Drug Repurposing against SARS-CoV-2 RDRP - a computational quest against CoVID- 19

Hirak Jyoti Chakrabortya (✉ hj.chakraborty@gmail.com)

Department of Biotechnology, CIFRI, Barrackpore, Kolkata

Prasenjit Paria

Department of Biotechnology, CIFRI, Barrackpore, Kolkata

Aditi Gangopadhyay

Department of Chemical Technology, University of Calcutta

Sayak Ganguli

Department of Biotechnology, St. Xavier's College (Autonomous) Kolkata

Research Article

Keywords: Drug Repurposing, CoVID-19, SARS-CoV-2, PPI network, RDRP, molecular docking, DrugBank

Posted Date: April 9th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-22079/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The present CoVID-19 pandemic was first detected in December 2019 in Wuhan, China, and is rapidly spreading worldwide. To date, it has affected 465,915 individuals in 200 countries, and has been responsible for 21,031 deaths. In the absence of definitive treatment strategies, there is a pressing demand for drug discovery against CoVID-19. Drug repurposing is a cost-effective and time-saving strategy which essentially involves the identification of novel targets for known drug candidates. This reduces the time and cost of drug discovery, as the pharmacokinetics and toxicity profiles of the drugs are already known, which makes phase-I clinical trials redundant. Here, we employed a computational drug repurposing strategy for identifying drug hits against the RNA-dependent RNA polymerase (RDRP) protein of CoVID-19. Analysis of the human-virus protein-protein associations revealed that the viral RDRP (NSP12) is associated with multiple host proteins that partake in cellular processes, which indicated that NSP12 could be a potential target for drug discovery. This, combined with the fact that the RDRP protein is a potential antiviral target in several viral diseases, led us to consider the NSP12 as a potential drug target for CoVID-19. Owing to the absence of an experimentally-derived structure in the PDB, we constructed the NSP12 protein of CoVID-19 by homology modelling, and the potential druggable sites were analysed. The 13,533 entries in DrugBank were initially screened using the sequence of CoVID-19 NSP12. The 7 hits thus identified were subjected to a consensus docking and scoring strategy for identifying hits against the druggable site of CoVID-19 NSP12. Analysis of the docking scores and protein-ligand interactions revealed that two hits – N-alpha-[(benzyloxy)carbonyl]-n-[(1r)-4-hydroxy-1-methyl-2-oxobutyl]-l-phenylalaninamide and S-[5-(trifluoromethyl)-4h-1,2,4-triazol-3-yl] 5-(phenylethynyl) furan-2-carbothioate, had stronger binding affinity than remdesivir, which is being presently tested in clinical trials for its antiviral activity against CoVID-19. This indicated that these two compounds might be effective against CoVID-19, however, further experimentation is necessary for obtaining substantial evidence. We believe that the results of this study could offer a novel avenue for drug development against CoVID-19.

1. Introduction

Since its first emergence in Wuhan, China, in December 2019, the novel SARS-CoV-2 has spread rapidly across the globe. At present, there are 465,915 confirmed cases of CoVID-19, and 21,031 deaths. It took over three months to reach the first 100,000 confirmed cases, and only 12 days to reach the next 100,000 cases, and the number of infections are doubling every 2–10 days¹. Owing to the massive and rapid upsurge in the number of cases of CoVID-19, the WHO is calling for “urgent, aggressive actions” for contesting the viral pandemic². On 30th January 2020, the WHO declared the CoVID-19 outbreak as the “sixth public health emergency of international concern”². This situation urgently demands the rapid discovery of drugs for combating the massive surge in CoVID-19 infections. Drug repurposing offers a faster and cheaper strategy for repositioning drugs against novel targets^{3,4}. The strategy comprises the identification of novel targets for the existing drug candidates. This involves reusing an existing drug for treating a disease different from that the drug was originally intended for. As the pharmacokinetics of

these drugs are already known, it eliminates the necessity of phase I clinical trials, thus reducing the time and cost of drug discovery, which is of utmost necessity in the present global scenario^{4,5}. Despite the desperate necessity for the rapid identification of inhibitors against CoVID–19, it is important to ensure the safety of the drugs or vaccines that are proposed⁶. As the pharmacokinetics and toxicity profiles of drugs are already known, drug repositioning offers a safer strategy for identifying inhibitors against SARS-CoV–2.

