Novel Inhibitors to ADP Ribose Phosphatase of SARS-CoV-2 Identified by Structure-based High Throughput Virtual Screening and Molecular Dynamics Simulations

Dhrumi C. Patel[1]*, Katherine R. Hausman*[1], Muhammad Arba[2]*, Annie Tran[1], Phillip M. Lakernick[1], Chun Wu[1,#]

[1]Department of Molecular & Cellular Biosciences, College of Science and Mathematics, Rowan University, Glassboro, New Jersey 08028, United States.

[2]Faculty of Pharmacy, Universitas Halu Oleo, Kendari 93232, Indonesia.

*Contribute equally

#To whom correspondence should be addressed: wuc@rowan.edu

ABSTRACT

The outbreak of a new coronavirus (SARS-CoV-2) was first identified in Wuhan, People's Republic of China, in 2019, which leads to a severe, life-threatening form of pneumonia (COVID-19). Research scientists all around the world have been trying to find small molecule drugs to treat COVID-19. In the present study, a conserved macrodomain, ADP Ribose phosphatase (ADRP), of a critical non-structural protein (Nsp3) in all coronaviruses was probed using large-scale Molecular Dynamics (MD) simulations to identify novel inhibitors. In our virtual screening workflow, the recently-solved X-ray complex structure, *6W6Y*, in addition to a substrate-mimics was used to screen 17 million ZINC15 compounds using drug property filters and Glide docking scores. The top twenty output compounds each underwent 200 ns MD simulations (i.e. 20x200 ns) to validate their individual stability as potential inhibitors. Eight out of the twenty compounds showed stable binding modes in the MD simulations, as well as favorable drug properties from our predications. Therefore, our computational data suggest that the resulting top eight out of twenty compounds could potentially be novel inhibitors of ADRP for ADRP of SARS-CoV-2.

Keywords: SARS-CoV-2, COVID-19, Nsp3, macrodomain, docking, MD simulation

1. Introduction

Since the first cases in Wuhan, People's Republic of China, at the end of year 2019, the spread of Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2), which causes the disease COVID-19, has escalated worldwide. It has infected approximately 212M people and killed approximately 4.43M, leading the World Health Organization (WHO) to classify it as a pandemic. SARS-CoV-2, belonging to a group of beta-coronavirus, resembles the other two members: SARS-CoV and Middle East Respiratory Syndrome (MERS-CoV). SARS-CoV and MERS-CoV were also responsible for earlier outbreaks of deadly respiratory diseases, but there is no single drug to treat them. Currently, only one drug, Remdesivir, has been approved by the US food and drug administration (FDA) to treat SARS-CoV2, but its efficacy is not very high. The demand for therapeutics keeps increasing as the delta variant leads to an increase in breakthrough infections among vaccinated people.

The genome of SARS-CoV-2 encodes an approximately 7,096-residue-long polyprotein, which consists of many structural and non-structural proteins (NSPs).[1] Like other coronaviruses, SARS-CoV-2 uses positive-sense RNA genome to code NSPs and structural proteins, such as the spike glycol-protein, envelope, and membrane.[2] The 16 NSPS found in SARS-CoV-2 form a large, membrane-bound replicase complex, in which Nsp3 is the largest component, with a residue range between 1-1,922.[2] Nsp3, which exists in all CoVs, is a multidomain (macrodomain) protein, ADP ribose phosphatase domain (ADRP, also known as macrodomain), which contains the N-terminal ubiquitin-like domain (Ubl), SARS-unique domain, Papain-like proteinase (PL^{pro}), nucleic acid binding (NAB) domain, also containing transmembrane domain and Y-domain (**Fig. 1A**). Macrodomain (Mac1) prevents host immune response to viral infection by removing ADP-Ribosylation from modified host protein.[3] Host

poly-ADP ribose polymerase (PARP) enzymes catalyze the transfer of the ADP-ribose phosphate group to their target proteins, primarily attached to the Mac1 to perform its function (assisting in DNA damage repair, cellular stress, and proper immune response) which allows the host to recognize and attack viruses like SAS-COV2.[2, 4] The binding of SARS-CoV-2 to Mac1 is crucial, as it initiates virulence and RNA replication.[5] Type I interferons (IFN-I) impel the intrinsic induction by blocking the phosphorylation dimerization and resulting in nuclear translocation of the host IFN regulatory factor 3 (IRF3), which is a transcription regulator important for innate immunity.[2, 6] The role of Mac1 is crucial, as their inhibition helps to reduce the viral load, facilitate recovery and interferes with the host's immune response to SARS-CoV-2 (which makes it an attractive protein target).[2, 6] Although SARS-CoV-2 has higher transmission efficiency from human to human than previous viruses (for instance, SARs and MERs), the Mac1 protein exists in both SARs and MERs, which makes it a common drug target for both viruses.[7, 8]

