Targeting RdRp of SARS-CoV-2 with de novo molecule generation

Amal Vijay, Venkata Sai Sreyas Adury, and Arnab Mukherjee*
Department of Chemistry, Indian Institute of Science Education and Research, Pune-411008, India.

ABSTRACT: Viruses are known for their extremely high mutation rates, allowing them to evade both the human immune system and many forms of standard medicine. Despite this, the RNA Dependent RNA Polymerase (RdRp) of the RNA viruses has been largely conserved, and any significant mutation to this protein is unlikely. The recent COVID-19 pandemic presents the need for novel therapeutics. We have designed a de novo drug design algorithm that generates strong binding ligands from scratch, based only on the structure of the target protein's receptor. In this article, we applied our method to target SARS-CoV-2 RdRp and generated several de novo molecules. We then chose some drug molecules based on the structural similarity to some of our strongest binding de novo molecules. Subsequently, we showed, using rigorous all-atom explicit-water free energy calculations in near-microsecond timescales using state-of-the-art well-tempered metadynamics simulations, that some of our de novo generated ligands bind more strongly to RdRp than the recent FDA approved drug Remdesivir in its active form, remdesivir triphosphate (RTP). We elucidated the binding mechanism for some of the top binders and compared with RTP. We believe that this work will be useful both by presenting newer lead structures for RdRp inhibition and by delivering key insights into the residues of the protein potentially involved in the binding/unbinding of these small molecule drugs, leading to more targeted studies in the future.

INTRODUCTION

Coronavirus 2019 (COVID-19) is a disease caused by the SARS-CoV-2 virus. The SARS-CoV-2 genome is a single-strand positive-sense RNA[1], and its replication is facilitated through a multi-subunit replication/transcription complex of viral non-structural proteins (NSPs)[2-4]. The key aspect of the NSP complex is an RNA-dependent RNA polymerase (RdRp; NSP 12)[1, 5-6]. The RdRp requires extra factors, such as NSP7 and NSP8, for its activity[7]. The active site of NSP12 is situated in the middle of the substrate domain, where the synthesis of RNA takes place as an RNA template is accessed from the template input channel. A structural representation of NSP7, NSP8, and NSP12 is shown in Fig. 1. The central RdRp domain is divided into three subdomains: the thumb, palm, and right-handed cup-like fingers.[8-9] The relevant finger and thumb regions which constitute the open/closed structure of the RdRp protein are highlighted in Fig. 2. Due to the central role of RdRp in the replication process of viruses and its conserved structure, it is an attractive target for drug development against SARS-CoV-2 infections.

Since the pandemic, a plethora of studies have focused on finding suitable therapeutic remedies for Covid-19, and the search is still ongoing. Among several possible targets, including spike protein, main protease (M^{pro}), etc., RdRp is a promising target for drug design. Some antiviral molecules have already been identified by targeting key residues in the active site of RdRp and other proteins involved in the lifecycle of the virus.[10-13] At the same time, various drug repurposing studies that target RdRp using computational tools like docking have also been reported.[14-19] All these studies resulted into a great number of possible drugs which could be repurposed for the treatment of Covid-19, pending rigorous theoretical or experimental tests.

However, the search for a cure for Covid-19 is far from over. Among the possible drugs for Covid-19 that target RdRp, Remdesivir is the one with FDA approval.[11] Although the effectiveness of Remdesivir is still disputed, there have been many studies on its inhibition effects[20-22]. It was found by experimental and computational studies that the active form of Remdesivir, Remdesivir triphosphate (RTP), is an effective drug that prevents the replication of RNA at the active site of the RdRp.[23-27] Studies indicate that the binding affinity of RTP is in the range of ~5 kcal/mol to ~7 kcal/mol[28-29], and it completely inhibits the polymerization process at the catalytic site.[20, 30]

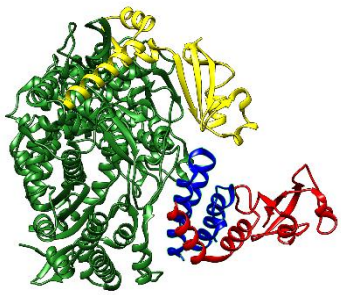


Figure 1. Structure of RdRp protein (PDB ID: 6NUR) with its key components indicated by NSP12 (Green), NSP7 (Yellow), and NSP8 (Blue).

In the present study, we used an atomistic receptor-based de novo generation algorithm[31] followed by structural mapping and free energy calculations to obtain a new set of molecules and repurpose drugs that can possibly bind more effectively than the RTP. Initially, we analyzed the active region of RTP-bound 3D structure of RdRp for identifying the hotspot region, which was used as the target location for generating small molecules. To confirm the binding, we have performed molecular dynamics simulation and free energy calculations to compare the binding strength of our de novo molecules and RTP. From the multi-dimensional free energy surfaces, we found five molecules, some of which are known drug molecules, that can bind more strongly than RTP at the catalytic site of RdRp.

METHODS

Our strategy for finding suitable molecules relies on the following three stages of the protocol.