SARS-CoV–2 is a single-stranded, positive-sense RNA virus belonging to the coronaviridae family. The viruses replicate in the host cytoplasm, which relies on the function of the RNA- dependent RNA polymerase (RDRP), also called the NSP12 protein in SARS-CoV–2. The RDRP protein is crucial to the life cycle of RNA viruses, as they mediate the replication of their RNA genomes⁷. The RDRP protein is a potential antiviral target, and has been targeted in numerous studies for the inhibition of several viral diseases, including dengue, hepatitis C, zika, West Nile disease, and Japanese encephalitis^{8–11}. In this study, we aimed to identify drugs against the RNA-dependent RNA polymerase (RDRP) of CoVID–19 by employing an *in silico* drug repositioning strategy using the DrugBank database. The RDRP protein is a potential viral target that has been targeted for the treatment of several viral diseases^{11,12}. Out of the seven drugs initially screened from DrugBank, we propose that two candidates, N- alpha-[(benzyloxy)carbonyl]-n-[(1r)-4-hydroxy-1-methyl-2-oxobutyl]-l-phenylalaninamide and S-[5-(trifluoromethyl)-4h-1,2,4-triazol-3-yl] 5-(phenylethynyl) furan-2 -carbothioate, had stronger binding affinity than remdesivir, which is being presently tested in clinical trials for its antiviral activity against SARS-CoV–2. These compounds could potentially target the RDRP of SARS-CoV–2, however, further experimental studies are necessary for validating the inhibitory potential of these compounds. We believe that this study provides an avenue for drug development against SARS-CoV–2 for impeding the massive global outbreak.

2. Materials And Methods

2.1 Virus–host PPI networks

The virus-human protein-protein interaction (PPI) networks were analysed using the VIRUSES web interface, version 10.5, of STRING^{13–15}, that was used to predict the associations between the proteins of the human SARS coronavirus (NCBI taxonomy Id: 694009) with *Homo sapiens* (NCBI taxonomy Id: 9606). In the constructed network, the line colour of the edges indicated the type of interaction evidence, the maximum number of interactors that were shown were not more than 10, and the minimum interaction score was 0.400.

2.2 Sequence retrieval, multiple sequence alignment (MSA) of viral RDRP proteins

The sequence of the RDRP of SARS-CoV–2 was retrieved from NCBI (YP_009725307). The sequences of the RDRP proteins of SARS-CoV–2, murine coronavirus (P0C6X9), human coronavirus (P0C6X6), bovine coronavirus (P0C6W9), middle ERSR coronavirus (K9N7C7), bat coronavirus (P0C6W2), and the human

SARS coronavirus (P0C6X7) were aligned using the MUSCLE algorithm and an MSA was subsequently constructed. A phylogenetic tree was constructed using the Neighbor-Joining algorithm, using a bootstrap value of 1000 with the CLC Genomics workbench, version 8.5.

2.3 Analysis of motifs and conserved domains

The motifs in the RDRP of SARS-CoV-2, murine coronavirus, human coronavirus, bovine coronavirus, middle ERSR coronavirus, bat coronavirus, and human SARS coronavirus were identified and compared using the Dynamic Motifs search approach in CLC Genomics Workbench version 8.5. The conserved domains in the RDRP protein were detected using the Conserved Domain Search (CD-Search) service of NCBI¹⁶.

2.4 Homology modelling, molecular dynamics (MD) simulations, validation, druggability analysis

The three-dimensional structure of the RDRP protein of SARS-CoV-2 was constructed by homology modelling with Modeller version 9.11^{17,18}. Template identification was achieved by performing BLASTp¹⁹ against the Protein Data Bank (PDB)²⁰. The structure of the NSP12 of SARS-coronavirus (PDB ID: 6NUR) was selected as the template for modelling the NSP12 of SARS-CoV-2²¹. The model thus obtained was validated by the Discrete Optimized Potential Energy (DOPE) score²². The structural stability of the constructed model of SARS-CoV-2 RDRP was confirmed by 10 ns MD simulations using the CABS-flex server²³, according to the method described by Chakraborty et al²⁴. The resulting model was validated by analysing the Ramachandran plot using ProCheck²⁵. The druggable sites in the RDRP was identified by SiteMap version 4.3²⁶.

