Open Access
Inkwell Innovations in Life Sciences
ISSN: 3079-3335 (Print) 3079-3343 (Online)
https://inkwellliterature.com/index.php/lifescience
Volume 1, Issue: 1, Pages: 13-26, Date Published: Dec 31, 2024

Research Article

Computational Identification of New Inhibitors for 3C-Like Protease: A Potential Drug Target in Coronavirus

Muhammad Numan[1,#], Amar Ajmal[1,#], Iqra Akbar[1], Syed Mubassir Shah[2,*], Arif Ali[3]

[1]Department of Biochemistry, Abdul Wali Khan University, Mardan, KPK, Pakistan
[2]Department of Biotechnology, Abdul Wali Khan University, Mardan, KPK, Pakistan
[3]Laboratory of Structural Bioinformatics, State Key Lab of Microbial Metabolism, Department of Bioinformatics & Biostatistics, School of life Sciences & Biotechnology, Shanghai Jiaotong University, Shanghai, China,
[#]These authors contributed equally.

*Correspondence to; Syed Mubassir Shah, Email; syedmubassir@awkum.edu.pk

ABSTRACT

SARS CoV-2 is a single-strand RNA, positive sense and enveloped beta coronavirus that causes respiratory, nervous, hepatic, and human gastrointestinal diseases. Due to the presence of spike glycoprotein, it appears crown or corona-like. SARS-CoV-2 is more contagious and has caused more deaths and infections. Old-age people and immune-compromised patients face the greatest risk of death. The main protease (Mpro), also known as the 3CL pro, is responsible for virus replication. Because it differs from human proteases, the main protease of SARS-CoV-2 is considered to be an attractive target for drugs. The process of developing a new drug is difficult, time-consuming, expensive, and usually takes more than ten years and billions of dollars. On the other hand, computer-aided techniques have accelerated the process of drugs. The present study aims to identify new inhibitors for the Main protease. The South African natural compound database containing 1012 compounds was docked against the Main protease. The compound with a good docking binding energy score was simulated for a total of 100 ns, and the result was compared with the control compound. The compound SANC00361 had a binding energy score of -7.30 Kcal/mol. The compound SANC00361 was found to be more stable as compared to the control compound. The identified compound SANC00361 may possess the potential to inhibit the Mpro of SARS CoV-2 and can be helpful in treating SARS CoV-2-associated infections.

Keywords: SARS CoV-2, drug target, South African natural compound database, Docking, MD simulations.

1. INTRODUCTION

Coronavirus is a member of the largest family of RNA viruses with a 27-32 kb genome size and belongs to the coronaviradeae family of viruses. Coronaviruses cause upper and lower respiratory tract infections and they even existed in the past [1]. Cough, sore throat, high fever, running nose, difficulty in breathing, sneezing, severe pneumonia, and loss of taste and smell are the different symptoms of coronavirus infections [2, 3]. These infections are viewed as enzootic but sometimes they have been able to crack the barrier between the animals and human species that manifests themselves as virulent human viruses. Without the RT-PCR

diagnosis assay, one can't distinguish the novel coronavirus from the influenza respiratory syndrome [4]. Whole-genome sequencing of SARS CoV-2 and bat coronavirus revealed that both viruses are 96% identical, indicating that bats could be the intermediate host of the novel coronavirus [5]. The genome of SARS CoV-2 encodes four structural proteins, according to a structural point of view. The spike protein helps anchor the virus to the host cell receptor, and S1 and S2 are the two subunits of the spike protein. The N terminal region of S1 acts as a receptor binding region, while the C terminus of the S2 subunit promotes the fusion activity [6]. Understanding the human coronavirus pathogenicity and the activation of the S protein is key. Matrix (M) glycoprotein interacts with S protein to promote their retention in the Golgi and endoplasmic reticulum and function as the assembly of virion [7]. The nucleotide N protein is responsible for the synthesis of nucleocapsid and has the ability to bind to the virus RNA genome. The envelope (E) protein is the smallest among all the structural proteins; it is partially incorporated into the virion envelope and is abundantly expressed in infected cells [8].

Apart from structural proteins, some non-structural proteins are present in the new coronavirus. It is well known that the viral RdRp regulates replication and transcription and requires some host factors for this process. Conformational changes in the RNA molecule can be triggered by the non-specific nucleic acid binding protein called chaperones that have a long, disordered structure. The nucleoproteins of the coronavirus act as chaperones and play an important role in template switching [9]. A recent study demonstrated that many essential enzymes such as 3CL-protease (3CL-pro), RNA-dependent RNA polymerase (RdRp), spike glycoproteins, and papain-like proteinase of both SARS and SARS CoV-2 have more than 90% sequence similarity [10]. To prevent the spread of SARS-CoV-2 Cov-2, it is important to develop vaccines and drugs, but before applying vaccines and drugs to the public, careful assessment of conceivable immune complications is required. Developing effective vaccines against SARS-CoV-2 could take several months or even years. For the new therapeutics option, the viral nsp12 and 3CL-protease (3CL-pro) have appeared as attractive drug targets. Drugs like favipiravir and remdesivir have proved effective in clinical trials [11]. In addition, the main protease inhibitors against this drug target must be helpful in the treatment of new coronavirus infections [12].

The main protease, also known as the 3CL pro, is responsible for virus replication [13]. Because it differs from human proteases, the main protease of SARS-CoV-2 is considered to be a good target for drugs [14]. The process of developing new medications is difficult, time-consuming, expensive, and high-risk; it usually takes more than ten years and billions of dollars in investments [15]. Only a small percentage of drugs get FDA approval and reach the market, despite the substantial research and financial investments [16]. The development of computer-aided techniques has accelerated the process of drugs. These approaches have shown to be helpful in cutting expenses, saving time, removing unqualified compounds in advance, and minimizing failures during the latter phase of drug development [17]. This study aims to identify new inhibitors against the SRAS CoV-2 Main protease using computational techniques.