A. Stage 1: Generation of novel de novo molecules around the active region of RdRp.
B. Stage 2: Similarity mapping of top molecules from stage 1 for possible repurposable drugs.
C. Stage 3: Validation of the binding affinity of top molecules from stage 1 and stage 2 by free energy calculations.

We discuss each step in more detail below.

a) De novo drug generation. We have developed an in-house program[31] that is designed to generate chemically rational molecules by placing atoms in the receptor one by one and connecting them by bonds based on standard force field, while optimizing their interactions to a target macromolecule (usually a protein). It is designed to work in any parametrized chemical environment and explores vast regions of chemical space. Given the structure of the protein in PDB[32] and GRO[33] formats with residue names denoting a particular hotspot, the program picks a random location within the hotspot and starts generation with a seed atom. A second atom is chosen to bind to this seed atom and places it within the bonding sphere of the first. Then a third atom is chosen and is placed within the bonding sphere of the second and within the base of a cone formed by the first three atoms. The fourth atom is placed following the allowed dihedral angle of the first four atoms, and the other rules for bonding and bending angles are mentioned for the first three atoms. Thus, the construction of the entire molecule follows one atom at a time following the restrictions of a given classical force field and applying the approach of similar to the Configurational Bias Monte Carlo (CBMC)[34] until a molecule of a given target size is completed. Since we follow a particular classical force field (here, the program has been fine-tuned to work with the CHARMM-27[35]), we use atom-type instead of the atom, which provides both bonding and nonbonding interactions within itself and with the protein. Based on the force field parameters our program can calculate the nonbonded interaction energies between any two (or more) molecules. This interaction energy is used for optimizing the molecules generated within a given region, i.e., hotspot. This is rather simplistic for a scoring function. However, for biasing a rapidly growing molecule, it is a good starting point. A quick local optimization is performed using gradient descent[36] to get a more accurate score for the molecule. There is a scope to add in more filters at this stage to make the final screening more accurate, which the algorithm allows as an external input. A brief intuition of this algorithm can be gleaned from the flowchart shown in the supplementary information (SI) Fig. S1.

We applied our program to target RdRp, the structure of which is obtained from the Protein Data Bank (PDB ID: 6NUR). We used Modeller 9.21[37] to model missing residues in the protein. The active site of RdRp (which contains the following residues Gly616, Trp617, Asp618, Tyr619, Leu758, Ser759, Asp760, Asp761, Ala762, Lys798, Tys799, Trp800, Glu811, Phe812, Cis813, and Ser814) was used for the generation of small molecular binders.

b) Similarity matching. Because the molecules generated by our program are built from scratch, they may be different from the standard available molecules. Sometimes it could also be away from known synthetic routes, and devising synthetic methods for these new molecules might be expensive and time-consuming. Most importantly, our de novo molecules are not screened for ADMET (adsorption, distribution, metabolism, excretion, and toxicity) criteria required to be a drug.[38 39] Therefore, we augmented our list of strong binders (based on the interaction energy mentioned before) with known drug molecules with similar chemical structures, with the assumption that molecules with similar structures would exhibit similar functions,[40,41] as shown in our previous study[31]. Thus, we mapped our molecules to approved or investigational (repurpose) drugs from the DrugBank[42] database with a Tanimoto similarity score of 0.4 or more, which has been shown to be significant[43]. We screened the drug molecules selected based on structural similarity to our high scoring de novo molecules further for docking scores using AutoDock Vina version 1.1.2[44-45] to ensure the preliminary idea of their binding strength. A box centered near the mid-point of the Asn691 and Cys622 side chains was created with dimensions 32.6 Å×33.3 Å×36.5 Å. The docking score of the molecules is provided in Table S1 of the Supporting Information (SI). The chemical structures of both de novo and repurpose drugs are shown in Fig. S2.

c) Free energy validation. Finally, we calculated the binding free energy for each of these molecules (selected based on strong interaction energy for de novo molecules and high docking score for structurally similar drug molecules) using all-atom, explicit water metadynamics simulations as described later in this article.

i. Forcefield generation: Following the guideline of General Amber Force Field (GAFF)[46], we optimized all the molecules using Hartree-Fock (HF) theory with 6-31G* basis set using Gaussian 09 software[47]. We then used the antechamber[48] module of AMBER18[49] for the RESP charge calculation of drug atoms. For protein, we used AMBER99SB force field.[50] The topology and coordinates were then converted into the GROMACS format by using a python script acpype.py (available at https://github.com/t-/acpype)[51].

ii. System setup for simulation of all molecules: All the simulations were performed using the GROMACS 2019.6[52-53] molecular dynamics software patched with plumed 2.6[54-55]. Each protein-ligand complex system was solvated by ~76000 TIP3P water molecules[56] in a box with dimensions 16 X 14 X 11 nm^3. The physiological concentration (150 mM) of Na^+ and Cl^- ions along with extra Cl^- ions, were used to neutralize the system.

