecular dynamics simulations and demonstrated to be stable, indicating a reasonable binding mode for remdesivir. The relative binding free energy of remdesivir is calculated to be -8.28 ± 0.65 kcal/mol, much stronger than the natural substrate ATP (-4.14 ± 0.89 kcal/mol) which is needed for the polymerization. The ~800-fold improvement in the K_d from remdesivir over ATP indicates an effective replacement of APT in blocking of the RdRp binding pocket. Key residues D618, S549 and R555 are found to be the contributors to the binding affinity of remdesivir. These findings demonstrate that remdesivir can potentially act as a SARS-CoV-2 RNA-chain terminator, effectively stopping its RNA reproduction, with key residues also identified for future lead optimization and/or drug resistance studies.

Keywords:

COVID-19, SARS-CoV-2, RNA-dependent RNA polymerase (RdRp), remdesivir, homology model, molecular dynamics, free energy perturbation

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

Introduction

Coronavirus disease 2019 (COVID-19) is a disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) – a novel coronavirus that has been causing a once-in-a-century pandemic[1], threatening millions, if not billions, of people. SARS-CoV-2[2] belongs to the family *Coronaviridae*, which includes RNA viruses such as severe acute respiratory syndrome coronavirus (SARS-CoV[3], which caused a pandemic in 2003), Middle East respiratory syndrome-related coronavirus (MERS-CoV[4], which has caused a continuing pandemic since 2012), and also human coronaviruses[5] that cause symptom known as a "common cold". Without a decisive mortality rate yet (currently in the range of ~1-6%), COVID-19 is believed to be less deadly than SARS (~10%)[6] or MERS (~40%)[6]; however, its reproductive number (R_0) has been estimated to be 2.0-6.5[7], higher than SARS and MERS. COVID-19 has been spreading to all continents with multiple epicenters. While certain physical treatment has been shown to assist patients to fight this disease with their own immune systems[8], no proven remedies exist so far, causing high mortality rates especially in senior groups[9]. This raises high and urgent demand to screen for potential drugs through either drug-repurposing or novel drug development[10-14].

Remdesivir is a nucleotide analogue that mimics the structure of adenosine. It was originally developed by Gilead Sciences, Inc. to treat Ebola[15]. Even though it hasn't passed the phase 3 clinical trial of Ebola treatment, it showed moderately promising improvement over the mortality rate of this deadly disease[16]. In the case of Ebola, remdesivir was found to act as an RNA-dependent RNA polymerase (RdRp) binding substrate that replaces ATP in the polymerization before terminating this process[15], also known as a "chain terminator". The active form of remdesivir was found to be hydrolyzed and decorated with triphosphates, using the core of remdesivir as a nucleoside[15]. We term this hydrolyzed and phosphorylated remdesivir as "RemTP". Like other nucleotide analogues, remdesivir could potentially be utilized as a broad-spectrum antiviral drug[17] due to the structural similarities of RdRp's from various viruses[18]. It was clinically tested against MERS-CoV and showed significant efficacy[19]. Because of this, a phase 3 trial of remdesivir is currently under progress in China[20] and the U.S.[21] The new and rapid development of COVID-19 around the world raises pressing need of more information about the drug, including its detailed inhibition mechanism. Therefore, we carried out a physics-based molecular modeling study on the binding mechanism between remdesivir and SARS-CoV-2 RdRp. A previous study revealed that ATP serves as the main substrate to NSP12 of SARS-CoV RdRp[22] (the RdRp complex has multiple nonstructural protein (NSP) units, such as NSP12, NSP8, and NSP7), thus, we narrowed down our search of the RemTP binding pocket to NSP12 of SARS-CoV-2 RdRp (denoted as "COVID-19 NSP12"). In the absence of the crystal structure of COVID-19 NSP12, we employed homology modeling to first construct its tertiary structure. The initial binding mode of ATP to COVID-19 NSP12 was subsequently determined by structural alignment to the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

ATP bound poliovirus RdRp due to the structural resemblance of viral RdRp's. We then performed molecular dynamics (MD) simulations to validate the identified binding mode. Upon locating the stable binding mode, we further carried out free energy perturbation (FEP) calculations to not only estimate the binding affinity of RemTP and ATP to COVID-19 NSP12, but also identify the key residues in the binding process.

Method

Homology modeling

The structure of SARS-CoV NSP12 RdRp was obtained from protein data bank (PDBID: 6NUR)[18]. The sequence of SARS-CoV-2 NSP12 RdRp (COVID-19 NSP12) was obtained from entry YP_009725307.1 at NCBI[23]. Sequence alignment and homology modeling were performed with MODELLER 9.23[24], with unresolved structures on the N-terminus and C-terminus truncated (gray residues on Figure 1A), which should not affect our current binding affinity calculations.

Molecular dynamics (MD) simulation

Based on the co-crystal structure of poliovirus RdRp with ATP[25] (PDBID: 2ILY), we prepared the initial structure of COVID-19 NSP12-ATP complex by aligning the "fingers" domain of the two proteins (i.e., one from poliovirus and one from COVID-19). The complex structure of RemTP was then aligned with that of ATP for the corresponding simulations. A chelating Mg^{2+} ion is often needed for these RdRps, which is positioned in our simulations based on previous studies on class I RNA polymerase ribozyme[26]. A total of 9 different simulations were performed: one for the apo form of COVID-19 NSP12, two (independent runs) for COVID-19 NSP12-ATP (no Mg^{2+}), two for COVID-19 NSP12-ATP (with Mg^{2+}), two for COVID-19 NSP12-RemTP (no Mg^{2+}) and two for COVID-19 NSP12-RemTP (with Mg^{2+}). The parametrization of RemTP was generated by CHARMM CGenFF[27]. All other parameters were taken from CHARMM36[28]. Simulations were performed with GROMACS 5.1.2[29]. Van der Waals interactions were treated with a switching distance of 10 Å and a smooth cutoff distance of 12 Å. Electrostatic interactions were treated with Particle Mesh Ewald with a grid size of 1 Å. All simulations lasted about 100 ns before the free energy perturbation calculations.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

Free energy perturbation (FEP) calculation

The thermodynamic cycles of our FEP calculations were illustrated in Figure S1. A softcore potential was applied in FEP and more details can be found in our previous studies[30]. The λ windows were set as (0.00 0.00001 0.0001 0.001 0.01 0.05 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 0.95 0.99 0.999 0.9999 0.99999 1.00) to avoid the well-known FEP end-point catastrophe. NAMD 2.13[31] was used to perform the FEP calculations. The starting FEP complex structures of COVID-19 NSP12 binding with NTP (ATP or RemTP) were selected using the clustering algorithm based on the RMSD of NTP while COVID-19 NSP12 was aligned. The largest cluster of each simulation was taken to carry out the FEP calculations. All other simulation parameters were the same as MD simulations. Each simulation was repeated 5 times with distinct random seeds to obtain the standard error of the free energy changes.