2.5 Screening drug hits, molecular docking

The sequence of RDRP-SARS-CoV-2 (YP_009725307) was used as the target sequence for screening drug hits from DrugBank, version 5.1.5, by performing BLAST. A gap opening penalty of -1, a mismatch penalty of -3, and a gap extension penalty of -1 were used. The hits that were identified from DrugBank were docked to the druggable site identified in the structure of RDRP. A consensus docking and scoring strategy was employed, using the docking algorithms and empirical scoring functions of Glide SP and Glide XP^{27,28}. A grid of dimensions 33 Å × 33 Å × 33 Å was prepared around the druggable site that mapped to the RDRP domain of NSP12. The x, y, and z coordinates of the grid centre were 165.23, 141.089, and 142.37, respectively.

3. Results

3.1 Analysis of virus-host PPI networks

Analysis of the virus-host PPI network revealed that the RDRP protein is associated with the human hub proteins XPO1 and SPECC1, via the viral proteins 9b and nsp8, respectively (Figure 1). The combined interaction score between XPO1 and 9b was 0.8, while that between SPECC1 and nsp8 was 0.828,

indicating a strong association (Supplementary Table 1). The combined interaction scores of RDRP with nsp8 and 9b were 0.825 and 0.8, respectively, which also indicated a strong association. It might be possible that RDRP mediates the interactions of the viral proteins 9b and nsp8 with the human proteins, XPO1 and SPECC1. As the proteins XPO1 and SPECC1 participate in cellular processes^{29–31}, it might be possible that the viral RDRP affects the cellular processes of the host by mediating the association of 9b and nsp8 with XPO1 and SPECC1, respectively. The interactome network further revealed that eight viral proteins, namely, 9b, nsp8, PL-PRO, NC, nsp1, 7a, 3a, and ns2 are directly associated with the host proteins XPO1, SPECC1, PPP1CA, FGL2, JUN, SGTA and BCL2L1, STAT3 and CAV1, and PHB, respectively (Figure 1). The interaction scores are provided in supplementary Table 1. The viral S glycoprotein is indirectly associated with the human proteins CAV1 and STAT3 via the viral protein 3a. The viral M protein is indirectly associated with STAT3 and CAV1 via the viral S glycoprotein and viral protein 3a, and is indirectly associated with SGTA and BCL2L1 via viral protein 7a (Figure 1).

3.2 MSA and motif analysis

An MSA of the RDRP proteins of SARS-CoV-2, murine coronavirus, human coronavirus, bovine coronavirus, middle ERSC coronavirus, bat coronavirus, and the human SARS coronavirus was performed (Supplementary Figure S1). The results of motif analysis revealed that the RDRP proteins of all the viruses contained 4 motifs with sequence-specific conservedness, namely, motifs E, A, B, and D (Figure 2a). Analysis of the sequence-specific conservedness of motif A revealed that motif A was conserved in two regions across all the viruses analysed herein. Motif A was highly repetitive in SARS-CoV-2, which contained 5 copies of motif A. The sequence composition of motif A was DXXXXDXXXX, indicating that it is rich in aspartic acid (Figure 2b). The phylogenetic tree analysis revealed that the sequence of the NSP12 protein of SARS-CoV-2 was most closely related to the RDRP proteins of bat coronavirus and the human SARS coronavirus, as depicted in Figure 2c.

3.3 Refinement of homology model, validation

The DOPE score of the model (figure 3a) constructed by homology modelling was - 107039.9844. The Ramachandran plot of the resulting refined model revealed that 3.3% of the residues were in the disallowed regions (Supplementary Figure S2). The model was further refined by 10 ns MD simulations and the best model of was selected after clustering (figure 3b). Analysis of the inter-residue interactions revealed that the intra-protein contacts remain stable over the 10 ns simulation trajectory (Figure 3c). Analysis of the root mean square fluctuations (RMSF) revealed that the RMSF of the residues at the N-terminus and C-terminus was the highest, owing to the fact that they were located at the terminal ends of the protein (Figure 3d).

3.4 Compounds screened from DrugBank

The sequence of the RDRP of SARS-CoV-2 (YP_009725307) was used to screen the 13,533 entries in DrugBank. A total of 7 compounds were identified from DrugBank version 5.1.5 (Table 1). Of these, six compounds were experimental small molecules (compounds 1–6 in Table 1), while compound 7 was an

investigational small molecule. The pharmacokinetic properties of the compounds are provided in supplementary Table 2. The 2D structure of the compounds were retrieved from DrugBank in SMILES format and were converted to the 3D structures were obtained using OpenBabel³².