2. MATERIALS AND METHODS

2.1. Molecular docking

2.1.1. Preparation of ligands and receptor

The Protein Data Bank was accessed to obtain the 3D structure of SARS-CoV-2 3CLPRO (PDB ID: 6LU7) [18]. Water molecules and heteroatoms were eliminated from the protein structure [19]. Next, the protein was saved in the pdbqt format for docking study [20].

Furthermore, the entire South African natural compound database with a total of 1012 compounds was retrieved from the webserver (https://sancdb.rubi.ru.ac.za/) in the SDF format [21]. All the ligands were then saved in the pdbqt format. Then energy minimization was carried out for all the ligands. The compounds in the lowest energy form were then used for docking study.

2.1.2. Active site identification

The PDB database provided the crystal structure of the SARs-CoV-2 Mpro in a complex with an N3 inhibitor (PDB ID: 6LU7). We performed a virtual screening of the phytochemical library using the N3 binding site.

2.1.3. Docking protocol

A molecular docking investigation was conducted using AutoDock Vina software. Five poses in total were produced for each phytochemical [22]. For the molecular interactions investigation, the lowest binding energy score hits against 3CLPRO were considered [23]. Based on the binding energy score, binding mechanism, and visible ligand interaction, the best hit was selected for MD simulation.

2.2. MD simulations

The dynamic behavior of the complex may be effectively understood by the application of molecular dynamics simulation [24]. The best hits, along with control, were subjected to MD simulation using Amber v20. In the Amber 20 package, the ff14SB protein force field was used. Solvation of each system was carried out, and sodium ions were incorporated to balance the system charges, allowing for sufficient solvation and neutralization of each system [25]. The steepest descent minimization algorithm was then used to minimize energy. After that, the temperature was gradually raised to 300 K until the density was balanced for two nanoseconds under weak constraints [26]. Another two nanoseconds were spent, bringing the entire system to steady-state pressure. A Langevin thermostat was employed to maintain the temperature at 300 K. Additionally, a 100 ns MD was run on the systems that were at equilibrium. Ultimately, PMEMD.CUDA was used to perform a 100 ns MD simulation of the complexes [27].

3. RESULTS AND DISCUSSIONS

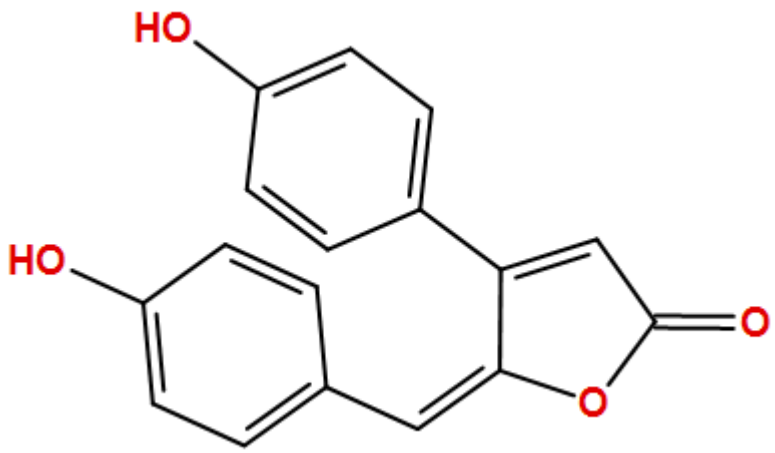
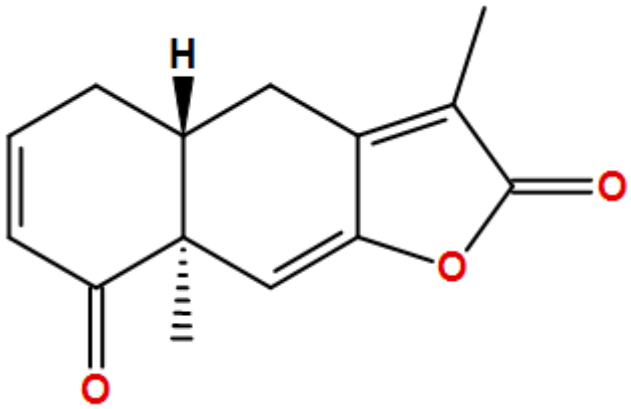
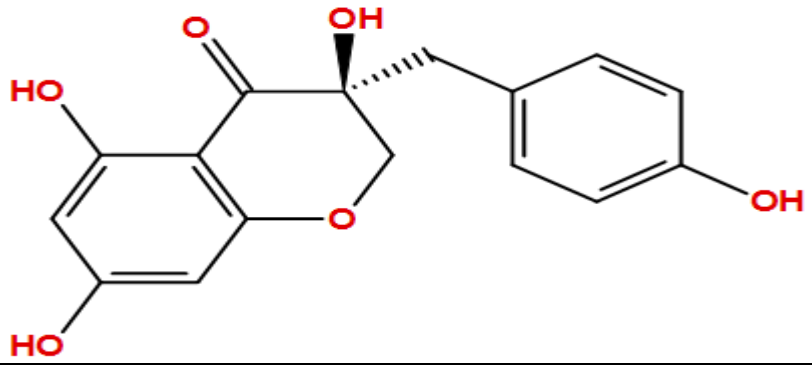
More than 80% of people around the World have used plants as their major source of medicine to address a variety of illnesses since ancient times [28]. The plants have been used to treat illnesses either by themselves or in combination with other medicinal herbs to create a polyherbal remedy. To identify the potent inhibitors against the MPro drug target of the SARS CoV-2 the South African natural compounds database was used. Docking is presently being used to investigate ligand-protein interactions at the atomic level, illuminating significant biological processes and small compounds' potential as drugs [29]. Since the ligand-binding pocket in the target is the foundation for interactions such as hydrogen bonding, pi-pi stacking, and salt bridge contacts, evaluating it is crucial to understand the binding affinities of protein ligands. The phytochemicals obtained from the South African natural compounds were docked with the Mpro using computer-aided drug design investigation to see if they may be used as SARS CoV-2 inhibitors. The results of the docking investigation showed that SANC00361 had a good binding energy score. The binding energy score of hit compound SANC00361 was -7.30 Kcal/mol. Eight hydrogen bonds were made with the CYS145, PHE140, ASN142, and GLY143 binding site residues of Mpro. The compound SANC00170 was the 2nd most potent compound with a binding energy score of -6.02 and the compound formed six hydrogen bonds with the active site residues. The residues CYS145, ASN142, GLY143, SER144, THR25, and THR26 were involved in hydrogen bonding. Among the phytochemicals the compounds