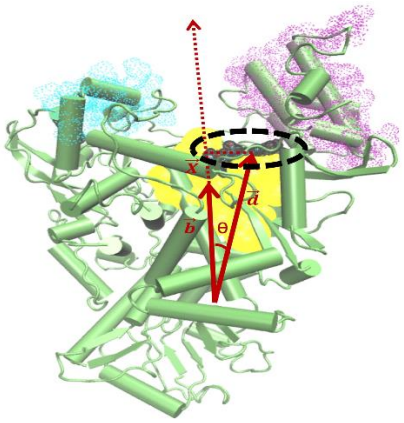


Figure 2. Structure of NSP12 protein. The active site of the protein is indicated by the yellow region. A part of the finger sub-region (residue 500 to 525) and thumb sub–region (residue 850 to 900)

constituting the ring opening/closed junction is indicated by cyan and purple, respectively. A sample small molecule is shown in black colour covered by black dotted circle near the active site of the protein. The vectorial representation of collective variables used for metadynamics simulations is shown by red arrows (see methods).

iii. Equilibration and Simulation: Initially, each system was energy minimized using steepest descent[36] for 10000 steps, followed by heating to 300K in 200 ps using Berendsen thermostat and barostat[57] with a coupling constant of 0.6 ps. Restraints of 25 kcal/mol/Å^2 were applied on heavy atoms during the heating process. Thereafter, equilibration was carried out for 2 ns at constant temperature (300 K) and pressure (1 bar) without any restraints using the same thermostat and barostat with coupling constants of 0.2 ps for each. The last 100 ps of NPT simulation was used to calculate the average volume, which was used in the final 5 ns unrestrained NVT equilibration using the Nosé-Hoover[58] thermostat with a coupling constant of 0.2 ps. During the simulation, the LINCS[59] was used to constrain all the bonds, and the Particle Mesh Ewald (PME) method[60] was used for electrostatics. The distance cut-offs for the van der Waals (vdW) and electrostatic long-range interaction were both kept at 10 Å. The timestep for all simulations was taken to be 2 fs. The equilibrated ligand-bound protein structures (systems) were initially simulated for 5 ns. For a particular system, if the ligand was found to be bound after 5 ns simulation, we proceeded with the free energy calculation. Otherwise, the ligand was rejected as being too weak.

iv. Free energy calculations. To calculate the binding free energy of the molecules, well-tempered metadynamics[61] simulation was performed after equilibration. We used two collective variables (CVs) distvec (X) and angvec (θ) for the well-tempered metadynamics[61] simulations. $X = \hat{b}.\vec{d}$ and $\theta = \cos^{-1}(\hat{b}.\vec{d}/|d|)$, where $\hat{b}$ is the unit vector from the Centre of Mass (COM) of residues 127, 130, 131, 134, 139, 140 – 147, 157, 162, 175, and 787 to the COM of residues 545, 553, 555, 556, 617-619, 622, 623, 680, 682, 687, 691, 759-761, 811, 813, and 814 that lie more toward the hotspot region (yellow region in the figure 2). $\vec{d}$ is the vector from the COM of 127, 130, 131, 134, 139, 140 – 147, 157, 162, 175, and 787 to the COM of the drug molecule. A hill height of 0.2 kJ/mol and a bias factor of 10 was chosen with a hills deposition rate of 2 ps for the metadynamics simulations. Gaussian widths for X and θ were taken to be 0.6 Å and 0.002 rad, respectively. These CVs have been previously used in DNA intercalation and protein-small molecule binding studies that produced free energy values in good agreement with experiments.[31, 62-63]

An intuitive representation of the CVs is shown in Fig. 2. An increase in X ensures that the molecule moves away from the equilibrated bound structure towards the solution medium. The θ is designed to capture the directionality of the unbinding of the molecule to complement X. The minimum free energy path from the ligand-bound minimum to the unbound state in the free-energy landscape was obtained using the program MULE[64]. The free energy convergence for unbinding of the drug in our enhanced sampling simulation takes anywhere between 50 and 200 ns. The metadynamics simulation length for each of the molecules is shown in Table S2. The 2D representation of small molecules interacting with the protein surface is constructed using ligplot[65].

RESULTS & DISCUSSION

a) Screening of DeNovo molecule. Our de novo program generated a total of 5375 molecules, with a target of picking at least 160 molecules of the number of heavy atoms ranging between 11 and 50. Molecules with more than 43 heavy atoms, however, could not be generated owing to the limited size of the active site. The strongest interaction energy (de novo score) was at -150.6 kJ/mol, and the weakest was at -22.1 kJ/mol. The distribution of interaction energies to the active site of RdRp is shown in Fig. S3 of the Supplementary Information (SI). We took the top molecules for the calculation of free energy using well-tempered metadynamics.

b) Free energy studies of RTP. The multi-dimensional free energy surface of RTP is shown in the Fig. 3a. The minimum free energy path connecting the bound minima state to the unbound state is shown by black dotted lines in the Fig. 3a. The unbinding proceeds with an increase in the CVs X and θ. A 2D representation of the RTP molecule with the possible stabilizing interaction in the minima with the residues of protein is shown in Fig. 3b. The top view of the RTP bound RdRp at its minima state is shown in the Fig. 3c. Structures of relevant intermediate states of RdRp-RTP complex along the minimum free energy unbinding path are shown in the Fig. 3d in accordance with the free energy profile RTP (States I to VI in Fig. 3a).