3.5 Selection of druggable site

Analysis of the druggable site of the RDRP protein of SARS-CoV-2 revealed two druggable sites (1 and 2) in the protein, having Dscore values of > 1 (Figure 4a). The SiteMap score is designed in such a way that druggable sites have a Dscore of > 1 and a Sitescore of > 0.8. Both these druggable sites had SiteScore and Dscore values greater than 0.8 and 1, respectively, indicating that both are potential druggable sites (Figure 4b). Conserved domain analysis revealed that residues 14–366 in the N-terminus of SARS-CoV-2 NSP12 protein mapped to the coronavirus RDRP domain, belonging to the Corona_RPol_N_superfamily (Pfam accession number pfam06478). The druggable sites were mapped to the conserved domains, which revealed that druggable site 2 mapped to the RDRP domain of NSP12 (Figure 4b). This indicated that site 2 is a potential druggable site, and targeting site 2 would inhibit the RDRP function of NSP12, which is a necessary component of SARS-CoV-2 replication. The don/acc ratio for site 2 was 1.053, indicating that the residues of this druggable region have a greater tendency to act as hydrogen bond donors in ligand interactions (Figure 4b). The results of SiteMap further revealed that site 2 was a tight binding site, revealed by the values of exposure and contact (Figure 4b). The volume of site 2 was 455.847 Å³. The hydrogen bond acceptor, hydrogen bond donor, hydrophobic, and hydrophilic surfaces of druggable site 2 are represented in Figure 4c.

3.6 Hit selection by consensus docking

The 7 compounds previously screened from DrugBank were further screened by consensus docking and scoring using Glide SP and Glide XP. Analysis of the hitlists from Glide SP (Table 2) and Glide XP (Table 3) revealed that the ranks of compounds 5 and 3 corresponded between Glide SP and Glide XP (Tables 2 and 3). Analysis of the Glide scores (Tables 2 and 3) revealed that the Glide SP and XP scores of compound 5 were the highest, while that of compounds 1 and 2 were the lowest among the seven compounds. There was a sharp decline in the Glide XP scores between compounds 7 and 4, and the difference between the Glide XP scores of compounds 7 and 4 was more than 1 kCal/mol (Table 3). The ranks of compounds 5 and 3 were consistent across the Glide SP and Glide XP hitlists. Therefore, compounds 5 and 3, with Glide XP scores of -7.05 and -6.589, respectively, were selected as the top scoring compounds from the hitlist.

The protein-ligand interactions of compounds 5 and 3 are compared with those of compound 7 (remdesivir) (Figure 5). Analysis of the protein-ligand interactions revealed that compounds 5 and 3 formed 2 and 1 hydrogen bonds, respectively, with the druggable site. The phenyl ring of compound 7 formed a pi-pi interaction with Y728. The analysis revealed that residues Y32, L207, L708, L240, and Y728 were involved in hydrophobic interactions with compounds 5, 3, and 7. The details of the

interactions are provided in Table 4. The protein- ligand interactions of the other compounds are provided in supplementary Figure S3.

4. Discussion

SARS-CoV-2 infects alveolar epithelial cells via receptor-mediated endocytosis by binding to the angiotensin-converting enzyme II (ACE2) receptor³³. The viruses multiply following entry into host cells. The multiplication of RNA viruses depends upon the simple strategies of replication and transcription. The RDRP protein, encoded by the *RdRp* gene, plays a central role in the replication of RNA viruses⁷. Therefore, the RDRP protein is a potential antiviral drug target⁷. Furthermore, as the 2019-nCoV/SARS-CoV-2 share the highest nucleotide sequence identity³⁴, the virus-human PPI networks were analysed using the SARS coronavirus database in STRING for identifying the potent druggable targets of SARS-CoV-2. Analysis of the virus-human PPI network revealed that the RDRP (NSP12) of SARS-CoV-2 is associated with the human proteins, XPO1 and SPECC1, via the viral proteins 9b and nsp8, respectively. The human protein, XPO1, also known as exportin-1, mediates the nuclear export of proteins and certain RNAs^{29,31}. The SPECC1 (sperm antigen with calponin homology and coiled-coil domains 1) protein, that is primarily expressed in the testes and plays a role in the organisation of the actin cytoskeleton³⁰. The association of RDRP with the viral proteins, 9b and nsp8, indicated that it might be possible that RDRP mediates the interactions of 9b and nsp8 with XPO1 and SPECC1, respectively. This, combined with the fact that RDRP is a potential antiviral target⁷, led us to consider that the RDRP (NSP12) of SARS-CoV-2 was a potential target for drug discovery. We therefore selected RDRP as a potential target for drug repurposing against SARS-CoV-2. Analysis of the motifs in the NSP12 protein revealed that five copies of motif A are present in the NSP12 protein of SARS-CoV-2. A study by Chou and Wang in 2015 revealed that the presence of D/E-rich repeats in proteins is associated with transcriptional regulation, DNA mimicry, and mRNA processing³⁹. It was found that motif A is rich in aspartic acid, and the presence of aspartic acid-rich repeats correlates with the fact that the RDRP domain catalyses RNA-dependent RNA transcription and plays a crucial role in viral replication. Several drugs, including chloroquine, baricitinib, remdesivir, lopinavir, and ritonavir, that are used to treat malaria, rheumatoid arthritis, and SARS-CoV infections, have been considered as potential drug candidates for the treatment of SARS-CoV-2 infections^{35,36}.