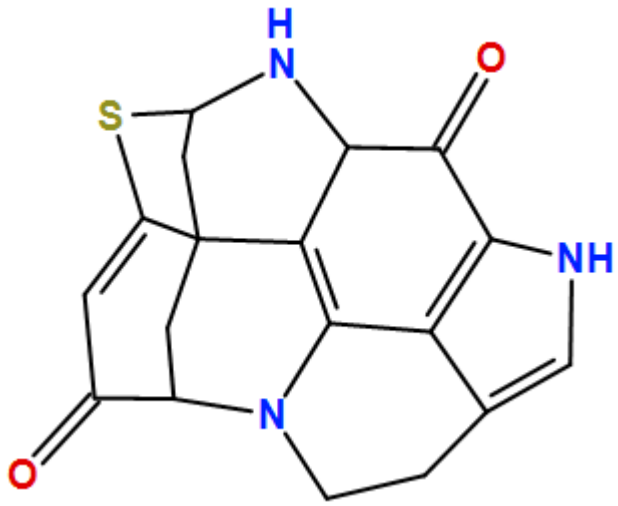
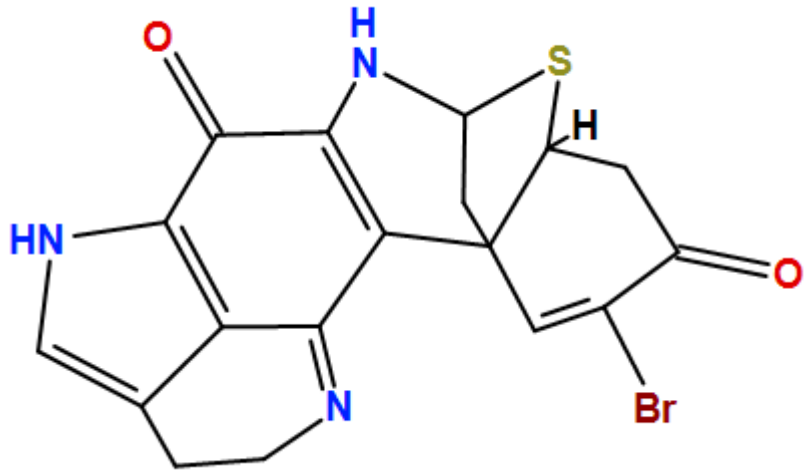
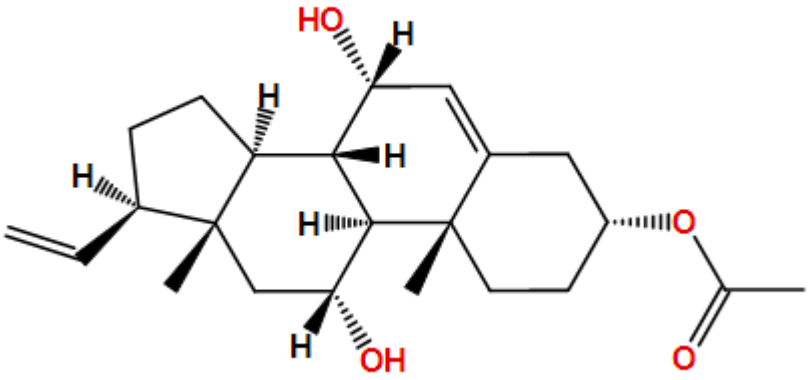
including SANC00361, SANC00170, SANC00354, SANC00117, SANC00140, SANC00313, SANC00141, SANC00130, and SANC00118 revealed more interactions as compared to the control compound. The control compound established only three hydrogen bonds with the GLU166, GLY143, and THR26 residues of the receptor. The binding energy score and interactions of the promising hits against the SARS CoV-2 are listed in Table 1. The structures of all the compounds are shown in Table 2. The 2D interactions of the most potent hits are shown in Table 3.