Results and Discussions

Homology model of COVID-19 NSP12

We constructed the homology model of COVID-19 NSP12 by first aligning the sequences between COVID-19 NSP12 (Wu *et. al.*[23], NCBI: YP_009725307.1) and the recently resolved SARS-CoV NSP12 RdRp structure[18] (PDBID: 6NUR). The sequence identity was determined to be 95.8%, with 131 unresolved terminal residues, 10 unresolved hinge residues and 24 mutated residues (Figure 1A). The 3D structure of COVID-19 NSP12 was then created with MODELLER package without the 131 unresolved terminal residues. The 10 unresolved hinge residues (residues 896 – 905) were estimated with the default algorithm, thus resulting in a random coil configuration. Because of the high sequence conservation, the backbone RMSD between 6NUR NSP12 and COVID-19 NSP12 was a mere 0.25 Å (Figure 1B). The newly added random coil and 24 mutated residues have been highlighted in Figure S2. It was previously discovered that RdRp's of various viruses share the same morphology with common building blocks[18]: fingers domain, a thumb domain and a palm domain (Figure 1C). The grip (a hole formed between these three domains) served as the binding site for RNA and nucleotide triphosphates (NTP). Additionally, the SARS-CoV NSP12 also featured with a unique N-terminal extension, which was mostly conserved in COVID-19 NSP12 (colored by green in Figure 1C). Comparing Figure 1C and Figure S2, we notice that majority of the mutations occur at the N-terminal extension and the palm domain. Meanwhile, the fingers domain and the thumb domain remain highly conserved. Consequentially, the RNA/NTP binding grip is highly conserved. As a simple test on the structural stability, we ran a 100 ns simulation on the apo form of COVID-19 NSP12. The RMSD (Figure 1D) is always below 3 Å, indicating a high stability of the constructed protein.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