SARS-CoV-2 is a novel strain of coronavirus that was first reported in Wuhan, China in December 2019, and has caused more than 21,031 deaths to date. The WHO is calling for “urgent, aggressive actions” for combating the pandemic. By employing a drug repurposing strategy, we aimed to identify plausible drug candidates from DrugBank against the NSP12 protein of SARS-CoV-2, which possesses the RDRP activity that is crucial for viral replication. We employed sequence-based and structure-based computational approaches for screening drug candidates from DrugBank. A homology model of the NSP12 protein of SARS-CoV-2 protein was constructed using the PDB template 6NUR, and refined by 10 ns MD simulations. Druggability analysis of the NSP12 protein of SARS-CoV-2 revealed the presence of two druggable sites. The druggable sites were mapped to the conserved domains in NSP12, which

revealed that druggable site 2 mapped to the RDRP domain of NSP12. This indicated that site 2 would be a potential target site for inhibiting SARS-CoV-2, as targeting this site it would inhibit the RNA-dependent replication of viral RNA. We therefore selected site 2 as a potential druggable site for targeting SARS-CoV-2. Initially, the 13,533 entries in DrugBank were screened by performing a BLAST using the sequence of CoVID-19 NSP12 as target. The 7 drug hits that were obtained were docked to the druggable site in the RDRP protein, using a consensus docking and scoring approach. The rank of the top scoring hits, compound 5 and compound 3, agreed between the Glide SP and Glide XP hitlists, indicating that the binding of these two compounds did not occur randomly. The compounds interacted with the RDRP domain via hydrogen bonds, hydrophobic interactions, and pi-pi stacking interactions. Compounds 5 and 3 are experimental small molecules and the docking scores of the compounds were consistently better than compound 7, in both the docking platforms. Compound 7, remdesivir, is being presently studied in other clinical trials against ³⁷. Previous studies have demonstrated that remdesivir (compound 7) can bind to the RDRP of SARS-CoV-2 and inhibit the virus *in vitro*^{35,38}. We observed that the Glide SP and Glide XP scores of compounds 5 and 3 were higher than that of remdesivir (compound 7), indicating that these compounds might have better inhibitory properties against the RDRP activity of SARS-CoV-2.

5. Conclusion

In conclusion, we report two compounds, N-alpha-[(benzyloxy)carbonyl]-n-[(1r)-4-hydroxy-1-methyl-2-oxobutyl]-l-phenylalaninamide (compound 5) and S-[5-(trifluoromethyl)-4h-1,2,4-triazol-3-yl] 5-(phenylethynyl) furan-2-carbothioate (compound 3), which showed stronger binding affinity to the RDRP domain of SARS-CoV-2, than remdesivir, which has been reported to inhibit SARS-CoV-2 *in vitro*³⁵. This suggests that these compounds might serve as potent inhibitors of CoVID-19, however, further experimental studies are necessary for validating their inhibitory potential.