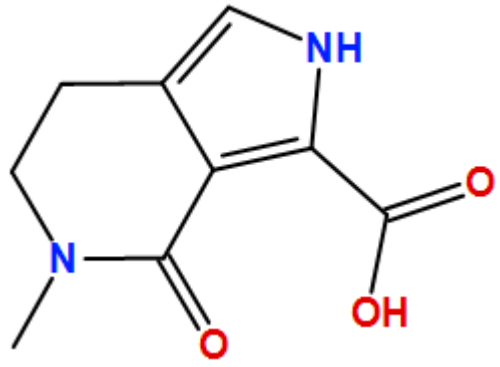
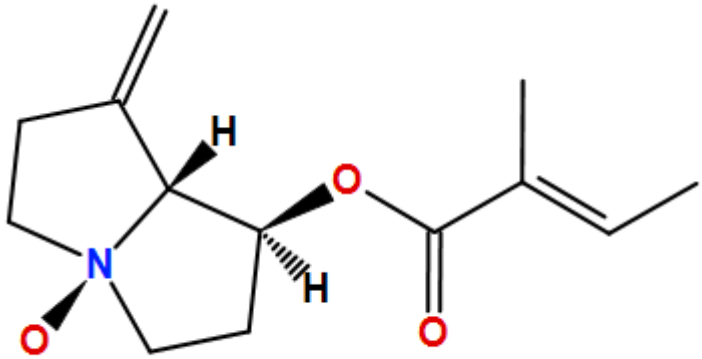
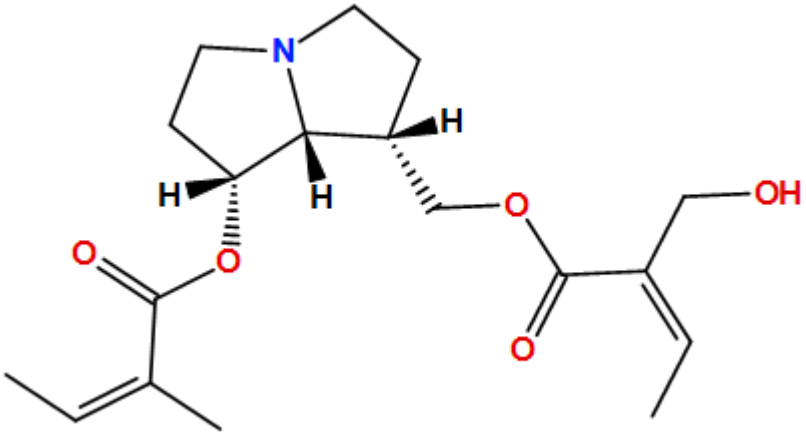
Table 1. Binding energy score and interactions of the potent compounds with the Mpro drug target in corona virus.

Compound ID	Interacting residues	Number of H-bonds	Other contacts	Binding energy
SANC00313	CYS 145, HIS 164, PHE 140, HIS 163, SER 144, CYS 145	5	1	-5.39
SANC00354	CYS 145, LEU 27, GLY 143, CYS 145, HIS 163, GLU 166,	6	Nil	-5.42
SANC00361	CYS 145, PHE 140, ASN 142, GLY 143	8	Nil	-7.30
SANC00141	SER 144, CYS 145, GLY 143	5	Nil	-5.56
SANC00140	LEU 141, SER 144, PHE 140, ASN 142, CYS 145	6	Nil	-5.72
SANC00170	CYS 145, ASN 142, GLY 143, SER 144, THR 25, THR 26	6	Nil	-6.02
SANC00130	HIS 163, MET 165, SER 144, GLY 143, CYS 145, HIS 41	5	1	-4.71
SANC00114	CYS 145, MET 49, HIS 163, GLU 166	5	Nil	-5.10
SANC00117	HIS 164, ASN 142, GLN 189, CYS 145, ASN 142, HIS 163	6	Nil	-6.15
SANC00118	HIS 164, ASN 142, GLY 143, CYS 145, ASN 142	5	Nil	-5.17
Control	GLU 166, GLY 143, THR 26	3	Nil	-6.18

Table 2. 2D structures of the most potent hits compounds and the toxicity analysis of compounds.

Compound ID	Interacting residues	Toxicity
SANC00313		No
SANC00354		No
SANC00361		No