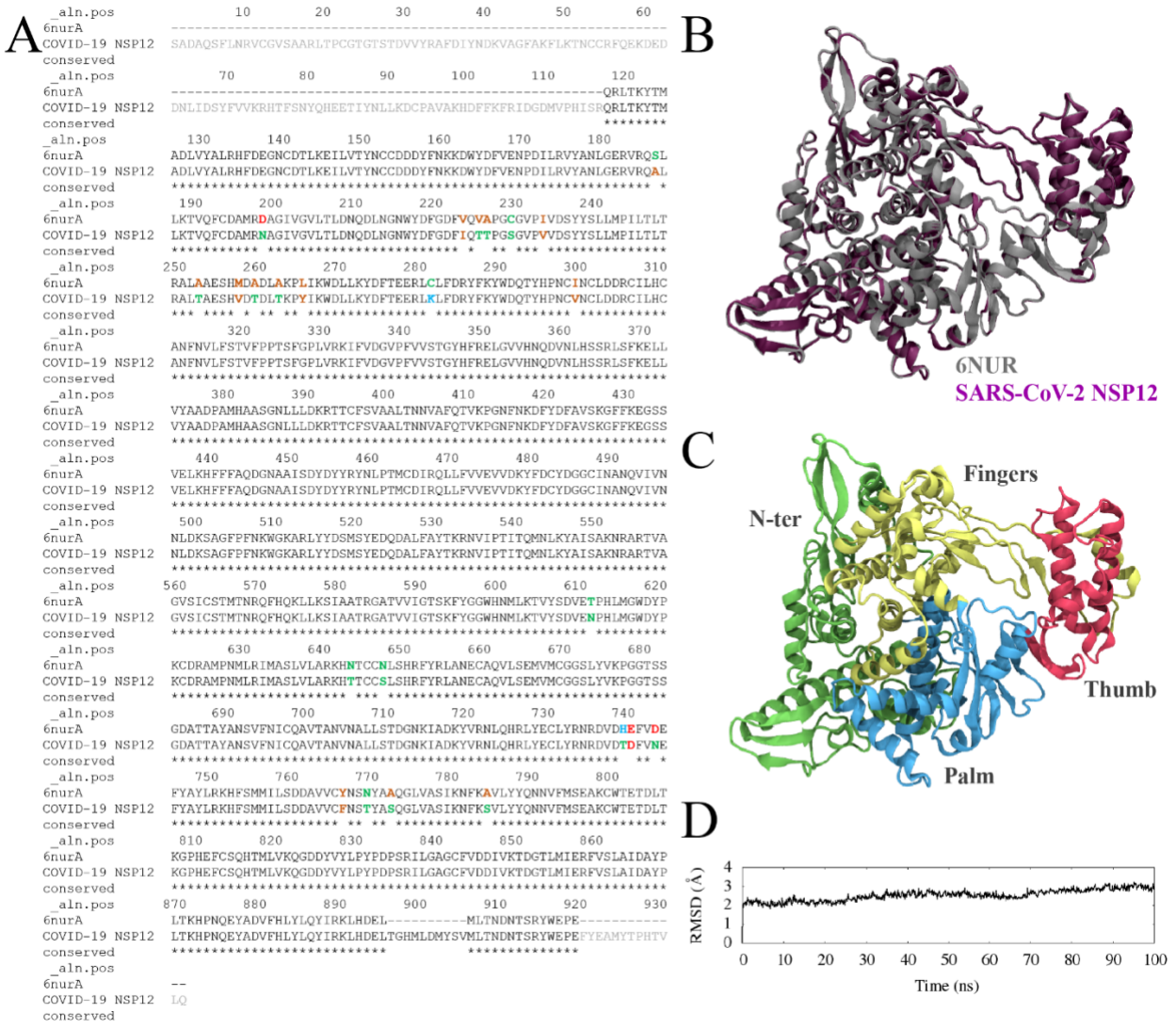

Figure 1. Sequence alignment between 6NUR (PDBID) and SARS-CoV-2 NSP12 RdRp (COVID-19 NSP12) is shown in (A). Conserved residues are colored in black and labeled with *. Mutated residues are colored with the following rules: brown as hydrophobic, green as hydrophilic, red as positively charged and blue as negatively charged. N-terminus and C-terminus are truncated in COVID-19 NSP12 due to the lack of structural information in 6NUR. Residues (r.) 896 – 905 are filled with a random coil. Homology model of COVID-19 NSP12 (magenta) is overlaid with 6NUR (silver) in (B). Based on the study on SARS-CoV NSP12 (6NUR), the domains of COVID-19 NSP12 are colored as the following in (C): green for N-terminal extension (r. 1–397, green), yellow for fingers domain (r. 398–581, 628–687), blue for the palm domain (r. 582–627, 688–815), and red for the thumb domain (r. 816–919). The RMSD from the simulation using apo form of COVID-19 NSP12 is plotted in (D).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

Binding mode of the substrates

It was previously identified that adenosine triphosphate (ATP) was the natural substrate of SARS-CoV NSP12[22]. Logically, ATP should be the natural substrate of COVID-19 NSP12 due to the conservation of the grip. However, no crystal structures exist yet for the ATP-bound SARS-CoV NSP12. We instead referred to another ATP-bound RdRp structure from poliovirus (PDBID: 2ILY) utilizing the structural conservation of viral RdRp's. The binding structure of ATP to poliovirus RdRp is shown in Figure 2A, with residues in close contact with ATP labeled. The triphosphate part of ATP mostly interacts with positively charged residues, such as K159, R174, R163, K167, K172 and K359. The negatively charged residue D323 in the vicinity implies the existence of a chelated ion (Mg^{2+}), suggested by other studies[26]. The nucleoside part of ATP (adenosine) interacts with more diverse residues, such as K61, I176, E177, D238 and S288, which are mostly hydrophilic with only one hydrophobic residue. To guesstimate the binding mode of ATP in COVID-19 NSP12, we aligned the fingers domain of poliovirus RdRp with that of COVID-19 NSP12 because this was the main constituent of the grip. The aligned COVID-19 NSP12 is shown in Figure 2B with both ATP and RemTP (NTP, note that residue numbers are drastically different because of the size of the protein: 932 residues in COVID-19 NSP12 versus 461 residues in 2ILY). As expected, the alignment ensured that triphosphates of NTP's are in close contact with positively charged residues K545, R551, R553 and R555. Nucleosides of NTP's are surrounded by residues like T556, V557, A558, C622, D623 and S682, which are still mostly hydrophilic with only some hydrophobic. Our alignment of the binding pocket here indicated once again the conservation of the binding grip from viral RdRp's.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

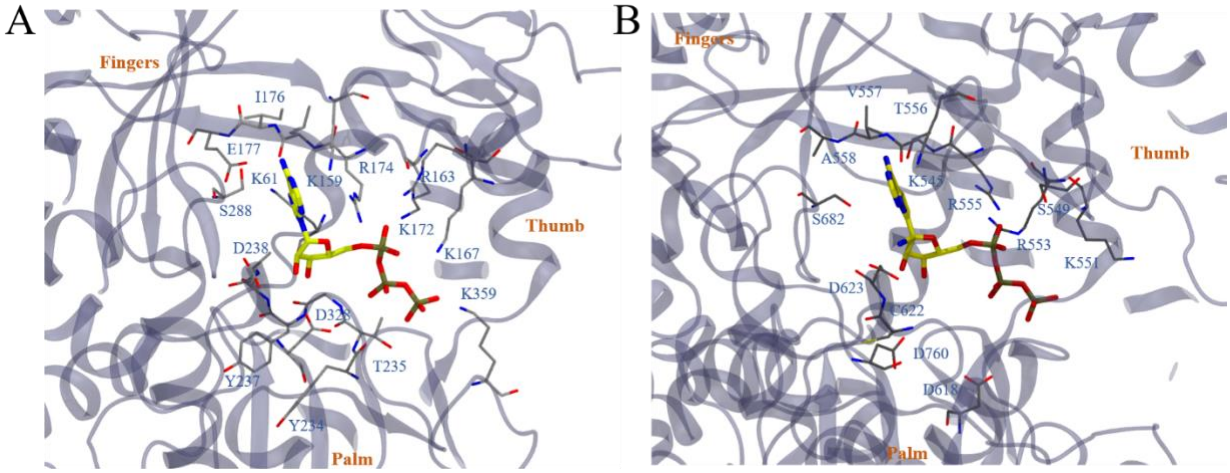

Figure 2. The binding mode of ATP to poliovirus RdRp (PDBID: 2ILY) is illustrated in (A). ATP molecule is highlighted in stick models with yellow (carbon), blue (nitrogen), red (oxygen) and tan (phosphorus). Residues that directly contact ATP (within 5 Å of the molecule) are illustrated in thin stick models colored by gray (carbon), blue (nitrogen), red (oxygen) and yellow (sulfur). The rest of the RdRp is illustrated with new cartoon model with VMD. Fingers domain, the thumb domain and the plam domain are labeled to guide the eyes. By aligning the fingers domain, the binding mode of ATP and RemTP to COVID-19 NSP12 is constructed and illustrated in (B). Note that in the figure the constructed ATP and RemTP completely overlap. All representations are the same as (A).