Although there have been studies (**Table 1**) which focused on identifying novel drugs targeting viral macrodomains using the virtual screening and pharmacophore approach, there have been no such studies conducted which utilized elaborate virtual screening methods. In a 2020 study conducted by Babar et al., a total of 64,043 drugs were screened, in which potential inhibitors were chosen based on their docking score and high binding affinity for key active site residues.[9] In a 2020 study conducted by Debnath, virtual screening was performed on the 113K MolPort database, from which six candidates were selected based on their XP glide score range, whose binding affinities were validated using free energy calculations and MD simulations.[10] However, Debnath did not analyze the stability of the top potential inhibitors using the MD

simulation.[9, 10] Using a similar approach, the present study uses a much larger database and longer MD simulations for more compounds.

In the details, the structure of ADP-ribose phosphatase in the complex of AMP (*PDB ID: 6W6Y*, **Fig. 1B-1C**) was used in the structure-based high throughput screening of zinc15 library with 17 million compounds. AMP as a substrate, which is used as a monomer in RNA, also plays an important role in intracellular signaling and cellular metabolic processes.[9] Next, we have utilized a long MD simulation (200 ns) to examine the top 20 hits from the virtual screening. Followed by advanced MD simulation and MM-GBSA binding energy calculation methods was used to obtain better estimate of the binding affinity. Eight compounds out of the 20 hits showed significantly improved binding free energy score and good drug properties. The present study adds important knowledge to the ongoing efforts of finding the potential drug target compounds and novel inhibitors to ADP-Ribose phosphatase of SARS-CoV-2.

2. Methods

A virtual screening workflow (VSW) in **Fig. 2** was developed to identify lead inhibitors to the ADP-Ribose phosphatase of SARS-CoV-2 from ZINC 15 drug-like library with 17 million entries. This VSW consists of ten essential steps including drug property prediction, molecular docking, and molecular dynamics simulation. The first step of the VSW was the input of the prepared protein structure and ligand library. Then, the compounds were filtered by drug property in setup 2 and docking with multiple Glide docking score functions with increasing accuracy (Glide HTVS, SP and XP) in step 2-5. Ligand similarity analysis was performed to identify different molecular scaffolds in step 6. In step 7, the ligands that have either a worse Glide XP score than the reference compound (the crystal ligand AMP) or more than one red flag in drug property (# of star, from QikProp) were removed, the top 20 compounds were manually

selected from the remaining compounds by maximizing the number of molecular scaffolds (i.e., different ligand cluster IDs). In steps 8 and 9, the 200ns MD simulation was carried out, followed by post simulation analyses including MMGBSA binding free energy calculation, simulation interaction diagram analysis, and protein conformation clustering analysis. In step 9, the prediction of ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) was used to check the human oral bioavailability of potential drug candidates. Finally, the compounds that had better MMGBSA binding free energy than the reference compounds were selected and presented in the main text. The details of these ten steps are presented in the following six modules.

2.1. Preparation of protein and ligand library

The crystal structure (6W6Y) of ADP ribose phosphatase of Nsp3 from SARS-CoV-2 was prepared using Maestro's Protein Preparation Wizard.[11] The protein was preprocessed to assign correct bond orders, add hydrogen atoms, create disulfide bonds, and to delete water beyond 5 angstrom (Å) from hetero groups. The charge state of the titratable residues was optimized using PROPKA at a pH of 7. A restrained minimization was done to relax the protein using an OPLS3 force field.[12] Epik, a tool based on accurate methodologies from Hammer and Taft, was used to generate the proper ionization state of each ligand.[11] The lowest tautomeric state for each ligand structure was selected and minimized to relax the ligands to a best fit structure. Lastly, a geometry optimization was performed using quantum mechanics methods in Jaguar.

