

Conservation analysis and screening of Ayurvedic compounds against SARS-CoV-2 Spike protein

Zarrin Basharat (✉ zarrin.iiui@gmail.com)

Jamil-ur-Rahman center for Genome Research, Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi-75270, Karachi, Pakistan

Muhammad Jahanzaib

Jamil-ur-Rahman center for Genome Research, Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi-75270, Karachi, Pakistan.

Noor Rahman

HEJ Research Institute of Chemistry, International Center for Chemical and Biological Sciences, University of Karachi-75270, Karachi, Pakistan.

Ishtiaq Ahmad Khan

Jamil-ur-Rahman center for Genome Research, Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi-75270, Karachi, Pakistan.

Azra Yasmin

Microbiology and Biotechnology Research Lab, Department of Biotechnology, Fatima Jinnah Women University, The Mall, Rawalpindi 46000, Pakistan.

Research Article

Keywords: Coronavirus, Spike protein, COVID-19, Drug design, Post-translational modification

Posted Date: August 24th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-63923/v1>

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

[Read Full License](#)

Abstract

Recent infections caused by the novel coronavirus (SARS-CoV-2) have led to global panic and mortality. Here, we analyzed the spike (S) protein of this virus using bioinformatics tools. We aimed to determine relative changes among different coronavirus species over the past two decades and to understand the conservation of the S-protein. Representative sequences of coronaviruses were collected from humans and other animals between 2000 and 2020. Evolutionary analyses found that the S-protein did not evolve overnight, but rather continuously over time. Post-translational modification (PTM) analysis using online tools and virtual screening of S-protein against a phytochemical database of Ayurvedic medicinal compounds (n = 2103) identified the S-protein inhibitors. Among these, top ranked were Gingerol (IUPAC name: 4'-Me ether, 3,5-di-Ac 3,5-di-Gingerdiols), 1-(5-Butyltetrahydro-2-furanyl)-2-hexacosanone and Ginsenoyne N ginseng that stimulates Caspase-3, Caspase-8, and the immune system. Gingerol is found in the fresh ginger and has reputation of being a potent antiviral. These compounds might prove useful to design drugs against COVID-19.

Introduction

The number of global infections due to the novel coronavirus (SARS-CoV-2) reached over 22,262,946 infections with over 784,107 deaths on August 20, 2020. Stats are continuously changing ("Coronavirus (COVID-19) - Google News," n.d.). The virus was initially transmitted to humans from other animals and then via person-to-person contact. It is not airborne, but aerosols and contact base transmission are the usual modes of transmission. The mechanism of the rapid transmission and why the same viral strain causes death in some persons and mild symptoms in others remain obscure. Coronavirus belongs to the subfamily *Coronavirinae* along with *Torovirinae*, and is grouped into the *Coronaviridae* family (Belouzard, Millet, Licitra, & Whittaker, 2012). Coronavirus (SARS-CoV-2) infections have recurred from time to time across various geographic regions. It is responsible for almost 30% of all respiratory infections in humans and other animals and causes great economic loss. Alpha and beta types of CoV mostly target human hosts, and other serological genera include beta, gamma and delta types (<https://talk.ictvonline.org/taxonomy/>). The RNA-based genome of CoV is the largest among known RNA viruses and it has high zoonotic potential for recombination and infecting new animal hosts including humans (Lai, Shih, Ko, Tang, & Hsueh, 2020). Viral sequences are under continuous stress to break through host barriers. High rates of interaction between humans and other animals provide a considerable trial and error experimental environment for viral pathogens to cross from one host to another, resulting in the emergence of new pathogens (Dolja & Koonin, 2018).