SANC00141		No
SANC00140		No
SANC00170		No

SANC00130		No
SANC00114		No
SANC00117		No

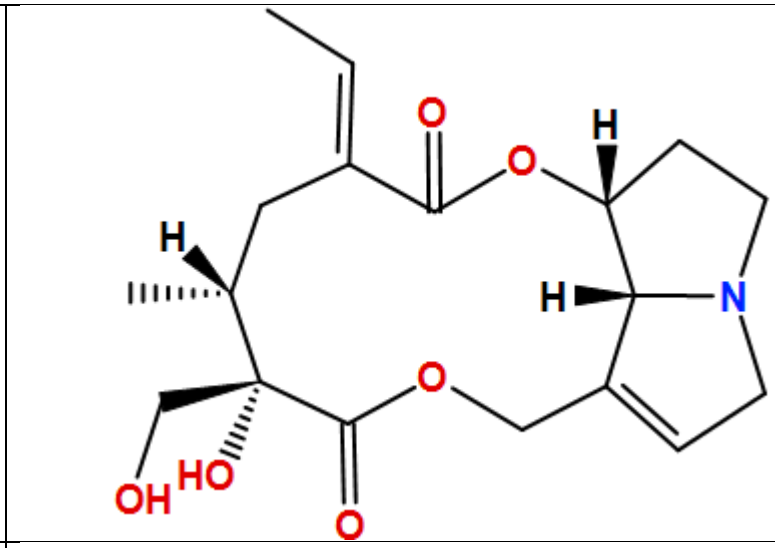
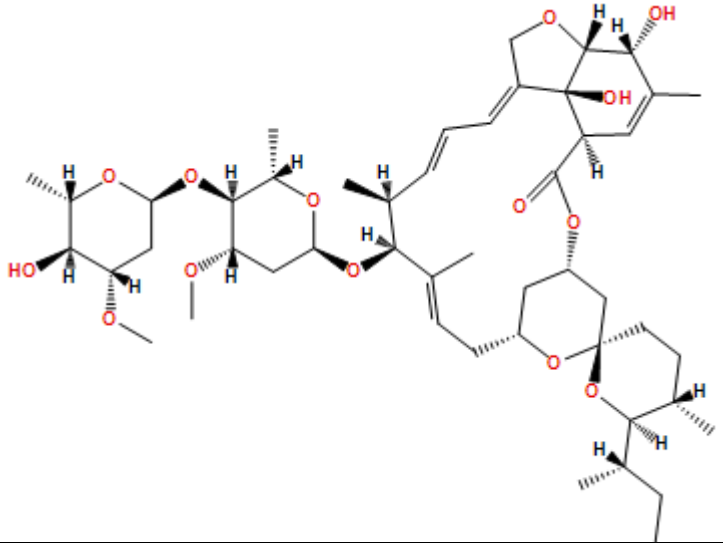