2.2. Filtering and Docking

The prepared merged protein-ligand complex was put through Schrodinger's Virtual Screening Interface, where it was prefiltered through Lipinski's Rule and filtered with ADMET

risk parameter assessments through QikProp. The parameters and cut-off values employed when screening using QikProp were described in **Table S1**.[13] The active binding site of ADP ribose phosphatase was defined using the center of the ligand, from which the grid file was generated using a van der Waals scaling factor of 1 and a partial charge cutoff of 0.25. The prepared compounds were docked into the generated grid of the protein receptor using an OPLS3 force field and their docking scores were calculated using both SP and XP scoring functions.[12] The default settings were used as the parameters for the scoring function: Active Epik state penalties to the docking score, dock sampling was flexible with sample nitrogen-inversions, same ring conformations, and bias sampling of torsions for amides, which only penalized non-planar conformations.[14, 15] The results of docking concluded with 20 top compounds with high docking scores, indicating that they all had high affinity for the receptor.

2.3. Ligand similarity clustering

The ligand similarity clustering was done on the Canvas program. First, digital fingerprints of 3D ligand structures were generated using 3-point pharmacophore.[16] Next, hierarchical clustering with default parameters was performed to group similar compounds into different clusters using their fingerprints and a cluster ID was assigned to each compound.[17, 18]

2.4. MD simulation

The twenty prepared receptor-ligand complexes were used to construct MD simulation systems. The complexes were solvated in an orthorhombic water box with a buffer distance of 10Å using a predefined SPC water model.[19] A 0.15 M NaCl salt concentration was added to neutralize the system. The systems were built with an OPLS3 force field using Desmond System Builder in Maestro on a Linux operating system.[12] The relaxation/minimization and production runs were set up using the Desmond module following our previous procedure.[20]

2.5. Post simulation analysis

2.5.1. Simulation interaction diagram (SID) analysis

The Desmond SID tool in Maestro was used to calculate the Root-Mean-Square Deviation (RMSD), the Root-Mean-Square Fluctuation (RMSF), the Secondary Structural Elements (SSE), and the residue-ligand interactions and contacts throughout the course of the simulation. The protein Cα and ligand RMSD plots obtained from the SID analysis were analyzed to ensure the convergence of each of the MD simulations. A relatively flat plot would imply that a steady state was reached.

2.5.2. Trajectory clustering analysis

The Desmond clustering tool was utilized to organize the complex structures from the trajectories.[21] The parameters included using the backbone RMSD matrix as the structural similarity metric and using a 2.5 Å merging distance cutoff and average linkage for the hierarchical clustering. The centroid structure, the structure with the most neighbors in the structural family, was chosen to represent each structural family. Of the centroids, the most abundant structures of the populated structural families were extracted and analyzed further. Clustering is used to identify the most abundant conformations and reduces complexity.

2.5.3 Binding energy calculations and decompositions

The ligand-binding affinities on the frames obtained in the last 50ns of each MD simulation were calculated using the surface-area-based Generalized Born model [22, 23] with an implicit membrane solvation model (VSGB 2.0).[24] as explained in our previous works [20].

2.6. ADMET Prediction

Prediction of ADMET properties for the eight best ZINC compounds were performed on the SwissADME web server (http://www.swissadme.ch/). This server is developed by the Swiss Institute of Bioinformatics, used to provide physiochemical descriptors, ADMET parameters, pharmacokinetic properties, and drug-like small molecule inhibitors to support drug discovery.[25] The SMILE codes for each compound were inserted into the webserver to receive their ADMET properties.

3. Results

MM-GBSA results reveal the top eight compounds according to their ligand binding energy. A total of 17M zinc compounds were screened from which compounds with docking scores higher than -9.7 kcal/mol were considered, discarding the others. The top twenty compounds were picked based on their SMILE code, numbers of STARs, cluster IDs, and centroid (**Fig. S3**). The various properties including the glide XP docking, MM-GBSA, and MD simulation helped in picking the top eight compounds out of twenty total compounds. The MM-GBSA method was used to estimate the binding free energy of the twenty compounds. The binding interaction between ligand-protein receptors is specified by the free energy binding. Using the crystal ligand as the control, the top picked compounds possess significantly higher binding energy against the Mac1. The results of other energy terms such as Van Der Waals energy (VDW), electrostatic energy (ELE), hydrophobic, and ligand-receptor RMSD strongly indicate that these top-picked compounds could be targeted against SARS-CoV-2. The hydrophobic value (-46.7±3.6 kcal/mol) was favorable for the binding of ZINC000096223736, followed by other compounds. ZINC000096223736 had superior binding energy than the other compounds which is visible in (**TABLE 2**). The protein-ligand contact pattern shows the protein

binding pocket, and the main residues responsible for the interaction. Protein-ligand contacts of the top eight compounds during MD simulation are shown in **(TABLE 3)**, the most abundant residues display the highest number of interactions throughout the results section.