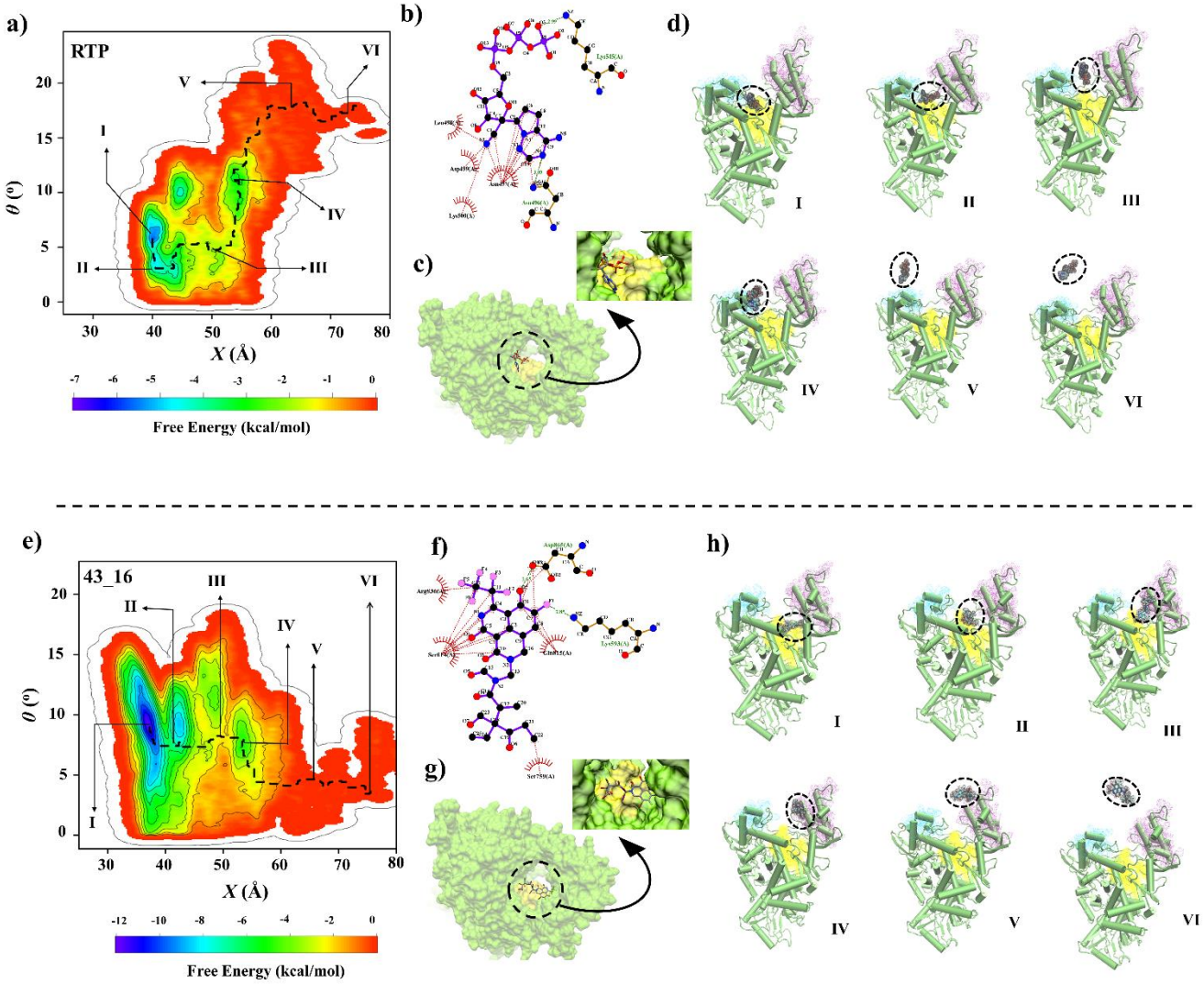


Figure 3. Comparison of free energy surfaces, key interactions and mechanism of RTP (top) and de novo molecule 43_16 (bottom). Free energy surface of a) RTP and e) 43_16. The minimum free energy path is shown by black dotted lines. Various intermediate states associated with the unbinding pathway are shown in the inset of the free energy surface (I-VI) for both RTP and 43_16. Representation of minima states of b) RTP and f) 43_16 with key residues of interactions with the RdRp protein. The top view of RdRp protein bound to c) RTP and g) 43_16 at their corresponding minimum free energy states. Structural representation of intermediate states (I-VI) in the unbinding pathway of d) RTP and h) 43_16, in accordance with the minimum free energy path shown in the free energy surface.