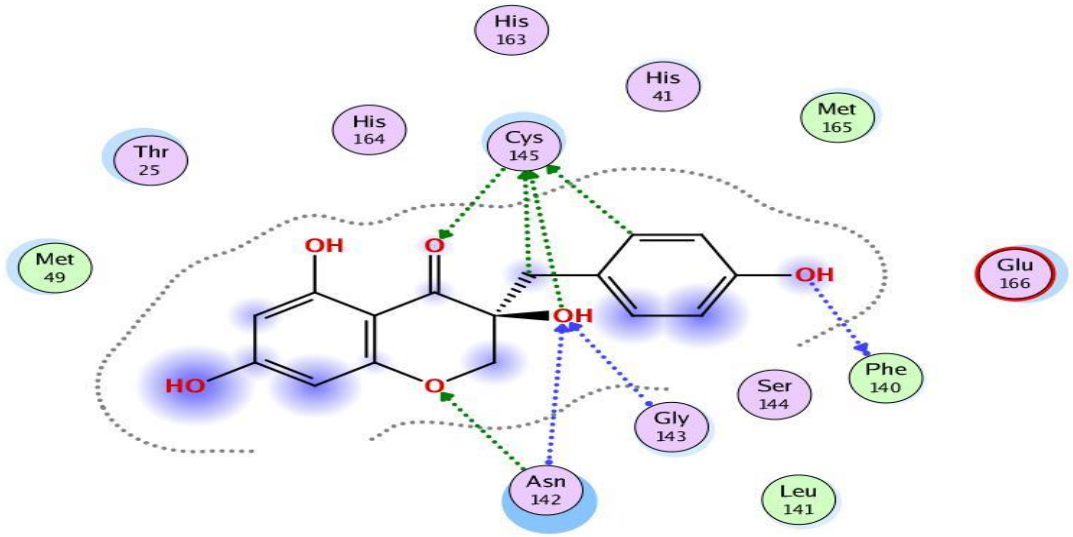
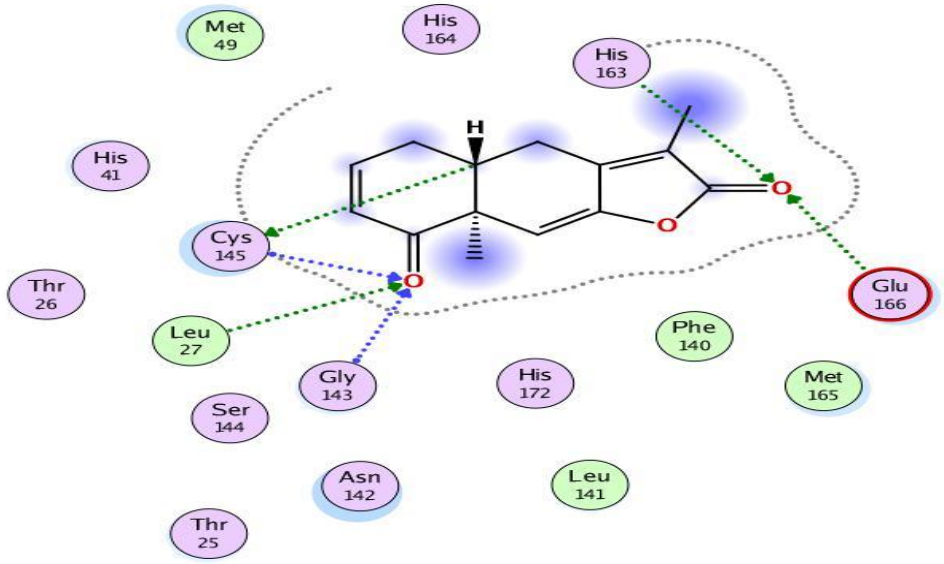
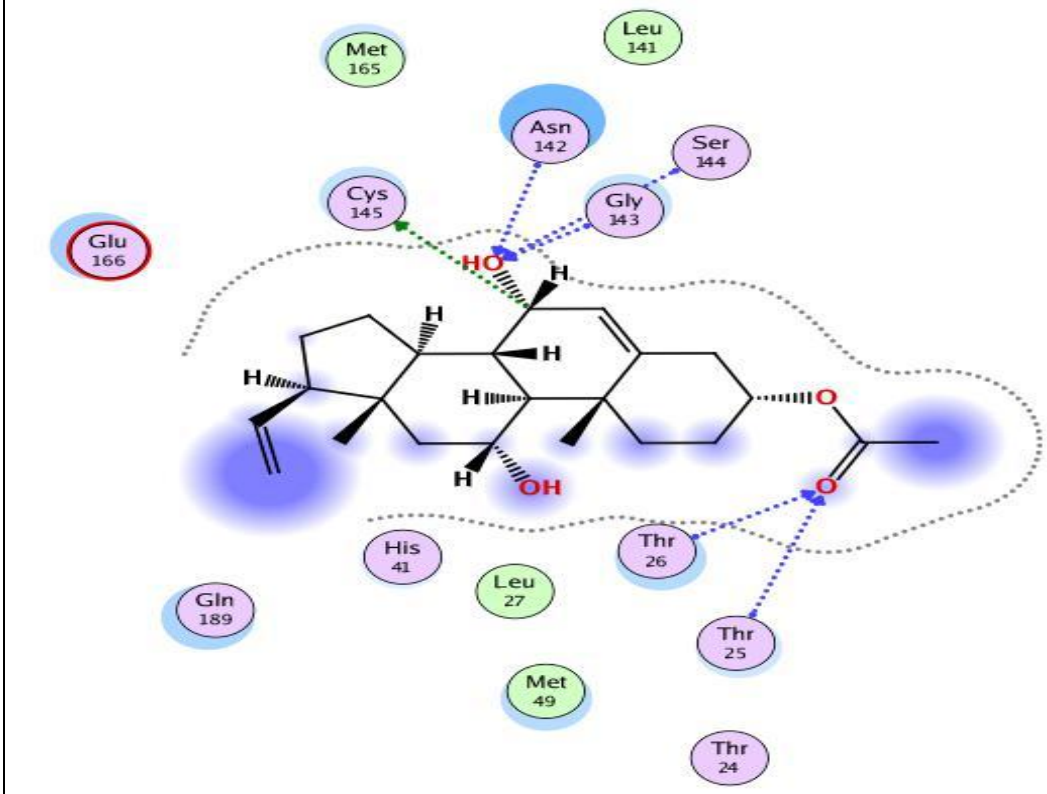
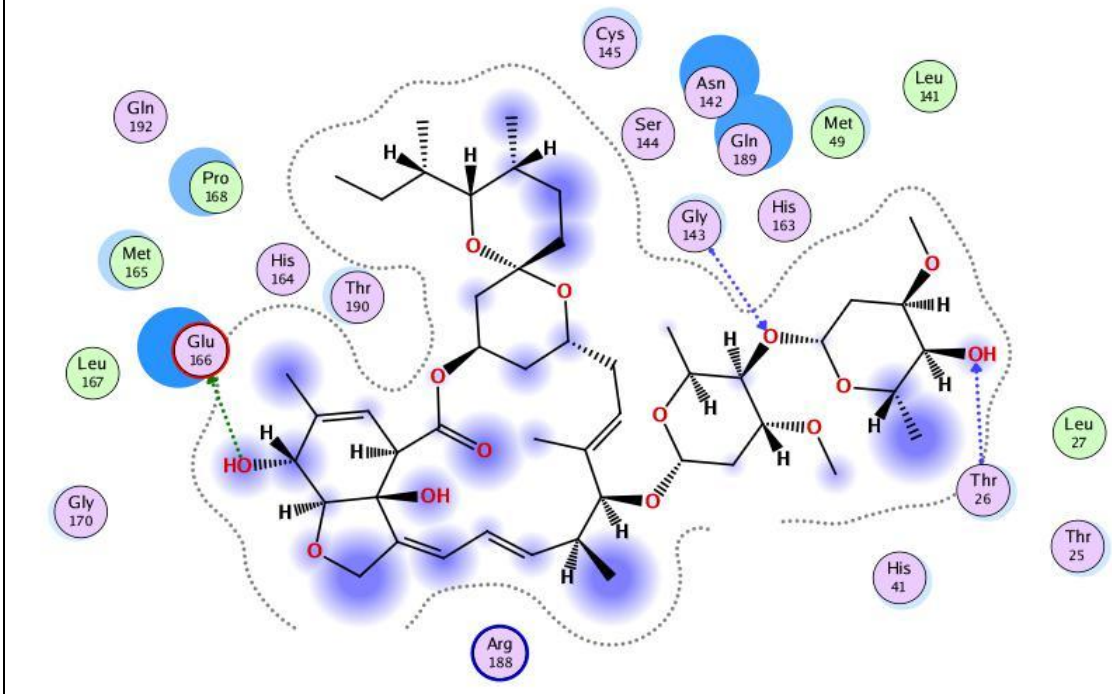
SANC00118		No
Control		No