The evolutionary basis of CoV has made devastating comebacks possible; for example, the severe acute respiratory syndrome coronavirus (SARS-CoV) outbreak of 2002 - 2003 (Drosten et al., 2003; Holmes & Rambaut, 2004) that infected over 8,000 people with varying morbidity and mortality rates (Guarner, 2020). A new strain that emerged in the Middle East during 2012 (Middle East respiratory syndrome coronavirus; MERS-COV) that killed over 780 people were thought to arise due to Arabs interacting with dromedary camel products and by-products (Reusken et al., 2013). In 2013, a porcine epidemic diarrhea

coronavirus (PEDV) with a 100% fatality rate that decimated 10% of the total pig population in the USA (Mole, 2013; Chen et al., 2014). The virus transformed and emerged within a decade of the PEDV epidemic in the form of COVID-19 with pneumonia-like symptoms. Such situations are becoming more problematic for authorities to manage. The epicenter of this virus is thought to be an animal market in Wuhan (Chang, Lu, Chen, Jin, & Yang, 2012)

SARS-CoV-2 has an ~29 kb genome, a GC content of 38%, and RNA that encodes various proteins including the structural spike (S), membrane (M), and envelope (E) proteins (Enjuanes, Almazán, Sola, & Zuñiga, 2006; Fehr & Perlman, 2015). The CoV S-protein is a glycoprotein that expresses pathogenic characteristics in hosts by interacting with various cellular receptors and invading cells (Kwak, Song, Lee, & Schiefelbein, 2015). S-proteins vary according to viral types and range from 1,160 to 1,400 amino acids that facilitate viral entry into cells by interaction with various receptors (Belouzard et al., 2012). S-proteins consist of an N-terminal domain, an S1 receptor-binding region, and a C-terminal S2 binding domain. The latter domain assists viral fusion with host receptors on cell membranes (Bosch, van der Zee, de Haan, & Rottier, 2003; Millet & Whittaker, 2015). The protein is cleaved during viral maturation and exocytosis in some viruses, which causes various distinctions among CoV isolates. S-protein is class 1 fusion protein with an α -helical structure that confers the features of similar coiled-coils such as influenza hemagglutinin protein HA (Bosch et al., 2003; Xu et al., 2004).

Receptor-mediated intracellular entry varies according to viral strains. Aminopeptidase N (APN) receptors are used by various α -CoV, HCoV-NL63 and SARS and they interact with angiotensin-converting enzyme 2 (ACE2) to mediate cellular entry, whereas the novel dipeptidyl-peptidase 4 (DPP4) confers invasive capability upon MERS-COV (Fehr & Perlman, 2015). Here, we analyzed the evolutionary changes in S-protein that allow the virus to penetrate host barriers and cause infection using specie based analysis. S-protein modifications and other mutations lead to protein evolution that helps viruses to invade new species. Viruses then undergo further modifications over several years that lead to the entry into and pathogenesis of more new hosts. In this study, we carried out an *in silico* based virtual screening based approach to find potential drug compound using natural compounds. Our findings will increase understanding of the mechanisms of CoV infection and possibly lead to therapeutic interventions against it. We also screened phytochemical inhibitors in a database of Ayurvedic compounds to identify agents that might act against S-protein.

Methods

2.1. Conservation analysis

Homologous sequences of S-protein and conservation profiles were analyzed using the protein sequences of major human and animal viral species of S-protein, obtained from the Viral Pathogen Resource (ViPR) <https://www.viprbrc.org/brc/home.spg?decorator=vipr>. Sequences collected annually between 2000 and 2020 were filtered to remove short sequences. Thereafter, sequences from several species were selected regarding year and aligned pairwise using the ClustalW function of MEGA-X

software (Kumar, Stecher, Li, Knyaz, & Tamura, 2018). Phylogenetic analyses were conducted using the maximum likelihood method. Accession numbers and aligned sequence data are provided in Supplementary File 1. Structural conservation analysis was done using ConSurf (Ashkenazy, Erez, Martz, Pupko, & Ben-Tal, 2010) and multiple sequences were aligned using ClustalW. Homologues were collected from UNIREF90 using the HMMER algorithm (E-value, 0.0001; maximum identity, 95; minimum identity, 35; iterations, 1). Conservation scores were calculated using maximum likelihood and Bayesian methods.