The free energy of RTP was found to be -7.1 kcal/mol, closely matching the reported results.[30] In the minima, the triphosphate of the RTP molecule is found to be stabilized by the hydrogen bond with Lys545 residue of the RdRp, while the other terminal end is found to be stabilized by the hydrogen bond by Asn496 (Fig. 3b). The core of RTP is stabilized by the hydrophobic interactions from the residues Asn497, Leu498, Asp499, and Lys500. The minimum free energy of the RTP is found to be located at the known hotspot region (yellow region shown in Fig. 2). RdRp core structure also consists of architecture containing finger and palm sub-domains which are essential in substrate recognition. After the formation of possible interactions in the hotspot region (state I to III), the RTP moves towards the finger sub-region (shown by cyan colour) as shown by the states IV and V in Fig. 3d. Finally, the departure of the molecule to the accessible solvent region is initiated from this finger subregion.

c) Free energy studies of de novo molecules and known repurposable molecules. As the objective of our study was to find the molecules having better binding affinity than RTP, we followed the same free energy validation protocol for selected de novo and repurpose drugs as discussed in the methods section.

We have found at least five molecules having stronger binding affinity than RTP (free energy < -7.1 kcal/mol). The strongest binding molecule, 43_16(named as per an internal nomenclature of our program), is found to have a free energy value of -11.8 kcal/mol. The free energy surface of 43_16 with minimum

free energy path to unbound state (black dotted lines) and relevant intermediate states (states I to VI) is shown in the Fig. 3e. The major contribution to the free-energetic stability of 43_16 is from hydrophobic stabilizing interactions (from the protein residues Arg836, Ser814, Ser759, and Gln815) and hydrogen bonding interactions (from protein residues Asp865 and Lys593) as shown in Fig 3f. Unlike RTP, the unbinding of 43_16 is facilitated by a decrease in θ with a concomitant increase in X, pointing to an alternate unbinding pathway for this molecule. The top view of the 43_16 bound RdRp is shown in Fig. 3g. The unbinding pathway with intermediate states (states I to VI) is shown in the Fig. 3h. Finally, the detachment of the molecule from the protein surface is initiated from this thumb region as shown in the state VI in Fig. 3h. Here, starting from the minima (state I in Fig. 3h), the unbinding is facilitated through the thumb sub-region of the protein (shown by violet colour) as shown in the states II to V in Fig 3h. Note that in the case of RTP, the unbinding is facilitated through finger sub-region.

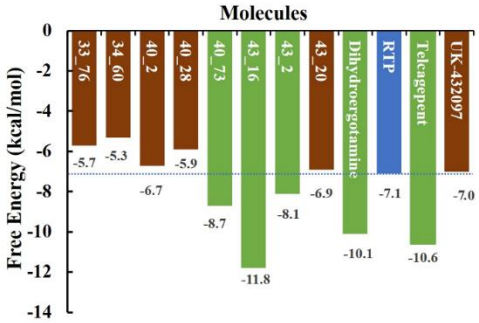


Figure 4. Bar diagram showing the free energetic stability of the ligands as obtained by all-atom, explicit water, well-tempered metadynamics simulations. The bars are coloured and annotated. The green bars indicate ligand with stronger binding affinity than the reference RTP (blue bar), while the brown bars represent ligands with weaker binding. Note that, two repurpose drugs show stronger binding than RTP.

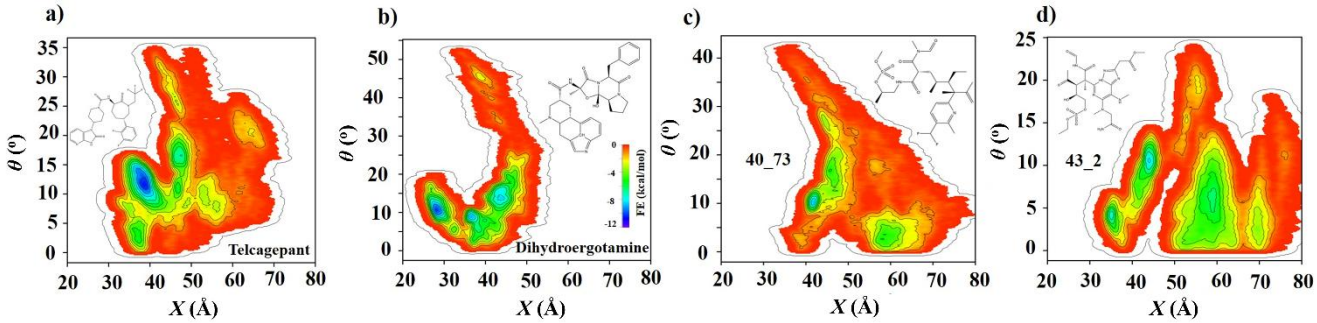


Figure 5. Free energy surface of molecules (de novo & repurpose drugs) with free energy lower than -8 kcal/mol. The chemical structure of the molecule is provided in the inset.

A bar plot showing the free energy of binding of all validated molecules by our simulation protocol is shown in Fig. 4. Other molecules with better binding affinity include Telcagepant and Dihydroergotamine (present-day drugs), and 40_73 and 43_2 (generated de-novo) with free energy values -10.6 kcal/mol, -10.1 kcal/mol, -8.7 kcal/mol, and -8.1 kcal/mol in that order. The free energy surfaces of these molecules are represented in the Fig. 5, along with the chemical structure of the molecules. The free energy surface of all others with lesser binding affinity than RTP is shown in Fig. S4 of the Supporting Information.