Binding mode optimization

MD simulations were then carried out from the initial binding complex structures prepared above. To testify the importance of Mg^{2+}, we ran two sets of simulations for both ATP and RemTP: one set without Mg^{2+} and one set with Mg^{2+}. The initial position of Mg^{2+} was constructed based on previous studies on class I RNA polymerase ribozyme[26].

We performed clustering analysis based on the RMSD of NTP's during the simulations, with COVID-19 NSP12 aligned. We noticed that by adding the Mg^{2+}, the cluster of ATP in the COVID-19 NSP12 binding grip was slightly larger than that without Mg^{2+} (Figure S1A vs S1B). This indicates that Mg^{2+} might have stabilized the binding between ATP and COVID-19 NSP12. As a result, the determined "bound state" (defined as the largest binding cluster observed in our simulations) was likely from the first set of simulations of COVID-19 NSP12-ATP with Mg^{2+} (Figure S1B, upper panel). A closer examination (Figure 4A) showed that the "bound state" did not deviate much from the initial structure (which was aligned to 2ILY). The triphosphate of ATP was found to bind S549, K551, R553, R555, R836 and Mg^{2+}. The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

adenosine of ATP was found to bind M542, T556, V557, A558 and S682, similar to its initial structure. This indicates a remarkable consistence between the binding modes of ATP to RdRp's from two species: poliovirus and SARS-CoV-2. Additionally, we plotted the solvent accessible surface of COVID-19 NSP12 upon binding to ATP (Figure 4B). ATP was found to reside well inside a local binding pocket within the grip. However, the "extra space" as seen on the right side of the adenosine group might suggest a possible druggable target. To further test this assumption, we performed molecular docking with several adenosine analogues (see more details in the Supporting Information). Interestingly, several molecules, including RemTP, with an enlarged nucleoside, occupied the top of the list ordered by docking scores (see Table S1 for details).

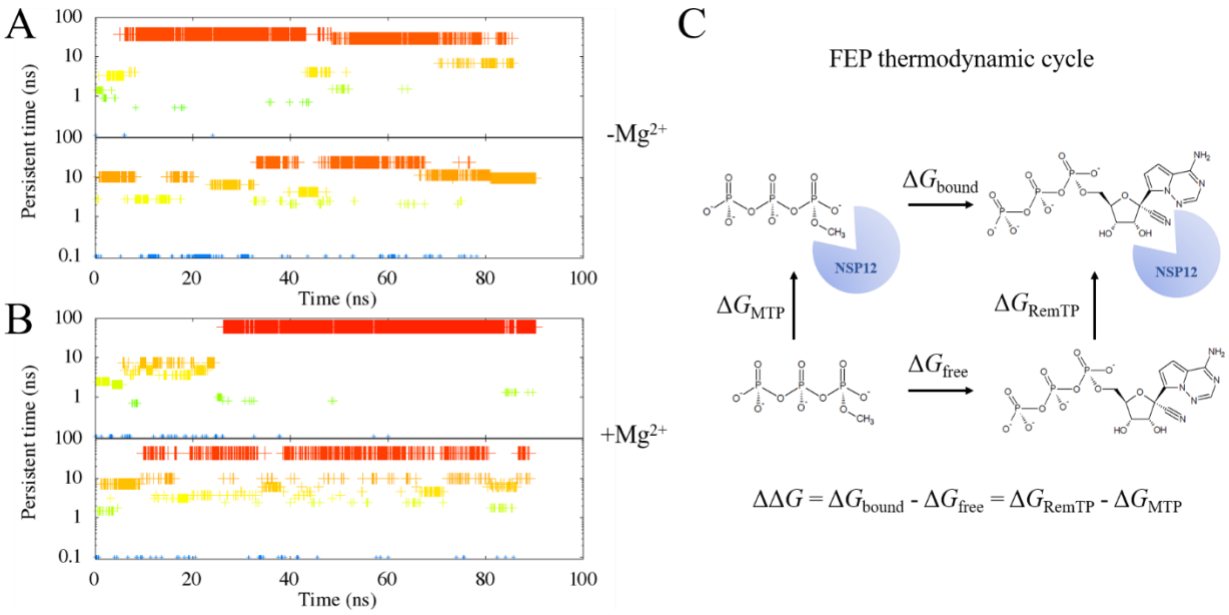

Figure 3. We identify the clusters of the ligand binding state with the RMSD calculations on the ligand alone (in this case, RemTP), while COVID-19 NSP12 is aligned. A cutoff of 2 Å is chosen. The clustering results from the two simulations with COVID-19 NSP12-RemTP complex without Mg^{2+} are plotted in (A). The clustering results from the two simulations with COVID-19 NSP12-RemTP complex and Mg^{2+} are plotted in (B). The clusters are plotted using their "persistence time", defined as the simulation time during which the cluster persists. Larger clusters are plotted with big crosses and red colors. Smaller clusters are plotted with small crosses and blue colors. Upon obtaining the largest clusters of each simulation, we perform free energy perturbation (FEP) calculations using the thermodynamic cycle in (C).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

Likewise, we performed the MD simulations for the putative binding structures of RemTP in COVID-19 NSP12. Similar to that of ATP, the binding cluster of RemTP with Mg^{2+} is significantly larger than that without Mg^{2+} (Figure 3A vs 3B). Similarly, the "bound state" was therefore chosen as the largest cluster of the first set of simulations with Mg^{2+} (Figure 3B, upper panel). A closer inspection of this structure showed that it deviated from the ATP binding site somewhat (Figure 5A). The triphosphate part of the molecule mainly interacted with positively charged residues K551, R553, R555, K621, K798, R836 along with Mg^{2+} and S549. The remdesivir part of the molecule bound with S549 and R555 as well, interacting with additional residues like K545, A547 and V557. Interestingly, D618, a previously identified residue that was crucial for the SARS-CoV RdRp activity[22] was found to be directly affected by the binding of RemTP (by RemTP forming a hydrogen bond with K798, which could have originally formed a hydrogen bond with D618). Another interesting finding was that the configuration of RemTP in the binding grip of COVID-19 NSP12 seemed to "block" the grip (Figure 5B). Therefore, it might poise to potentially slow down the efficiency of RdRp, on top of its supposed function that RemTP could act as a terminating nucleotide on the RNA[15].