ADMET properties show good human oral bioavailability

The predicted ADMET properties for the eight compounds show that there is high intestinal absorption with only one of them showing chances to distribute into the brain. However, some of the compounds do inhibit the cytochrome P450 enzymes (CYPs) including CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4 inhibitors, which indicated that these compounds could be metabolized. CYP3A4 possesses the highest activity in the small intestine and liver and metabolizes 50% of the medicines.[26] The CYP inhibitors also indicate an increase in the plasma concentration of the drug. BBB permeability protects the exposure of molecules that are toxic to the neurons in the brain. All eight compounds fulfill the conditions of drug-likeness properties without violation of Lipinski rule of five including MW < 500, calculated octanol-water partition coefficient (LogP) $\leq$ 5, some hydrogen bonding acceptors $\leq$ 10, as well as several hydrogen bonding donors $\leq$ 5. The PAINS (pan-assay interference compounds) alert system also gave off zero alerts to all eight compounds, which indicates a low chance of false positives from occurring (TABLE 4). The docked complex of the top eight ligands compared with the crystal complex structure is performed to view the closest interaction-based measures and pose predictions. (**Fig. 3**) The top eight ligands show a good binding pose with the crystal structure, further validating the top eight hits.

Crystal complex structure was stable in the MD simulation. The simulation interaction diagram of the crystal structure shows the RMSD plot obtained from the MD simulation. Monitoring the RMSD of the protein gives insight into the structural conformation

obtained throughout the simulation. The RMSD plot has equilibrated and fluctuated towards the

end of the simulation, it is acceptable for the crystal structure is a globular protein. (**Fig. 4A**) A

ligand atom interaction with protein residues from the MD trajectory provides insight on four

different hydrophobic interactions such as ASP32, PHE156, ALA154, and ILE23, which are the

four residues interacting with the ligand atom. **(Fig. 4B)** The protein SSE plot shows the

recurring arrangements of close amino acids through distribution by residue position throughout

the protein structure. The residue indicated 32.40% of the helix and 19.68% of the strand, which

made the total percentages of SSE of the residue position to be 52.08%, with the rest of the area

being random coil. **(Fig. 4C)** The peaks in the RMSF plot analyses the portion of the protein that

fluctuates the most during the simulation. The tails (*N*- and *C*-terminal) fluctuates more than any

other parts of the protein. **(Fig. 4D)** The interaction fraction plot indicates the ligand-protein

interaction throughout the simulation. The four residues with the 2D ligand-protein interaction in

Fig. 4B, show a greater amount of hydrogen bonds in the plot **(Fig. 4E)** during the simulation.

RMSD analysis of protein and ligand is stabilized. After the screening of all drugs, the

best hits were subjected to MD simulation to better understand their behavior. RMSD is used to

measure the crystallographic binding pose. Structural fluctuation patterns during 200ns MD

simulation are described in terms of C RMSD of the top eight compounds. The structures are

compared by the binding interaction and the energy between protein and ligand. (**Fig. 5**) In this

case, a low RMSD score is acceptable. Protein RMSD gives insight into its structural

conformation throughout the simulation, whereas ligand RMSD indicates how stable the ligand

is concerning the protein and its binding pocket. Most changes are shown in between 1-3 Å

which is perfectly acceptable for a small protein like these. The ligand RMSD values stayed

within the RMSD of the protein range, which indicates that the ligand has not diffused away

from its initial binding site. The average RMSD value of the compound (ZINC014116837) was found to be relatively high. The protein-ligand convergence of the compound (ZINC096223736) was observed between 50 and 150ns. It was observed to be at 1Å. The convergence was observed between 0 and 50ns in the compound (ZINC082673). All eight compounds showed minor fluctuations throughout the graphs. These results show that the lower the RMSD fluctuates, the better the model is in comparison to the target structure.