CONCLUSION

By targeting the catalytic site of RdRp protein using a novel strategy of generating molecules and validation by rigorous free energy calculations using well-tempered metadynamics, we were able to get at least five molecules that can bind more strongly than RTP, indicating the success of our approach. The free energy landscape using the X and θ collective variables was able to provide key insights into the directionality of movement for each candidate in the unbinding process. It is found that the unbinding of the molecules happens through the small opening at the finger and thumb subregion of the RdRp protein. We found some repurpose drugs that show, in our free energy calculation, stronger binding than RTP.

We show that starting with only the receptor, our approach can propose novel molecules and also repurpose drugs for binding to a receptor. Therefore, the present approach is a unique way to explore chemical space for drug design, while also capturing the known and readily useable repurpose drugs. The computational validation using all-atom, explicit water is expensive; however, reliable in terms of binding free energy prediction. The success rate from de novo score to the free energy stability could encourage one to try the method on other known receptors for lead generation.

We, however, think that there is a huge scope of improvement. We need to consider protein flexibility and the inclusion of water along with better scoring function. Also, to directly use as a drug, we need to screen the binders for potential drug likeness and synthesizability. Work along those directions are under progress.

ASSOCIATED CONTENT

Supporting Information. Text and figures discussing the generation of new molecules, structure of the computer generated and repurposable molecules, docking scores.

AUTHOR INFORMATION

Corresponding Author

*E-mail: arnab.mukherjee@iiserpune.ac.in Phone: +91 20 2590 8051

Author Contributions

The manuscript was written through contributions of all authors.

Funding Sources

Department of Biotechnology (DBT), India (BT/PR34215/AI/133/22/2019) is acknowledged for funding. AV (fellow id: 09/936(0209)/2019-EMR-I) thank CSIR for his fellowship.

ACKNOWLEDGMENT

Authors acknowledge the support and the resources provided by 'PARAM Brahma Facility' under the National Supercomputing Mission, Government of India at the Indian Institute of Science Education and Research, Pune. Authors thank Rituparno Chowdhury for discussion.

ABBREVIATIONS

FE, Free energy; RTP, Remdesivir triphosphate; COM, Center of mass; CV, Collective variable; RdRp, RNA Dependent RNA Polymerase