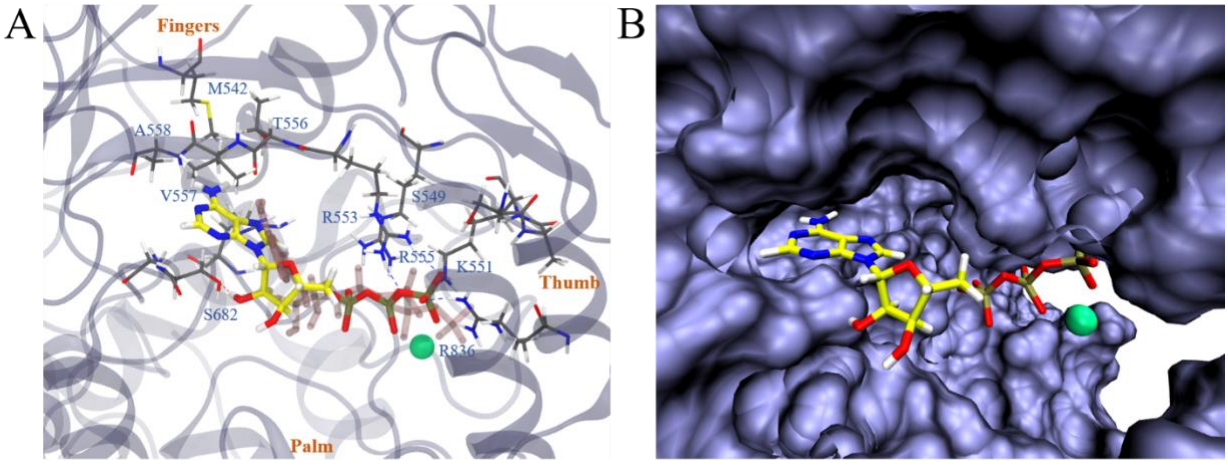

Figure 4. In (A), we illustrate the most stable binding mode of COVID-19 NSP12-ATP (with Mg^{2+}) based on the clustering result (Figure S1B, upper panel). All representations are the same as Figure 2 except that Mg^{2+} is drawn with a green sphere, and that the initial structure of ATP is drawn in transparent pink. In (B), the representation of COVID-19 NSP12 is switched to the solvent accessible surface, highlighting the local binding pocket.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

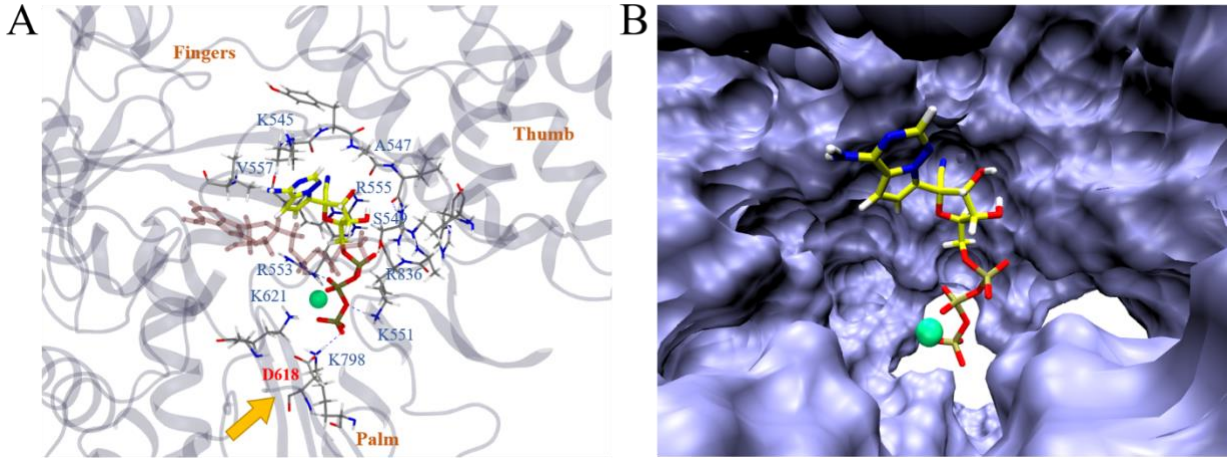

Figure 5. In (A), we illustrate the most stable binding mode of COVID-19 NSP12-RemTP (with Mg^{2+}) based on the clustering result (Figure 3B, upper panel). All representations are the same as Figure 2 except that Mg^{2+} is drawn with a green sphere, the most stable binding structure of *ATP* (from Figure 4A) is drawn in transparent pink, and that D618 is emphasized in red with an orange arrow. In (B), the representation of COVID-19 NSP12 is switched to the solvent accessible surface, highlighting the local binding pocket.

RemTP is ~800 times more potent than ATP

We followed the thermodynamic cycle illustrated in Figure S1C and Figure 3C to calculate the relative binding free energy of ATP and RemTP (with the 0-point energy reference as the MTP molecule: methanol triphosphate). The binding free energy of the most stable binding structures was then calculated and averaged over 5 runs of FEP calculations. The relative binding free energy ($\Delta\Delta G$) of ATP was found to be -4.14 ± 0.89 kcal/mol, while that of RemTP was -8.28 ± 0.65 kcal/mol. The difference between the two was ~4.14 ± 1.10 kcal/mol, equating to ~834 times of difference in K_d values under 310K. It should be noted that FEP can converge slowly particularly on systems where the environments of the target alchemical modification undergoes slow response fluctuations; for this purpose, various advanced free energy computing strategies were developed, such as Funnel Metadynamics[32], and Orthogonal Space Random Walk (OSRW)[33]. In this study, we adopted a simple approach following our previous studies[34-37] [cites] with multiple runs starting from the most probably binding cluster. It would be also desirable in the future to study the ligand-RdRp binding process with more advanced sampling techniques, such as replica

 doi:10.20944/preprints202003.0267.v1

exchange[38], replica exchange with solute-tempering [39] and metadynamics [40] to help reduce the enormous computational resources required (it takes hours if not days in experiment to measure Kd values typically), and better understand the underlying mechanism of the therapeutic effects of remdesivir.