Table 3. Two-dimensional interactions of the most potent compounds within the active site of drug target.

Compound ID	2D interaction images of complexes
SANC00361	
SANC00354	

SANC00170	
Control	

Post simulation analysis

RMSD analysis

In MD simulations, Root Mean Square Deviation (RMSD) is typically utilized to assess the stability of the receptor-ligand complex [30]. We performed RMSD calculations for both the control and SANC00361 complexes in order to shed additional light on the stability of the complexes. Figure 1 presents an illustration of the RMSD findings. The RMSD analysis of the SAN showed a reduced deviation throughout the MD simulation compared to the standard complex. In the control complex, deviations were seen from 25-45ns and 60-80ns, then the system moved to a stable conformation and indicated stability till 100 ns. In the system SAN, small deviations were seen only from 80-85 ns, and apart from these deviations, no other fluctuations were seen, and the system was found stable throughout the 100 ns MD run. The average RMSD value for the control complex was found to be 4 Å, while that of the SAN was found to be 2.5 Å.

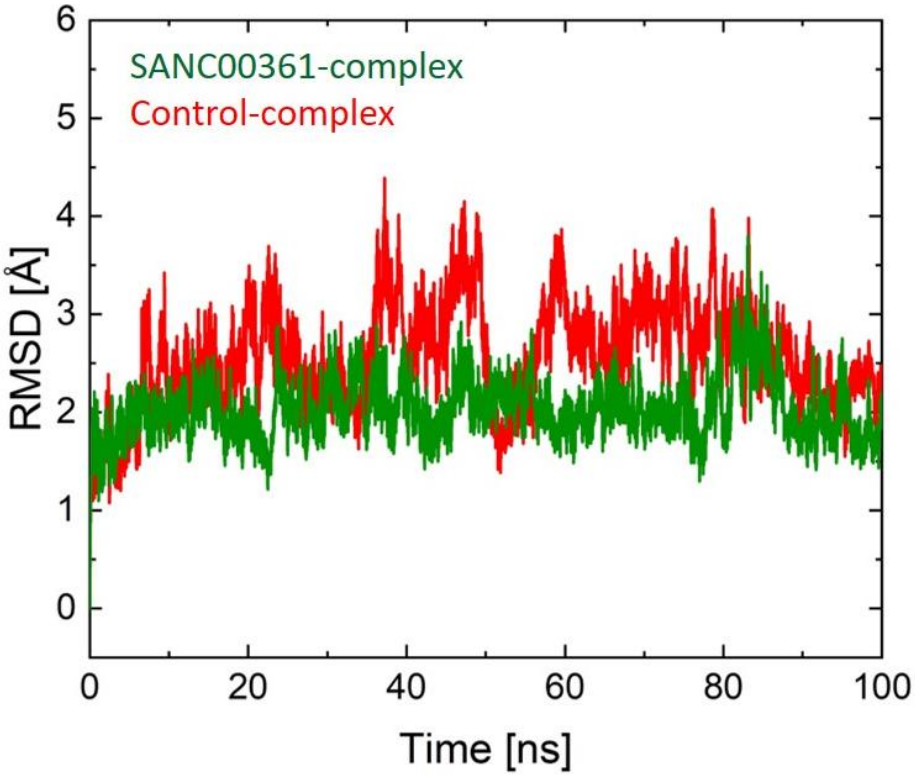


Figure 1. RMSD plot for SANC00361 (green) and the control complex (Red). A total of 100 ns MD is performed for each complex.

RMSF analysis

The RMSF of the main protease bound with SAN and control compound were examined. The binding site residues including Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, Glu166, Gln189 indicates high stability during 100 ns MD simulation. The residues from 51-60, 100-102, and 250-255 indicate fluctuations and unstable behavior as shown in Figure 2.

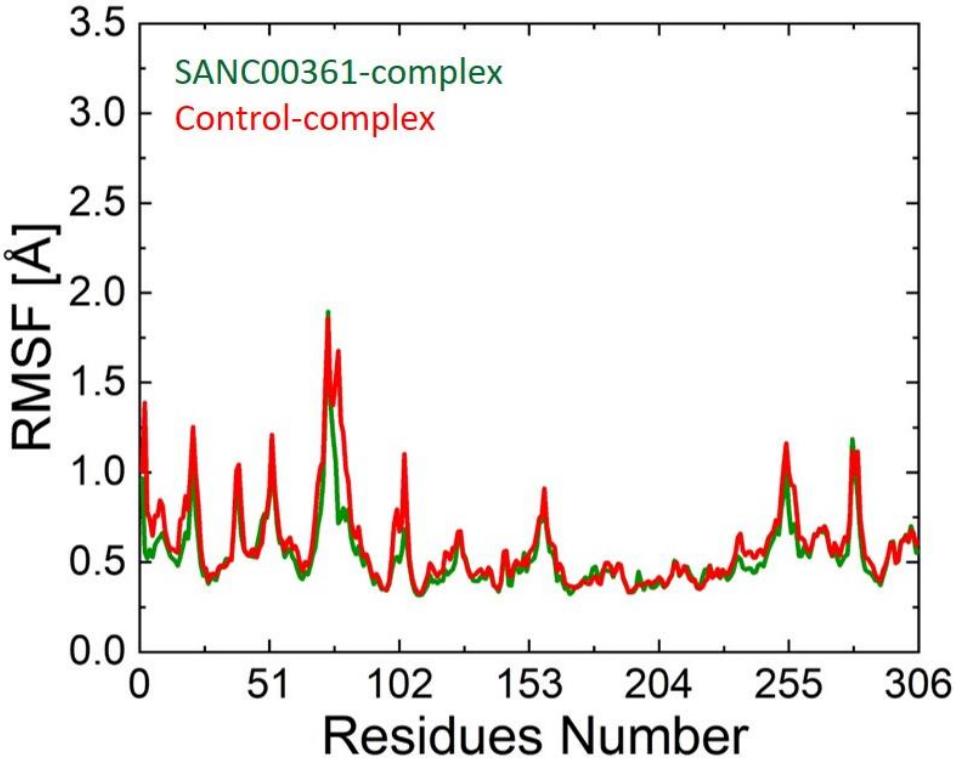


Figure 2. RMSF plot for SANC00361 (green) and the control complex (Red). A total of 306 residues of the Mpro are shown on the X-axis.

4. CONCLUSION

The COVID-19 pandemic spreads at an alarming rate worldwide. Compared to SARS and MERS, SARS-CoV-2 is more contagious and has caused more deaths and infections. Old-age people and immune-compromised patients face the greatest risk of death. Over the last 50 years, different coronaviruses have emerged which cause human and veterinary diseases. These viruses are likely to arise and evolve continuously due to their ability to replicate, mutate, and infect various species and cell types, causing human and veterinary outbreaks. In this study, new phytochemicals from the South African Natural compound database were predicted that formed strong interactions with the Mpro drug target. The compound SANC00361 was found to be more stable as compared to the control compound. The identified compound SANC00361 may possess the potential to inhibit the Mpro of SARS CoV-2 and can be helpful in treating SARS CoV-2-associated infections. It is further recommended that in-vitro and in-vivo studies be performed on the predicted hits to confirm their inhibitory potential.

Data availability

All data is available in the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Competing interests

The authors declare no competing interests.