2.2. S-protein retrieval

SARS-COV2 S-protein structure PDB-ID RVYB was retrieved from protein data bank (PDB) having a resolution of 3.20 Å, initial preprocessing was done using AutoDock Tools utility (Morris et al., 2009). Pocket prediction of all chain in open conformation and single chain was also performed using DoGSiteScorer (Volkamer, Kuhn, Rippmann, & Rarey, 2012). Both chains combinely forming pocket and individual chain pocket formation analysis was performed to find difference in ligand accommodating pockets offered. Comparison of pockets was also performed between SARS-CoV-1, SARS-CoV-2, and MERS-COV structures were retrieved from RCSB having PDB-ID 5X58, RVYB, and 5X5C respectively.

2.3. Ayurvedic medicine database screening against target S-protein

About 2,103 compound were retrieved from Ayurvedic medicine database available at (<http://www.way2drug.com/>). Compound were screened for their drug like features using reference and were tested for Lipinski's thumb rule of five (Lipinski, 2004). Molecular Operating Environment (MOE v2019.0102) was used for docking and visualization. Reference scale was also produced in order to compare drug efficacy, standard were Ritonavir and Remdesivir. Docking of was carried out using parameter: placement = triangle matcher, rescoring 142 1 = London dG, refinement = forcefield, rescoring 2 = affinity dG. Out of 21,103 compounds, around 500 were retained for docking after filtering based on various criteria like Lipinski's drug-likeness etc. and top ranked three were selected for further analysis. Selection was based on S-scores and root-mean-square deviation (RMSD) values. MOE uses built-in function that by default calculates binding energy (S value), which show affinity of ligand with receptor. While in case for RMSD the scoring functions uses reference to compare conformation against docked. Top selected compounds were having higher S-values and lower RMSD score than the reference which could be developed as a potential inhibitor for S-protein (Tahir ul Qamar et al., 2016). ADMET analysis was performed using AdmetSAR 2.0 (<http://lmmmd.ecust.edu.cn/admetsar2/>).

Results And Discussion

3.1. Conservation and phylogenetic evolution

The present study we analyzed S-protein of SARS-CoV-2, which recently emerged as a new threat and rapidly became a global pandemic. The GC content of SARS-CoV-2 S gene was 37.39%. The ideal range of GC content is between 30% and 70%. Any peaks outside this range will adversely affect transcriptional and translational efficiency. We analyzed a total of 114 amino acid sequences to infer annual evolution. To suggest evolutionary relationships with various other species while considering the possibility of recombination, all species were considered in this analysis using the maximum likelihood method and amino acid substitutions were corrected using the JTT substitution matrix-based model (Jones *et al.*, 1992). We estimated pairwise distance using a JTT model applied to Neighbor-Join and BioNJ with an automated heuristic search and final tree with the highest log -100343.12 (*Fig. 1*). Our findings indicated that SARS-CoV-2 evolved from bats and that it has continuously evolved. We also postulate that other recombinations are underway that might result in a new species.

3.2. SARS-CoV-2 S-protein potential drug pockets

Pocket analysis was performed to evaluate potential drug binding sites and volume, for calculation of volume protein co-ordinates file was submitted into online web server DoGSiteScorer (Volkamer *et al.*, 2012). Server predicted various pockets with a series of pockets data having various parameters of druggability. Druggability score ranges from 0–1 pocket, pocket higher score is more potential with respect to others. Pocket with highest drug score was 0.85 having volume of 895.01 Å³ for single chain while multiple chain were forming a pocket with a volume of 4155.19 Å³ (*Fig. 2*). The combine pocket has similar regions involved in pocket formation with minor variation for SARS-CoV1 and MERS-CoV sharing similarity in potential drug pocket with highest score. While other SARS-CoV-1 have different binding pocket with highest score as compared to SARS-CoV-2 and MERS-CoV. SARS-CoV-1 forms a channel pocket having a volume of 10587.05 Å³ being the largest docking pocket structurally it very much like multiple Y's connected with S shaped tunnels depicted in *Fig. 3*. This difference in ratio of pocket volume and surface volume clearly indicate that interaction would be clearly different.