MD simulation shows improvement in the binding pose of the top eight ligands. The trajectories of the receptor-binding domain (RBD) of the top eight ligands were analyzed by comparing the ligand XP docking binding pose before and after MD simulation. (**Fig. 6**) MD simulation is utilized to find the most dissimilar conformational changes. During the simulation, a ligand may significantly change from the originally bound conformation to optimize the overall interactions with the receptor. Rotatable bonds in the ligand may lead to high RMSD concerning initial bound conformation. The compound ZINC096223736 shows a stronger binding pose before and after the simulation followed by ZINC426746041 and ZINC079784201 also according to its MM-GBSA score shown in (**Table 2**). The 2D interaction of the ligand atom with protein residues from the MD trajectory of the top eight ligands with different active site amino acids is shown. (**Fig. 7**) The results show that the selected hits give a good binding affinity towards the active site of the protein.

Protein-ligand interaction analysis reveals novel residue interactions in all eight compounds. Protein interactions with the ligand were monitored throughout the MD simulation. Protein-ligand interactions can be categorized into four different types as shown. (**Fig. 8**) The highest amount of hydrogen bonding was observed on ZINC082673, ZINC014116837, and ZINC217844024. Leu126, ALA154, and ILE23 are the main residues containing the hydrogen

bonds in all eight compounds. It can lead to a strong influence on drug specificity, metabolization, and absorption. Hydrophobic interaction generally involves a hydrophobic amino acid and an aromatic or aliphatic group on the ligand. ZINC082673, ZINC121003678, and ZINC014116837 are observed to have the highest number of hydrophobic contacts. No ionic or polar interactions were observed in the eight compound graphs. Water bridges were observed in every plot where there was a hydrogen-bonded protein-ligand interaction mediated by a water molecule. The last 50ns of each 200ns simulation show little deviation, indicating convergence **(Fig. S5)**. All compounds showed higher hydrogen-bonding and hydrophobic interactions compared to the crystal structure, which leads to higher GI absorption of the compounds shown. (**Table 4**)

The secondary structure examination portrays minor differences. Protein SSE was monitored throughout the MD simulation. The plots summarize the SSE distribution by residue position throughout the protein structure. It is categorized into Alpha-helices, Beta-strands, and random coils (**Fig. 9**). Alpha-helices mainly have hydrophobic residues which are found in the core of the protein or are transmembrane proteins. On the other hand, beta-strands contain patterns of hydrophobic and polar amino acids. The random coil is a polymer conformation, where it is not one specific shape but instead a statistical distribution of all chains indicated by the white spaces in the plots. SSE is more rigid than the unstructured part of the protein, therefore they fluctuate less than the loop regions.

The RMSF shows fluctuation in localized regions of the protein and ligand. The flexibility of each compound was acquired by using RMSF. Protein C RMSF characterizes the local changes along the protein backbone during simulation. **(Fig. 10)** It measures the deviation between the particle position and the reference position. The peaks indicate areas of protein that

fluctuate the most during the simulation. The N- and C- terminal tails show more fluctuation than the other parts of the protein, due to the higher amount of proteins binding with the residues. ZINC014116837 showed the greatest fluctuation nearly at 5Å around residue 100 positions. All the other compounds fluctuated around the same residue position nearly from 1-3 Å, which is also reflected in trajectory clustering analysis. Another small fluctuation of ZINC014116837 and ZINC096232566 was observed at residue position 130. Lastly, a fluctuation of ZINC217844024 and ZINC079784201 took place towards the end at the residue position 155. Using the crystal structure as a positive control, the complexes showed much higher residual fluctuation at different positions throughout the graph.

4. Discussion

The cluster of pneumonia caused by SARS-CoV-2 created great challenges in public health all around the world. This pandemic has not only challenged people in terms of their health, but also in terms of the economy. To subsist with the virus, researchers and scientists are focusing on different aspects of the infections caused by it. The SARS-CoV-2 genome consists of 16 Nsps and 4 structural proteins. There are many studies that targeted domains like PLpro, MTase, NendoU, RdRp protein, and 3CLpro while others focused more on the host proteins of the SARS-CoV-2 genome. Out of the 16 Nsps, Nsp3 plays a major role in transcription and translation by seizing the host immune system. Nsp3 is the largest protein encoded by SARS-CoV-2, it binds to viral RNA as well as other viral proteins.