References:

1. Reusken, C.B., et al., *Laboratory readiness and response for novel coronavirus (2019-nCoV) in expert laboratories in 30 EU/EEA countries, January 2020.* Eurosurveillance, 2020. **25**(6): p. 2000082.
2. Chan, N.C. and J.I. Weitz, *COVID-19 coagulopathy, thrombosis, and bleeding.* Blood, The Journal of the American Society of Hematology, 2020. **136**(4): p. 381-383.
3. Huang, X., et al., *Epidemiology and clinical characteristics of COVID-19.* Archives of Iranian medicine, 2020. **23**(4): p. 268-271.
4. Wu, J.T., et al., *Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China.* Nature medicine, 2020. **26**(4): p. 506-510.
5. Zhou, J., et al., *Mental health response to the COVID-19 outbreak in China.* American Journal of Psychiatry, 2020. **177**(7): p. 574-575.
6. Bosch, B.J., et al., *Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides.* Proceedings of the National Academy of Sciences, 2004. **101**(22): p. 8455-8460.
7. Neuman, M.J. and S. Powers, *Political prioritization of early childhood education during the COVID-19 pandemic: A comparative policy analysis of low-and middle-income countries.* Early Childhood Research Quarterly, 2022. **60**: p. 287-297.
8. Gupta, M.K., et al., *In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel.* Journal of Biomolecular Structure and Dynamics, 2021. **39**(7): p. 2617-2627.
9. Abeldano Zuniga, R.A., et al., *Clinical effectiveness of drugs in hospitalized patients with COVID-19: a systematic review and meta-analysis.* Therapeutic advances in respiratory disease, 2021. **15**: p. 17534666211007214.
10. Mullard, J.C., *Authenticity and recognition: Theorising antiracist becomings and allyship in the time of COVID-19 and Black Lives Matter.* Anthropological Theory, 2024. **24**(2): p. 111-135.
11. Chen, Q., et al., *Mental health care for medical staff in China during the COVID-19 outbreak.* The Lancet Psychiatry, 2020. **7**(4): p. e15-e16.
12. Gao, J., et al., *Mental health problems and social media exposure during COVID-19 outbreak.* Plos one, 2020. **15**(4): p. e0231924.
13. Samad, A., et al., *Identification of novel inhibitors for SARS-CoV-2 as therapeutic options using machine learning-based virtual screening, molecular docking and MD simulation.* Frontiers in Molecular Biosciences, 2023. **10**: p. 1060076.
14. Sharma, A., et al., *A comprehensive review on strategies for new drug discovery and enhanced productivity in research and development: recent advancements and future prospectives.* Mini-Reviews in Organic Chemistry, 2021. **18**(3): p. 361-382.
15. Niazi, S.K. and Z. Mariam, *Computer-aided drug design and drug discovery: a prospective analysis.* Pharmaceuticals, 2023. **17**(1): p. 22.
16. Gurung, A.B., et al., *An updated review of computer-aided drug design and its application to COVID-19.* BioMed research international, 2021. **2021**(1): p. 8853056.
17. Hatada, R., et al., *Fragment molecular orbital based interaction analyses on COVID-19 main protease− inhibitor N3 complex (PDB ID: 6LU7).* Journal of chemical information and modeling, 2020. **60**(7): p. 3593-3602.
18. Lafridi, H., et al., *In silico evaluation of molecular interactions between macrocyclic inhibitors with the HCV NS3 protease. Docking and identification of antiviral pharmacophore site.* Journal of Biomolecular Structure and Dynamics, 2023. **41**(6): p. 2260-2273.

19. Hatherley, R., et al., *SANCDB: a South African natural compound database.* Journal of cheminformatics, 2015. **7**: p. 1-9.
20. Ajmal, A., et al., *In Silico Prediction of New Inhibitors for Kirsten Rat Sarcoma G12D Cancer Drug Target Using Machine Learning-Based Virtual Screening, Molecular Docking, and Molecular Dynamic Simulation Approaches.* Pharmaceuticals, 2024. **17**(5): p. 551.
21. Ullah, H., et al., *Synthesis, in vitro β-glucuronidase inhibitory potential and molecular docking study of benzimidazole analogues.* Chemical Data Collections, 2023. **44**: p. 100996.
22. Alotaibi, B.S., et al., *New drug target identification in Vibrio vulnificus by subtractive genome analysis and their inhibitors through molecular docking and molecular dynamics simulations.* Heliyon, 2023. **9**(7).
23. Wadood, A., et al., *Machine learning-based virtual screening for STAT3 anticancer drug target.* Current Pharmaceutical Design, 2022. **28**(36): p. 3023-3032.
24. Ajmal, A., et al., *Prospective virtual screening combined with bio-molecular simulation enabled identification of new inhibitors for the KRAS drug target.* BMC chemistry, 2024. **18**(1): p. 57.
25. Ajmal, A., et al., *In silico design of peptide inhibitors for Dengue virus to treat Dengue virus-associated infections.* Scientific Reports, 2024. **14**(1): p. 13130.
26. Alagu Lakshmi, S., et al., *Ethnomedicines of Indian origin for combating COVID-19 infection by hampering the viral replication: using structure-based drug discovery approach.* Journal of Biomolecular Structure and Dynamics, 2021. **39**(13): p. 4594-4609.
27. Subramaniyan, V., et al., *Molecular docking and ADME properties of bioactive molecules against human acid-beta-glucosidase enzyme, cause of Gaucher's disease.* In Silico Pharmacology, 2018. **6**: p. 1-11.
28. Nasir, A., et al., *Identification of novel and potential inhibitors against the dengue virus NS2B/NS3 protease using virtual screening and biomolecular simulations.* International Journal of Biological Macromolecules, 2024: p. 132855.