3.3. Virtual screening based on pharmacophore using Ayurvedic database

Virtual screening approach give advantage of screening through thousands of compound with investment of less time, and helps in discovery of novel drugs compounds. Using best hits in reference drugs make it more powerful as druggable compounds could be compared with already known active drugs. Like in this case considered compounds with one hydrophobic, one aromatic, one H-accepter, and one H-donor were selected, and were given preference over others. Compound were further screened through Lipinski's thumb rule of five and number was further brought down by just selecting compound fulfilling at least three rules. Lipinski's rule sets five parameters for compound to be a drug like compound which include log p-valued <5, molecular weight < 500 Daltons, H-bond acceptor and H-bond donor <5. Drug should be fulfilling these properties are important for human pharmacokinetics (Lipinski, 2004; Yang

et al., 2018). Best compound hits were ran through force-field MMFF94x, and Gradient: 0.05 for energy minimization running using MOE minimizing algorithm.

3.4. Molecular Docking

Molecular docking has great importance for drug discovery and widely accepted. Current studies shows that ACE2 binding sites show higher conservation, about 8 out of 14 are highly conserved. While remaining 6 positions are (semi) conservatively substituted: R426_{SARS-CoV}|N439_{SARS-CoV-2}, Y442_{SARS-CoV}|L455_{SARS-CoV-2}, L472_{SARS-CoV}|F486_{SARS-CoV-2}, N479_{SARS-CoV}|Q493_{SARS-CoV-2}, Y484_{SARS-CoV}|Q498_{SARS-CoV-2}, and T487_{SARS-CoV}|N501_{SARS-CoV-2}. This conservation clearly explains why SARS-CoV-2 have similar binding affinities like previous SARS-CoV with ACE2 receptor (Walls et al., 2020). Ayurvedic library was screened using receptor centric approach and entire surface was analyzed for possible binding sites and pockets were kept in focus of S-protein. Various compound showing interaction *Fig. 4* were studied with various parameters, further ranking was carried using S-value. Compound showing good interaction with higher S-values and lower RMSD were also compared with standard compound and their S-values were compared against each other *Table. 1*. Binding interactions of S-protein best hits were evaluated using MOE Ligplot algorithm. Compounds were showing hits near binding sites S-protein in SARS-CoV-2, compound with the highest S-value showed binding with ASN331, THR333, SER530, and GLN580 considered as a good inhibitory compound. Three compounds were selected out of all hits that were showing good features for ADMET analysis and their drugability profile was generated.

Table 1. Top ranked compounds that could serve as inhibitors of SARS-CoV-2 from Ayurvedic compounds, along with reference compounds.

Serial no.	Compound	Compound Name	S-value
1	Compound 1	4'-Me ether, 3,5-di-Ac 3,5-di-Gingerdiols	-6.8
2	Compound 2	1-(5-Butyltetrahydro-2-furanyl)-2-hexacosanone	-6.6
3	Compound 3	Ginsenoyne N	-6.5
4	Standard 1	Remdesivir	-6.4
5	Standard 2	Ritonavir	-5.4

3.5. ADMET validation of top hit compounds

AdmetSAR was used for further evaluation of top hits and standards for ADME and toxicity analysis. Along with reference compounds, top compounds were found to be within applicability domain

(consisting of six physicochemical properties: molecular weight, $\log P$, number of atoms, number of rings, H-bond acceptors, and H-bond donors) and negative for Ames mutagenicity.

Various administration and distribution matters directly influence metabolism and elimination in overall life cycle and effectiveness of drug. Oral bioavailability and Blood brain barrier (BBB) is endothelial cell barrier that prevent drug from entering the brain and is most important factor influencing a drug to be distributed (Alavijeh, Chishty, Qaiser, & Palmer, 2005; Thomas et al., 2006). Drug-like compound were positive for their blood brain barrier (BBB) penetration, bioavailability, HIP (Human intestinal preparation), ROCT (Renal Organic cation transporter), CaCO₂ permeability and P-Glycoprotein substrate. Compounds with top hits were non-toxic, non-inhibitor of CYP enzymes, non-carcinogenic, and non-mutagenic. Compounds selected could serve as novel drug compounds potentially active against S-protein of SARS-COV-2. The three compound 4'-Me ether, 3,5-di-Ac 3,5-di-Gingerdiols, 1-(5-Butyltetrahydro-2-furanyl)-2-hexacosanone, and Ginsenoyne N were all in acceptable range of ADMET parameters (*Table 2*).