Nevertheless, our calculations strongly suggest that RemTP would almost completely replace the "native ligand" ATP when both presenting at the binding grip of COVID-19 NSP12. This finding supplements the previous study where RemTP was found to bind with the Ebola RdRp[41]. It also suggests that if administrated correctly and metabolized correctly, remdesivir might indeed be a promising drug to significantly lower the reproduction of SARS-CoV-2.

Conclusion

In this study, we constructed the homology model of SARS-CoV-2 NSP12 RdRp with high sequence identity (95.8%). None of the key residues at the RNA/NTP binding site were mutated from SARS-CoV to SARS-CoV-2. Although there has been no co-crystal structure yet for ATP bound with either SARS-CoV or SARS-CoV-2 RdRps, we successfully constructed a model for ATP/remdesivir binding with SARS-CoV-2 NSP12 based on a previous co-crystal structure of poliovirus RdRp. The relative binding free energy of ATP (w.r.t. MTP) was subsequently calculated to be -4.14 ± 0.89 kcal/mol with the presence of Mg^{2+}. The active metabolite of remdesivir (RemTP) was found to have a relative binding free energy (w.r.t. MTP) of -8.28 ± 0.65 kcal/mol, which is significantly stronger than ATP. The ~800-fold difference in the K_d value might decisively block ATP out of the binding pocket when RemTP is in the vicinity. Subsequently, RemTP could act as an effective SARS-CoV-2 RNA-chain terminator, stopping its RNA reproduction. Additionally, the previously identified crucial residue D618 was affected by the binding of RemTP, indicating a possible secondary effect of this drug that would slow down the activity of RdRp, thus also helping cure the COVID-19 disease.

Author Contributions

R.Z. and L.Z. conceived and designed the study. L.Z. performed molecular dynamics simulations. L.Z. collected and analyzed data. L.Z. and R.Z. interpreted results and co-wrote the manuscript. All authors contributed to the general discussion of the project and manuscript.

Acknowledgement

The authors wish to acknowledge helpful discussions with Michael Levitt. R.Z. gratefully acknowledges the financial support from the IBM Bluegene Science Program (W125859, W1464125 and W1464164).

References

1. Cohen, J.; Kupferschmidt, K., Strategies shift as coronavirus pandemic looms. American Association for the Advancement of Science: 2020.
2. Gorbalenya, A. E., Severe acute respiratory syndrome-related coronavirus–the species and its viruses, a statement of the coronavirus study group. *BioRxiv* **2020**.
3. Smith, R. D., Responding to global infectious disease outbreaks: Lessons from sars on the role of risk perception, communication and management. *Social science & medicine* **2006,** *63* (12), 3113-3123.
4. Baharoon, S.; Memish, Z. A., Mers-cov as an emerging respiratory illness: A review of prevention methods. *Travel medicine and infectious disease* **2019**, 101520.
5. Forni, D.; Cagliani, R.; Clerici, M.; Sironi, M., Molecular evolution of human coronavirus genomes. *Trends in microbiology* **2017,** *25* (1), 35-48.
6. Lim, W. S., Influenza, pandemics and sars. *ERS Handbook of Respiratory Medicine* **2019**, 393.
7. Liu, Y.; Gayle, A. A.; Wilder-Smith, A.; Rocklöv, J., The reproductive number of covid-19 is higher compared to sars coronavirus. *Journal of travel medicine* **2020**.
8. Xu, K.; Cai, H.; Shen, Y.; Ni, Q.; Chen, Y., et al., Management of corona virus disease-19 (covid-19): The zhejiang experience. *Zhejiang da xue xue bao. Yi xue ban= Journal of Zhejiang University. Medical sciences* **2020,** *49* (1), 0.
9. Yanping, Z., The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (covid-19) in china. *Zhonghua liuxingbingxue zazhi* **2020,** *41* (2), 145.
10. Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J., et al., Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-ncov) in vitro. *Cell research* **2020**, 1-3.
11. Lim, J.; Jeon, S.; Shin, H.-Y.; Kim, M. J.; Seong, Y. M., et al., Case of the index patient who caused tertiary transmission of covid-19 infection in korea: The application of lopinavir/ritonavir for the treatment of covid-19 infected pneumonia monitored by quantitative rt-pcr. *Journal of Korean Medical Science* **2020,** *35* (6).
12. Li, G.; De Clercq, E., Therapeutic options for the 2019 novel coronavirus (2019-ncov). Nature Publishing Group: 2020.
13. Beck, B. R.; Shin, B.; Choi, Y.; Park, S.; Kang, K., Predicting commercially available antiviral drugs that may act on the novel coronavirus (2019-ncov), wuhan, china through a drug-target interaction deep learning model. *bioRxiv* **2020**.
14. Chang, Y.-C.; Tung, Y.-A.; Lee, K.-H.; Chen, T.-F.; Hsiao, Y.-C., et al., Potential therapeutic agents for covid-19 based on the analysis of protease and rna polymerase docking. **2020**.
15. Warren, T. K.; Jordan, R.; Lo, M. K.; Ray, A. S.; Mackman, R. L., et al., Therapeutic efficacy of the small molecule gs-5734 against ebola virus in rhesus monkeys. *Nature* **2016,** *531* (7594), 381-385.
16. Dyer, O., Two ebola treatments halve deaths in trial in drc outbreak. *BMJ: British Medical Journal (Online)* **2019,** *366*.
17. Furuta, Y.; Komeno, T.; Nakamura, T., Favipiravir (t-705), a broad spectrum inhibitor of viral rna polymerase. *Proceedings of the Japan Academy, Series B* **2017,** *93* (7), 449-463.
18. Kirchdoerfer, R. N.; Ward, A. B., Structure of the sars-cov nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nature communications* **2019,** *10* (1), 1-9.