The Nsp3 is reported to interact with Nsp2 which leads to intervention at an earlier or later stage of viral replication. It has been stated that the macrodomain of Nsp3 enzyme activity plays an essential role in pathogenesis.[27] The 16 domains (~Res: 1922) and regions of Nsp3 play an important role in transcription and translation, which interacts with the protein host by

taking over the host immune system. There is an essential process required to bind ADRP to this domain. Targeting this domain with high receptor binding inhibitors could help to reduce the viral implications caused by the pandemic. Therefore, using different drug targeting approaches like computer-aided drug design or rational drug design could help find a drug to target this domain and redeem the host immune system.

The emergency use of the FDA-approved drug Remdesivir is not as effective due to its associated problems, therefore its use is very limited. Therefore, there is still a need to either refine older drugs or search for new ones to treat SARS-CoV-2. In this regard, the use of structure-based high throughput virtual screening methods is the most useful way to find new drugs for COVID-19 based on their properties. In the present study, we used the virtual screening approach to discover the potential inhibitors to target ADP ribose phosphatase. Based on our findings using the bioinformatics tools, along with a drug similarity search, we analyzed compounds from Zinc15 and targeted novel inhibitors of ADP-Ribose phosphatase of SARS-CoV-2 using the virtual screening workflow which gave us a promising top twenty hits. The top twenty compounds were then further validated and shortlisted to eight of the inhibitors, which were further validated by the MM-GBSA score of binding free energy and MD simulation. The MM-GBSA score helped to specify the binding interaction between ligand-protein receptors.

The protein-ligand interaction further confirmed the top potential inhibitors. The large dataset and extended high throughput virtual screening method portray the best interactions between ligands of a molecular target to form a complex. The adverse effects and related articles of the selected compounds were checked through CAS Scifinder and PubChem, where the compounds showed no adverse effect. Moreover, the ZINC082673 compound is found to be useful in treatments of bacterial infection and the NadD inhibition leads to suppression of

bacterial growth. NadD synthetase uses energy from ATP and is widely used as a drug target in various microorganisms. These results further validate the top hits to be the potential inhibitors.

Based on the results, the potential binding and inhibiting effects of the top eight compounds are indicated. The use of the structure-based high throughput virtual screening method is the most useful approach to find the potential molecules that could target the macrodomain. After examining 17 million Zinc15 compounds using structure-based high throughput virtual screening methods, the most potential hits were validated by MD simulation. This study has the potential to assist and repurpose the drug design. In vitro experiments can be carried out as this study can facilitate the global efforts in the speedy development of potential drug candidates against SARS-CoV-2.

5. Conclusion

This study may assist in further investigating and testing the novel inhibitors of SARS-CoV-2. We performed a thorough investigation of a total of 17 million Zinc15 compounds to examine using a structure-based high throughput virtual screening method which provides the most potential hits. The molecular dynamic simulation further validated the top hits. In this study, twenty potential compounds were selected from which eight potential inhibitors of SARS-CoV-2 showed commendable docking scores ranging from -9.9 to -11.7 kcal/mol. The computational pipeline gives us the best top eight hits, which exhibit good binding affinity towards the active site. Based on the XP glide docking, the binding affinity, and ADMET properties, the top eight Zinc compounds are the potential inhibitors of ADP-ribose phosphatase of Nsp3 of SARS-CoV-2. This paper could provide great knowledge about the potential inhibitors to target SARS-CoV-2.

Competing Interests Statement

No potential competing interests were reported by the authors.