Table 2. ADMET profiling of the inhibitors of SARS-CoV-2 S-protein.

Parameter	compound1	compound2	compound3	Remdesivir	Ritonavir
Androgen receptor binding	+	-	+	+	+
Aromatase binding	-	-	+	+	+
Blood Brain Barrier	+	+	+	+	-
BRCP inhibitor	-	-	-	-	-
BSEP inhibitor	+	+	+	+	+
Caco-2	+	+	-	-	-
Carcinogenicity (binary)	-	-	-	-	-
CYP1A2 inhibition	+	-	+	-	-
CYP2C19 inhibition	+	-	-	-	+
CYP2C9 inhibition	-	-	-	-	-
CYP2C9 substrate	-	-	-	-	+
CYP2D6 inhibition	-	-	-	-	-
CYP2D6 substrate	-	-	-	-	-
CYP3A4 inhibition	-	-	-	-	+
CYP3A4 substrate	+	-	+	+	+
CYP inhibitory promiscuity	-	-	-	-	+
Eye corrosion	-	+	-	-	-
Eye irritation	-	+	-	-	-
Estrogen receptor binding	+	-	+	+	+
Glucocorticoid receptor binding	+	-	+	+	+
Hepatotoxicity	-	-	-	+	+
Human ether-a-go-go inhibition	+	-	+	-	+
Human Intestinal Absorption	+	+	+	+	+
Human oral bioavailability	-	-	-	-	-
MATE1 inhibitor	-	-	-	-	-
micronuclear Acute Oral Toxicity (kg/mol)					
	2.30	2.00	3.63	3.42	2.28
OATP1B1 inhibitor	+	+	+	+	+
OATP1B3 inhibitor	+	+	+	+	+
OATP2B1 inhibitor	-	-	-	-	+
OCT1 inhibitor	-	-	-	-	-
OCT2 inhibitor	-	-	-	-	-
P-glycoprotein inhibitor	+	-	+	+	+
P-glycoprotein substrate	+	-	+	+	+
PPAR gamma	-	-	-	+	+
Plasma protein binding (%)					
	0.85	0.77	1.26	1.18	1.11
Subcellular localzation					
	Mitochondria	Plasma membrane	Mitochondria	Lysosomes	Lysosomes

Thyroid receptor binding	-	+	+	+	+	
UGT catalyzed	-	-	-	-	-	
Water solubility(log S)		-4.75	-1.61	-4.14	-3.47	-3.22

3.6. Selection of drug based on interaction of protein ligand

All library showed binding with protein each and every compound showing binding were not going to act as drug. So selection of potential drug strength of binding was analyzed for all interacting compounds. Strength of binding is directly proportional to the S-score analyzed and compound were organized according to the highest to lowest. This approach was combined with reference based approach using well know affective as standards drugs for comparing interactions of our potential drug like compounds. Our top compound 4'-Me ether, 3,5-di-Ac 3,5-di-Gingerdiols was showing S-value of 6.8 while Remdesivir and Ritonavir were showing S-value 6.4 and 5.4 respectively. Top compound surface analysis using surface based analysis showed that the compound was quite embedded in the binding with good stability *Figure 5*.

4'-Me ether, 3,5-di-Ac 3,5-di-Gingerdiols or simply Gingerol, belongs to phenol phytochemical that is commonly found in fresh ginger and activate spice receptors. Gingerol belong to the family of capsaicin and piperine family of compound, which are alkaloids having different bioactivity pathway (Beltrán et al., 2013). Ginger has been used from ancient times in various Chinese, Ayurvedic and Tibb-Unani herbal medicines. It has been nonspecifically used for treatment of various unrelated condition including arthritis, inflammation, stomach problems like diarrhea, fever and parasite infections, and Gingerol has been specifically associated with the effectiveness (Ali, Blunden, Tanira, Toxicology, & 2008, n.d.). Antiviral activity has also been associated with Gingerol with various different pathway against SARS-CoV-1, Human Immuno- deficiency virus HIV, Ebola virus, and Influenza virus A (Mbadiko et al., 2020). All previous studies suggest that Gingerol could be potentially an effective drug option while treating CoVID-19 patients as safer option as it is least toxic substance that has been used for centuries by various medicinal culture throughout the world.