REFERENCES

1. McDonald, S. M., RNA synthetic mechanisms employed by diverse families of RNA viruses. *WIREs RNA* **2013,** *4* (4), 351-367.
2. Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; Bi, Y.; Ma, X.; Zhan, F.; Wang, L.; Hu, T.; Zhou, H.; Hu, Z.; Zhou, W.; Zhao, L.; Chen, J.; Meng, Y.; Wang, J.; Lin, Y.; Yuan, J.; Xie, Z.; Ma, J.; Liu, W. J.; Wang, D.; Xu, W.; Holmes, E. C.; Gao, G. F.; Wu, G.; Chen, W.; Shi, W.; Tan, W., Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020,** *395* (10224), 565-574.
3. Ivanov, K. A.; Ziebuhr, J., Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *J Virol* **2004,** *78* (14), 7833-8.
4. Ivanov, K. A.; Thiel, V.; Dobbe, J. C.; van der Meer, Y.; Snijder, E. J.; Ziebuhr, J., Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J Virol* **2004,** *78* (11), 5619-32.
5. 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.
6. 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), E3900-9.
7. 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), 2342.
8. Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; Ge, J.; Zheng, L.; Zhang, Y.; Wang, H.; Zhu, Y.; Zhu, C.; Hu, T.; Hua, T.; Zhang, B.; Yang, X.; Li, J.; Yang, H.; Liu, Z.; Xu, W.; Guddat, L. W.; Wang, Q.; Lou, Z.; Rao, Z., Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **2020,** *368* (6492), 779-782.
9. Vicenti, I.; Zazzi, M.; Saladini, F., SARS-CoV-2 RNA-dependent RNA polymerase as a therapeutic target for COVID-19. *Expert Opin Ther Pat* **2021,** *31* (4), 325-337.
10. Elfiky, A. A., Ribavirin, Remdesivir, Sofosbuvir, Galidesivir, and Tenofovir against SARS-CoV-2 RNA dependent RNA polymerase (RdRp): A molecular docking study. *Life Sciences* **2020,** *253*, 117592.
11. 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. *Journal of Biological Chemistry* **2020,** *295* (15), 4773-4779.
12. Hashemian, S. M. R.; Pourhanifeh, M. H.; Hamblin, M. R.; Shahrzad, M. K.; Mirzaei, H., RdRp inhibitors and COVID-19: Is molnupiravir a good option? *Biomed Pharmacother* **2022,** *146*, 112517.
13. Brunt, D.; Lakernick, P. M.; Wu, C., Discovering new potential inhibitors to SARS-CoV-2 RNA dependent RNA polymerase (RdRp) using high throughput virtual screening and molecular dynamics simulations. *Sci Rep* **2022,** *12* (1), 19986.
14. Ahmad, J.; Ikram, S.; Ahmad, F.; Rehman, I. U.; Mushtaq, M., SARS-CoV-2 RNA Dependent RNA polymerase (RdRp) - A drug repurposing study. *Heliyon* **2020,** *6* (7), e04502.
15. Elfiky, A. A.; Mahran, H. A.; Ibrahim, I. M.; Ibrahim, M. N.; Elshemey, W. M., Molecular dynamics simulations and MM-GBSA reveal novel guanosine derivatives against SARS-CoV-2 RNA dependent RNA polymerase. *RSC Adv* **2022,** *12* (5), 2741-2750.
16. Behera, S. K.; Vhora, N.; Contractor, D.; Shard, A.; Kumar, D.; Kalia, K.; Jain, A., Computational drug repurposing study elucidating simultaneous inhibition of entry and replication of novel corona virus by Grazoprevir. *Sci Rep-Uk* **2021,** *11* (1).
17. Alam, A.; Agrawal, G. P.; Khan, S.; Khalilullah, H.; Saifullah, M. K.; Arshad, M. F., Towards the discovery of potential RdRp inhibitors for the treatment of COVID-19: structure guided virtual screening, computational ADME and molecular dynamics study. *Struct Chem* **2022,** *33* (5), 1569-1583.
18. Celik, I.; Tallei, T. E., A computational comparative analysis of the binding mechanism of molnupiravir's active metabolite to RNA-dependent RNA polymerase of wild-type and Delta subvariant AY.4 of SARS-CoV-2. *J Cell Biochem* **2022,** *123* (4), 807-818.
19. Koulgi, S.; Jani, V.; Uppuladinne, V. N. M.; Sonavane, U.; Joshi, R., Natural plant products as potential inhibitors of RNA dependent RNA polymerase of Severe Acute Respiratory Syndrome Coronavirus-2. *PLoS One* **2021,** *16* (5), e0251801.
20. Yin, W.; Mao, C.; Luan, X.; Shen, D.-D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; Chang, S.; Xie, Y.-C.; Tian, G.; Jiang, H.-W.; Tao, S.-C.; Shen, J.; Jiang, Y.; Jiang, H.; Xu, Y.; Zhang, S.; Zhang, Y.; Xu, H. E., Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020,** *368* (6498), 1499-1504.
21. Frediansyah, A.; Nainu, F.; Dhama, K.; Mudatsir, M.; Harapan, H., Remdesivir and its antiviral activity against COVID-19: A systematic review. *Clin Epidemiol Glob Health* **2021,** *9*, 123-127.
22. Bravo, J. P. K.; Dangerfield, T. L.; Taylor, D. W.; Johnson, K. A., Remdesivir is a delayed translocation inhibitor of SARS-CoV-2 replication. *Mol Cell* **2021,** *81* (7), 1548-1552 e4.
23. Elfiky, A. A., SARS-CoV-2 RNA dependent RNA polymerase (RdRp) targeting: an in silico perspective. *Journal of Biomolecular Structure and Dynamics* **2020**, 1-9.
24. Elfiky, A. A., Anti-HCV, nucleotide inhibitors, repurposing against COVID-19. *Life Sciences* **2020,** *248*, 117477.
25. Yu, R.; Chen, L.; Lan, R.; Shen, R.; Li, P., Computational screening of antagonists against the SARS-CoV-2 (COVID-19) coronavirus by molecular docking. *Int J Antimicrob Agents* **2020,** *56* (2), 106012.
26. Singh, P. K.; Pathania, S.; Rawal, R. K., Exploring RdRp-remdesivir interactions to screen RdRp inhibitors for the management of novel coronavirus 2019-nCoV. *SAR QSAR Environ Res* **2020,** *31* (11), 857-867.
27. Wang, J.; Reiss, K.; Shi, Y.; Lolis, E.; Lisi, G. P.; Batista, V. S., Mechanism of Inhibition of the Reproduction of SARS-CoV-2 and Ebola Viruses by Remdesivir. *Biochemistry* **2021,** *60* (24), 1869-1875.
28. Aranda, J.; Wieczor, M.; Terrazas, M.; Brun-Heath, I.; Orozco, M., Mechanism of reaction of RNA-dependent RNA polymerase from SARS-CoV-2. *Chem Catal* **2022,** *2* (5), 1084-1099.
29. Zhang, L.; Zhou, R., Structural Basis of the Potential Binding Mechanism of Remdesivir to SARS-CoV-2 RNA-Dependent RNA Polymerase. *J Phys Chem B* **2020,** *124* (32), 6955-6962.
30. Zhang, L.; Zhou, R., Structural Basis of the Potential Binding Mechanism of Remdesivir to SARS-CoV-2 RNA-Dependent

RNA Polymerase. *The Journal of Physical Chemistry B* **2020,** *124* (32), 6955-6962.

31. Chowdhury, R.; Sai Sreyas Adury, V.; Vijay, A.; Singh, R. K.; Mukherjee, A., Atomistic De-novo Inhibitor Generation-Guided Drug Repurposing for SARS-CoV-2 Spike Protein with Free-Energy Validation by Well-Tempered Metadynamics. *Chem Asian J* **2021,** *16* (12), 1634-1642.

32. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res* **2000,** *28* (1), 235-42.

33. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* **2008,** *4* (3), 435-47.

34. Siepmann, J. I.; Frenkel, D., Configurational bias Monte Carlo: a new sampling scheme for flexible chains. *Molecular Physics* **1992,** *75* (1), 59-70.

35. MacKerell, A. D.; Banavali, N.; Foloppe, N., Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **2000,** *56* (4), 257-265.

36. Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P., *Numerical Recipes in FORTRAN; The Art of Scientific Computing*. Cambridge University Press: 1993.

37. Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M.-y.; Pieper, U.; Sali, A., Comparative Protein Structure Modeling Using Modeller. *Current Protocols in Bioinformatics* **2006,** *15* (1), 5.6.1-5.6.30.

38. Zhong, H. A., ADMET Properties: Overview and Current Topics. In *Drug Design: Principles and Applications*, Grover, A., Ed. Springer Singapore: Singapore, 2017; pp 113-133.

39. Daoud, N. E.; Borah, P.; Deb, P. K.; Venugopala, K. N.; Hourani, W.; Alzweiri, M.; Bardaweel, S. K.; Tiwari, V., ADMET Profiling in Drug Discovery and Development: Perspectives of In Silico, In Vitro and Integrated Approaches. *Curr Drug Metab* **2021,** *22* (7), 503-522.

40. A.H, L.; Y.G.S, Concepts and Applications of Molecular Similarity. *Journal of Molecular Structure* **1992,** *269* (3-4), 376-377.

41. Hansch, C.; Fujita, T., p-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *Journal of the American Chemical Society* **2002,** *86* (8), 1616-1626.

42. Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J., DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* **2006,** *34* (Database issue), D668-72.

43. Baldi, P.; Benz, R. W., BLASTing small molecules--statistics and extreme statistics of chemical similarity scores. *Bioinformatics* **2008,** *24* (13), i357-65.

44. Eberhardt, J.; Santos-Martins, D.; Tillack, A. F.; Forli, S., AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J Chem Inf Model* **2021,** *61* (8), 3891-3898.

45. Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **2010,** *31* (2), 455-61.

46. Wang, J.; 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), 1157-74.

47. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 09 Rev. A.01*, Wallingford, CT, 2016.

48. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling* **2006,** *25* (2), 247-260.

49. D.A. Case, I. Y. B.-S., S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden,; R.E. Duke, D. G., M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang,; S. Izadi, A. K., T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J.; Mermelstein, K. M. M., Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R.; Qi, D. R. R., A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman, Amber18. *University of California, San Francisco.* **2018**.

50. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006,** *65* (3), 712-25.

51. Sousa da Silva, A. W.; Vranken, W. F., ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* **2012,** *5* (1), 367.

52. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J., GROMACS: fast, flexible, and free. *J Comput Chem* **2005,** *26* (16), 1701-18.

53. 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,** *1-2*, 19-25.

54. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G., PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **2014,** *185* (2), 604-613.

55. consortium, P., Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods* **2019,** *16* (8), 670-673.

56. 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.

57. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984,** *81* (8), 3684-3690.

58. Nosé, S., A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics* **1984,** *52* (2), 255-268.

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

60. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993,** *98* (12), 10089-10092.

61. Barducci, A.; Bussi, G.; Parrinello, M., Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Physical Review Letters* **2008,** *100* (2), 020603.

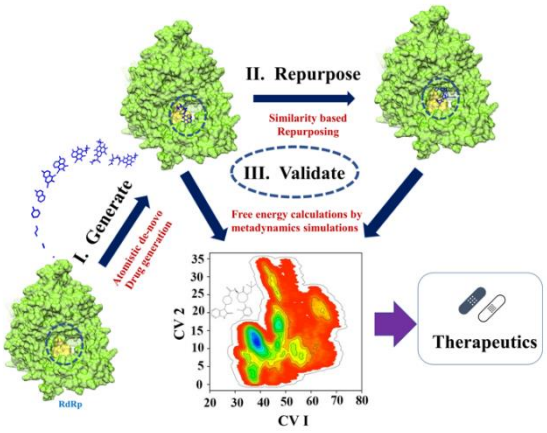
62. Sasikala, W. D.; Mukherjee, A., Structure and dynamics of proflavine association around DNA. *Phys Chem Chem Phys* **2016,** *18* (15), 10383-91.

63. Sasikala, W. D.; Mukherjee, A., Molecular mechanism of direct proflavine-DNA intercalation: evidence for drug-induced minimum base-stacking penalty pathway. *J Phys Chem B* **2012,** *116* (40), 12208-12.

64. Fu, H.; Chen, H.; Wang, X.; Chai, H.; Shao, X.; Cai, W.; Chipot, C., Finding an Optimal Pathway on a Multidimensional Free-Energy Landscape. *J Chem Inf Model* **2020,** *60* (11), 5366-5374.

65. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M., LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **1995,** *8* (2), 127-34.

TOC


II. Repurpose
Similarity based Repurposing
III. Validate
Free energy calculations by metadynamics simulations
I. Generate
Atomistic de-novo Drug generation
RdRp
CV 2
CV 1
Therapeutics