References

1. Novel coronavirus (COVID-19) situation. <https://experience.arcgis.com/experience/685d0ace521648f8a5beeeee1b9125cd> (2020).
2. WHO calls for urgent, aggressive actions to combat COVID-19, as cases soar in South-East Asia Region. <https://www.who.int/southeastasia/news/detail/17-03-2020-who-calls-for-urgent-aggressive-actions-to-combat-covid-19-as-cases-soar-in-south-east-asia-region> (2020).
3. Ashburn, T. T. & Thor, K. B. Drug repositioning: identifying and developing new uses for existing drugs. *Rev. Drug Discov.* **3**, 673–683 (2004).
4. Chong, C. R. & Sullivan, D. J. New uses for old drugs. *Nature* **448**, 645–646 (2007).
5. Nosengo, N. Can you teach old drugs new tricks? *Nature* **534**, 314–316 (2016).
6. Jiang, S. Don't rush to deploy COVID-19 vaccines and drugs without sufficient safety guarantees. *Nature* **579**, 321 (2020).

7. Venkataraman, S., Prasad, B. V. L. S. & Selvarajan, R. RNA dependent RNA polymerases: Insights from structure, function and evolution. *Viruses* **10**, (2018).
8. Lee, G. *et al.* Novel inhibitors of hepatitis C virus RNA-dependent RNA polymerases.
9. *Mol. Biol.* **357**, 1051–1057 (2006).
10. Lim, S. P. *et al.* Potent Allosteric Dengue Virus NS5 Polymerase Inhibitors: Mechanism of Action and Resistance Profiling. *PLoS Pathog.* **12**, (2016).
11. Riccio, F. *et al.* Development and validation of RdRp Screen, a crystallization screen for viral RNA-dependent RNA polymerases. *Open* **8**, (2019).
12. Malet, H. *et al.* The flavivirus polymerase as a target for drug discovery. *Antiviral* **80**, 23–35 (2008).
13. Velkov, T. *et al.* The RNA-Dependent-RNA Polymerase, an Emerging Antiviral Drug Target for the Hendra Virus. *Drug Targets* **15**, 103–113 (2014).
14. von Mering, C. *et al.* STRING: A database of predicted functional associations between proteins. *Nucleic Acids Res.* **31**, 258–261 (2003).
15. von Mering, C. *et al.* STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acid Res.* **1**, D433-7 (2005).
16. Cook, H. V., Doncheva, N. T., Szklarczyk, D., von Mering, C. & Jensen, J. Viruses.STRING: A virus-host protein-protein interaction database. *Viruses* **10**, 1–11 (2018).
17. Marchler-Bauer, A. & Bryant, S. H. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **32**, W331 (2004).
18. Andrej Šali. MODELLER A Program for Protein Structure Modeling. *protein Model. by Satisf. Spat. restraints.* 779–815 (1993).
19. Webb, B., Sali, A., Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. in *Current Protocols in Bioinformatics* 6.1-5.6.32 (John Wiley & Sons, Inc., 2014). doi:10.1002/0471250953.bi0506s47.
20. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Mol. Biol.* **215**, 403–410 (1990).
21. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242 (2000).
22. Kirchdoerfer, R. N. & Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Commun.* **10**, 1–9 (2019).
23. Shen, M. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524 (2006).
24. Jamroz, M., Kolinski, A. & Kmiecik, S. CABS-flex: Server for fast simulation of protein structure fluctuations. *Nucleic Acids Res.* **41**, W427-31 (2013).
25. Chakraborty, H. J., Gangopadhyay, A. & Datta, A. Prediction and characterisation of lantibiotic structures with molecular modelling and molecular dynamics simulations. *Rep.* **9**, 7169 (2019).
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. & IUCr. PROCHECK: a program to check the stereochemical quality of protein structures. *Appl. Crystallogr.* **26**, 283–291 (1993).

27. Schrödinger Release 2015-1: SiteMap, version 3.4, Schrödinger, LLC, New York, NY, 2015.
28. Halgren, T. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Med. Chem.* **47**, 1750–1759 (2004).
29. Friesner, R. A. *et al.* Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Med. Chem.* **49**, 6177–6196 (2006).
30. Köhler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Rev. Mol. Cell Biol.* **8**, 761–773 (2007).
31. Gaudet, P., Livstone, M. S., Lewis, S. E. & Thomas, P. D. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Bioinform.* **12**, 449–62 (2011).
32. Kudo, N. *et al.* Molecular cloning and cell cycle-dependent expression of mammalian CRM1, a protein involved in nuclear export of proteins. *Biol. Chem.* **272**, 29742–29751 (1997).
33. O’Boyle, N. M. *et al.* Open Babel: An Open chemical toolbox. *Cheminform.* **3**, 2–14 (2011).
34. Guo, Y.-R. *et al.* The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. *Med. Res.* **7**, 11 (2020).
35. Zhou, Y. *et al.* Network-based drug repurposing for novel coronavirus 2019- nCoV/SARS-CoV-2. *Cell Discov.* **6**, 14 (2020).
36. Wang, M. *et al.* Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **30**, 269–271 (2020).
37. Richardson, P. *et al.* Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *Lancet* **395**, E30–E31 (2020).
38. NIH clinical trial of remdesivir to treat COVID-19 begins | National Institutes of Health (NIH). <https://www.nih.gov/news-events/news-releases/nih-clinical-trial-remdesivir-treat-covid-19-begins> (2020).
39. Elfiky, A. A. Anti-HCV, nucleotide inhibitors, repurposing against COVID-19. *Life Sci.* **248**, 117477 (2020).
40. Chou, C. C. & Wang, A. H. J. Structural D/E-rich repeats play multiple roles especially in gene regulation through DNA/RNA mimicry. *Biosyst.* **11**, 2144–2151 (2015).