References

1. Naqvi, A.A., et al., *Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach.* Biochimica Et Biophysica Acta-Molecular Basis of Disease, 2020. **1866**(10): p. 165878.
2. Michalska, K., et al., *Crystal structures of SARS-CoV-2 ADP-ribose phosnhatase: from the apo form to ligand complexes.* IUCrJ, 2020. **7**: p. 814-824.
3. Egloff, M.P., et al., *Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains.* Journal of Virology, 2006. **80**(17): p. 8493-8502.
4. Ahmad, Z., *ATP synthase: A potential molecular drug target.* Faseb Journal, 2019. **33**: p. 470.1.
5. Fehr, A.R., et al., *Viral Macrodomains: Unique Mediators of Viral Replication and Pathogenesis.* Trends in Microbiology, 2018. **26**(7): p. 598-610.
6. Yanai, H., et al., *Revisiting the role of IRF3 in inflammation and immunity by conditional and specifically targeted gene ablation in mice.* 2018. **115**(20): p. 5253-5258.
7. Liu, J., et al., *A comparative overview of COVID-19, MERS and SARS: Review article.* International Journal of Surgery, 2020. **81**: p. 1-8.
8. Jian Lei, Y.K., Rolf Hilgenfelda, *Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein.* Antiviral Research, 2018. **149**: p. 58-74.
9. Dunn, J.G., M, *Physiology, Adenosine Triphosphate.* StatPearls Publishing, 2021.
10. Debnath, P., et al., *In Silico Identification of Potential Inhibitors of ADP-Ribose Phosphatase of SARS-CoV-2 nsP3 by Combining E-Pharmacophore and Receptor-Based Virtual Screening of Database.* ChemistrySelect, 2020. **5**(30): p. 9388-9398.
11. Sastry, G.M., et al., *Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments.* Journal of Computer-Aided Molecular Design, 2013. **27**(3): p. 221-234.
12. Harder, E., et al., *OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins.* Journal of Chemical Theory and Computation, 2016. **12**(1): p. 281-296.
13. Schrödinger, L., *QikProp, version 4.4*. 2015, Schrödinger, LLC: New York, NY.
14. Friesner, R.A., et al., *Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1739-1749.

15. Friesner, R.A., et al., *Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes.* Journal of Medicinal Chemistry, 2006. **49**(21): p. 6177-6196.

16. Wood, D.J., et al., *Pharmacophore Fingerprint-Based Approach to Binding Site Subpocket Similarity and Its Application to Bioisostere Replacement.* Journal of Chemical Information and Modeling, 2012. **52**(8): p. 2031-2043.

17. Zolfaghari, F., et al., *Hierarchical cluster analysis to identify the homogeneous desertification management units.* Plos One, 2019. **14**(12).

18. Leal, W., et al., *How frequently do clusters occur in hierarchical clustering analysis? A graph theoretical approach to studying ties in proximity.* Journal of Cheminformatics, 2016. **8**(1): p. 4.

19. Mark, P. and L. Nilsson, *Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K.* Journal of Physical Chemistry A, 2001. **105**(43): p. 9954-9960.

20. Kumar, V., H. Liu, and C. Wu, *Drug repurposing against SARS-CoV-2 receptor binding domain using ensemble-based virtual screening and molecular dynamics simulations.* Computers in Biology and Medicine, 2021. **135**: p. 104634.

21. Kevin J. Bowers, E.C., Huafeng Xu, Ron O. Dror, Michael P. Eastwood,; Brent A. Gregersen, J. L. K., Istvan Kolossvary, Mark A. Moraes, Federico D. Sacerdoti,; John K. Salmon, Y. S., David E. Shaw, *Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters.* Proceedings of the 2006 ACM/IEEE conference on Supercomputing, 2006: p. 84.

22. Ghosh, A., C.S. Rapp, and R.A. Friesner, *Generalized born model based on a surface integral formulation.* Journal of Physical Chemistry B, 1998. **102**(52): p. 10983-10990.

23. Yu, Z.Y.J., M. P.; Friesner, R. A., *What role do surfaces play in GB models? A new-generation of surface-generalized Born model based on a novel Gaussian surface for biomolecules.* Journal of Computational Chemistry, 2006. **27**: p. 72-89.

24. Li, J.N., et al., *The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling.* Proteins-Structure Function and Bioinformatics, 2011. **79**(10): p. 2794-2812.

25. Zoete, V., et al., *SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening.* J Chem Inf Model, 2016. **56**(8): p. 1399-404.

26. Bibi, Z., *Role of cytochrome P450 in drug interactions (Retraction of vol 5, 27, 2008).* Nutrition & Metabolism, 2014. **11**: p. 27.

27. Frick, D.N., et al., *Molecular Basis for ADP-Ribose Binding to the Mac1 Domain of SARS-CoV-2 nsp3.* Biochemistry, 2020. **59**(28): p. 2608-2615.