19. Sheahan, T. P.; Sims, A. C.; Leist, S. R.; Schäfer, A.; Won, J., et al., Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against mers-cov. *Nature Communications* **2020,** *11* (1), 1-14.
20. Gilead's remdesivir to enter trials for coronavirus treatment. **2020**.
21. Us begins trial of gilead's remdesivir in covid-19 patients. **2020**.
22. Te Velthuis, A. J.; Arnold, J. J.; Cameron, C. E.; van den Worm, S. H.; Snijder, E. J., The rna polymerase activity of sars-coronavirus nsp12 is primer dependent. *Nucleic acids research* **2010,** *38* (1), 203-214.
23. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W., et al., A new coronavirus associated with human respiratory disease in china. *Nature* **2020**, 1-5.
24. Fiser, A.; Šali, A., Modeller: Generation and refinement of homology-based protein structure models. In *Methods in enzymology*, Elsevier: 2003; Vol. 374, pp 461-491.
25. Thompson, A. A.; Albertini, R. A.; Peersen, O. B., Stabilization of poliovirus polymerase by ntp binding and fingers–thumb interactions. *Journal of molecular biology* **2007,** *366* (5), 1459-1474.
26. Sgrignani, J.; Magistrato, A., The structural role of mg2+ ions in a class i rna polymerase ribozyme: A molecular simulation study. *The Journal of Physical Chemistry B* **2012,** *116* (7), 2259-2268.
27. Vanommeslaeghe, K.; MacKerell Jr, A. D., Automation of the charmm general force field (cgenff) i: Bond perception and atom typing. *Journal of chemical information and modeling* **2012,** *52* (12), 3144-3154.
28. Huang, J.; MacKerell Jr, A. D., Charmm36 all‐atom additive protein force field: Validation based on comparison to nmr data. *Journal of computational chemistry* **2013,** *34* (25), 2135-2145.
29. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C., et al., Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015,** *1*, 19-25.
30. Zhuang, S.; Zhang, L.; Zhan, T.; Lu, L.; Zhao, L., et al., Binding specificity determines the cytochrome p450 3a4 mediated enantioselective metabolism of metconazole. *The Journal of Physical Chemistry B* **2018,** *122* (3), 1176-1184.
31. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E., et al., Scalable molecular dynamics with namd. *Journal of computational chemistry* **2005,** *26* (16), 1781-1802.
32. Limongelli, V.; Bonomi, M.; Parrinello, M., Funnel metadynamics as accurate binding free-energy method. *Proceedings of the National Academy of Sciences* **2013,** *110* (16), 6358-6363.
33. Zheng, L.; Chen, M.; Yang, W., Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105* (51), 20227.
34. Ahmed, R.; Omidian, Z.; Giwa, A.; Cornwell, B.; Majety, N., et al., A public bcr present in a unique dual-receptor-expressing lymphocyte from type 1 diabetes patients encodes a potent t cell autoantigen. *Cell* **2019,** *177* (6), 1583-1599. e16.
35. Chowell, D.; Morris, L. G.; Grigg, C. M.; Weber, J. K.; Samstein, R. M., et al., Patient hla class i genotype influences cancer response to checkpoint blockade immunotherapy. *Science* **2018,** *359* (6375), 582-587.
36. Das, P.; Li, J.; Royyuru, A. K.; Zhou, R., Free energy simulations reveal a double mutant avian h5n1 virus hemagglutinin with altered receptor binding specificity. *Journal of computational chemistry* **2009,** *30* (11), 1654-1663.
37. Zhou, R.; Das, P.; Royyuru, A. K., Single mutation induced h3n2 hemagglutinin antibody neutralization: A free energy perturbation study. *The Journal of Physical Chemistry B* **2008,** *112* (49), 15813-15820.
38. Sugita, Y.; Okamoto, Y., Replica-exchange molecular dynamics method for protein folding *Chem. Phys. Lett.* **1999,** *314* (1-2), 141 - 151.
39. Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J., Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102* (39), 13749.
40. Laio, A.; Parrinello, M., Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99* (20), 12562.

 doi:10.20944/preprints202003.0267.v1

41. 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.

 doi:10.20944/preprints202003.0267.v1

Supporting information for

Binding mechanism of remdesivir to SARS-CoV-2 RNA dependent RNA polymerase

Leili Zhang[1] and Ruhong Zhou[1,2]*

[1]*Computational Biology Center, IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA*

[2]*Department of Chemistry, Columbia University, New York, NY 10027, USA*

*All Correspondence should be addressed to: rz24@columbia.edu

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

Table S1. Autodock vina scores (vina score) of adenosine analogues from DrugBank. Note that the column of "#P" indicates the number of phosphorus in the molecule; "Charge" lists the net charge of the molecule; and "MW" lists the molecular weight of the molecule.

Drug Bank ID	Name	vina score	#P	Charge	MW
DB06213	Regadenoson	-7.6	0	0	404
DB06441	Cangrelor (no 3D)	-7.3	3	-4	790
NA	ATP	-7.2	3	-4	521
NA	Remdesivir-TP	-7.2	3	-4	549
DB00140	Riboflavin	-7.1	0	0	390
DB00118	Ademetionine	-6.9	0	-1	409
DB01610	Valganciclovir	-6.7	0	0	364
NA	Remdesivir-MP	-6.7	1	-2	385
DB01073	Fludarabine	-6.5	0	0	293
DB00442	Entecavir	-6.4	0	0	287
DB01280	Nelarabine	-6.4	0	0	305
NA	Remdesivir-noP	-6.4	0	0	303
DB00900	Didanosine	-6.3	0	0	244
DB01048	Abacavir	-6.3	0	0	296
DB00242	Cladribine	-6.2	0	0	294
DB00552	Pentostatin	-6.1	0	0	274
DB00631	Clofarabine	-6.1	0	0	312
DB00640	Adenosine	-6.1	0	0	275
DB01004	Ganciclovir	-6.1	0	0	263
DB00194	Vidarabine	-6.0	0	0	275
DB14126	Tenofovir	-6.0	1	-2	297
DB00426	Famciclovir	-5.8	0	0	333
DB13868	Adefovir (NA)	-5.8	1	-2	283
control	ethanol	-2.7	0	0	46