Declarations

Acknowledgements

Author Hirak Jyoti Chakraborty acknowledges Mrs. Anamika Chakraborty (Majumdar) for her support during figures quality improvement.

Conflicts of interest

The authors have no conflicts of interest to declare.

Tables

Due to technical limitations, Tables 1-4 are provided in the Supplementary Files section.

Figures

Figure 1

Virus-human protein-protein interaction (PPI) network showing the interactions between the proteins of the human SARS coronavirus and those of Homo sapiens. The red circles represent the viral proteins and the grey circles represent the human proteins. The colour of the lines (edges) indicates the type of

interaction, the thickness of the edges indicates the strength of the supporting data, and the different edges of different colours indicate the number of interactions.

Figure 2

a) MSA and comparison of motifs present in the RDRP proteins of SARS-CoV-2, murine coronavirus, human coronavirus, bovine coronavirus, middle ERSR coronavirus, bat coronavirus, and the human SARS coronavirus. Four motifs, motif E, A, B, and C were identified in all the viral RDRP sequences, that are indicated by yellow rectangles, blue triangles, red triangles, and green rectangles, respectively. The sequence-specific conserved motifs that were present in all the sequences are outlined by rectangles. b) Sequence composition of motif-E (yellow), motif-A (blue), motif-B (red), and motif-C (green). c) Phylogenetic cladogram of the viral RDRP protein sequences. The bootstrap percentage value obtained from 1000 samplings. The RDRP sequence of SARS-CoV-2 (CoVID-19) was phylogenetically closest to the RDRP sequence of the bat coronavirus and human SARS coronavirus. The sister group is indicated by a blue rectangle.

Figure 3

A) Model of the RDRP (NSP12) protein of SARS-CoV-2 that was constructed by homology modelling. The helices are coloured in blue, the beta sheets are represented in red, while the coiled regions are indicated in grey. and structural stability of model. B) Superimposition of the best models obtained from the clusters after the 10 ns MD simulations. C. The residue contact maps of the constructed model of SARS-CoV-2 RDRP. The map shows a matrix with the contacts. The colour scale, indicating the contact distance between the residues is provided below the map. D. RMSF of the constructed model over the 10 ns simulation trajectory.

Figure 4

a) The RDRP domain of the NSP12 protein of SARS-CoV-2 is represented in orange surface representation. Druggable sites 1 and 2 are in grey and purple representation, respectively. b) Graphical representation of the properties of the druggable sites 1 and 2. SiteScore and Dscore indicate the potential druggability of binding sites. Exposure indicates the degree of solvent exposure, while enclosure represents the degree of openness of the site to the solvent. Contact represents the strength of ligand-target van der Waals interactions. Phobic and philic indicate the hydrophobic and hydrophilic character of the binding site, while balance indicates the phobic:philic ratio. Don/Acc indicates the ratio of hydrogen bond donor property to hydrogen bond acceptor property. c) The hydrogen bond donor, hydrogen bond acceptor, hydrophobic, and hydrophilic surfaces of druggable site 2 are represented by blue, red, yellow, and green surfaces, and surface areas are indicated in the corresponding colours.

Figure 5

The binding poses of compound 5, compound 3, and compound 7 (upper panel). The interacting residues are represented in green colour, and the ligands are in grey colouration. The hydrogen bonds are indicated by purple dashed lines. The pi-pi interactions are indicated by light blue dashed lines. The protein-ligand interactions are provided in the lower panel.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinformation.pdf](#)
- [Tables.docx](#)