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

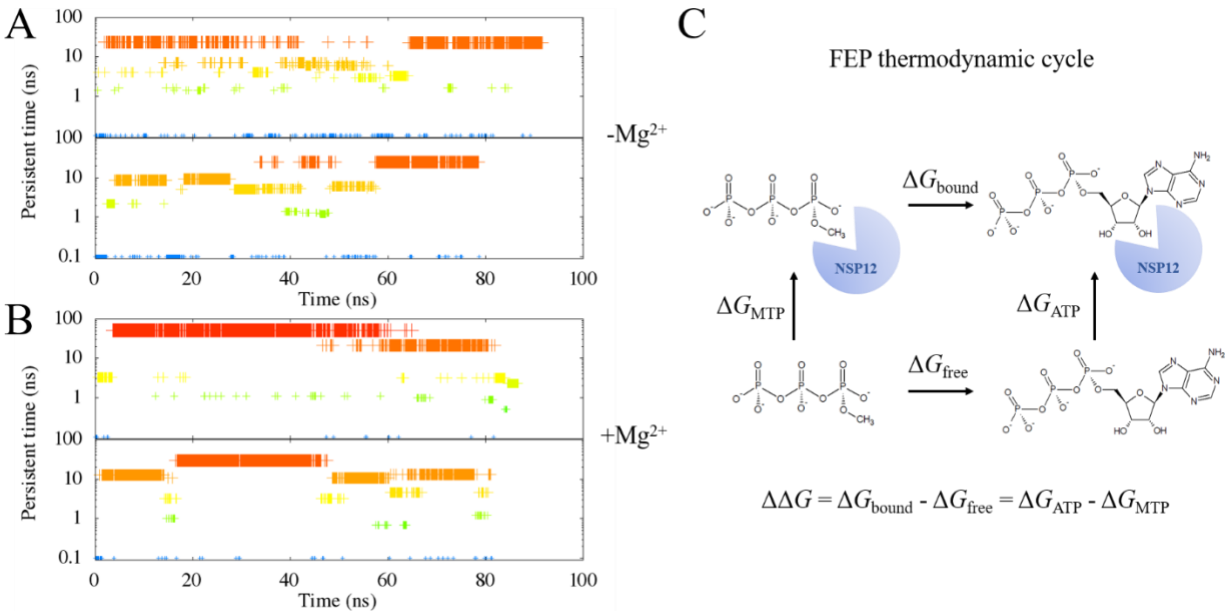

Figure S1. We identify the clusters of the ligand binding state with the RMSD calculations on the ligand alone (in this case, ATP), while COVID-19 NSP12 is aligned. A cutoff of 2 Å is chosen. The clustering results from the two simulations with COVID-19 NSP12-ATP complex without Mg^{2+} are plotted in (A). The clustering results from the two simulations with COVID-19 NSP12-ATP complex and Mg^{2+} are plotted in (B). The clusters are plotted using their "persistence time", defined as the simulation time during which the cluster persists. Larger clusters are plotted with big crosses and red colors. Smaller clusters are plotted with small crosses and blue colors. Upon obtaining the largest clusters of each simulation, we perform free energy perturbation (FEP) calculations using the thermodynamic cycle in (C).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2020 doi:10.20944/preprints202003.0267.v1

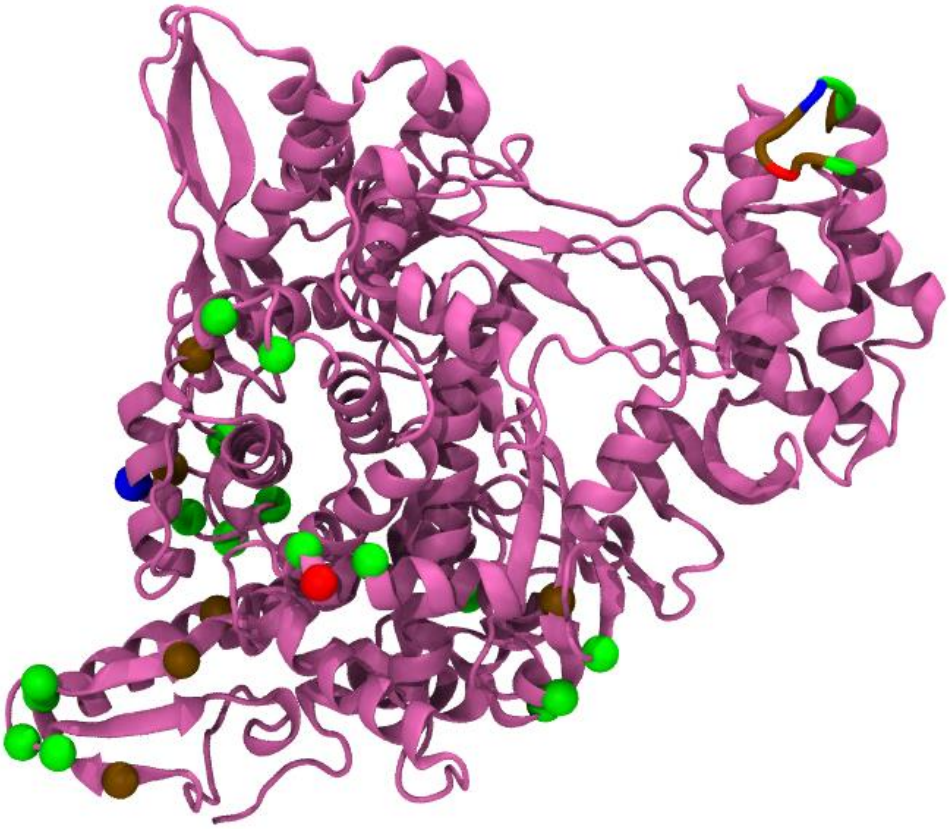

Figure S2. The homology model of COVID-19 NSP12 (pink). The mutated residues from 6NUR are highlighted in VDW balls (colored by brown for hydrophobic residues, green for hydrophilic residues, red for positively charged residues and blue for negatively charged residues). The added random coil is highlighted with a loop representation (colored the same way as the VDW balls).