Conclusion

In this study Ayurvedic medicine database was virtually screened for potential drug like compounds that could inhibit SARS = CoV-2 binding with host receptor. After screening library of more than 2000 compounds, 500 were chosen for docking based on various parameters. Three compounds were further analyzed for ADMET, as they ranked as top inhibitors against SARS-CoV-2 S-protein having higher interactions in comparison to standard drugs. According to our findings, 4'-Me ether, 3,5-di-Ac 3,5-di-Gingerdiols is a potential drug for SARS-CoV-2, with S-protein as target. It needs to be evaluated in vivo, in vitro and added in drug development pipeline in near future.

Declarations

Acknowledgements

The authors are thankful to Alexandra Elbakyan for providing literature support. Furthermore, we are grateful to all those scientists, research labs and universities who made computational biology tools, data and web servers freely available for public use.

Funding

This study did not receive any particular funding.

Conflict of Interest

The authors declare that they have nothing to disclose which might pose a potential conflict of interest.

Author contributions

ZB and MJ conceived and designed the experiments. MJ, NR and ZB conducted formal analysis, compiled information and drafted manuscript. AY and IAK provided resources for study and supervised the project.

References

1. Alavijeh, M. S., Chishty, M., Qaiser, M. Z., & Palmer, A. M. (2005). Drug metabolism and pharmacokinetics, the blood-brain barrier, and central nervous system drug discovery. *NeuroRx*, 2(4), 554–571. <https://doi.org/10.1602/neurorx.2.4.554>
2. Ali, B., Blunden, G., Tanira, M., Toxicology, A. N.-F. and chemical, & 2008, (n.d.). Some phytochemical, pharmacological and toxicological properties of ginger (*Zingiber officinale* Roscoe): a review of recent research. *Elsevier*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0278691507004243>
3. Ashkenazy, H., Erez, E., Martz, E., Pupko, T., & Ben-Tal, N. (2010). ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research*, 38(SUPPL. 2). <https://doi.org/10.1093/nar/gkq399>
4. Belouzard, S., Millet, J. K., Licitra, B. N., & Whittaker, G. R. (2012). Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses*, Vol. 4, pp. 1011–1033. <https://doi.org/10.3390/v4061011>
5. Beltrán, L. R., Dawid, C., Beltrán, M., Gisselmann, G., Degenhardt, K., Mathie, K., ... Hatt, H. (2013). The pungent substances piperine, capsaicin, 6-gingerol and polygodial inhibit the human two-pore

- domain potassium channels TASK-1, TASK-3 and TRESK. *Frontiers in Pharmacology*, 4 NOV. <https://doi.org/10.3389/fphar.2013.00141>
6. Bosch, B. J., van der Zee, R., de Haan, C. A. M., & Rottier, P. J. M. (2003). The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *Journal of Virology*. <https://doi.org/10.1128/jvi.77.16.8801-8811.2003>
 7. Chang, Z., Lu, S., Chen, L., Jin, Q., & Yang, J. (2012). Causative Species and Serotypes of Shigellosis in Mainland China: Systematic Review and Meta-Analysis. *PLoS ONE*, 7(12). <https://doi.org/10.1371/journal.pone.0052515>
 8. Chen, Q., Li, G., Stasko, J., Thomas, J. T., Stensland, W. R., Pillatzki, A. E., ... Zhang, J. (2014). Isolation and characterization of porcine epidemic diarrhea viruses associated with the 2013 disease outbreak among swine in the united states. *Journal of Clinical Microbiology*. <https://doi.org/10.1128/JCM.02820-13>
 9. Coronavirus (COVID-19) - Google News. (n.d.). Retrieved August 20, 2020, from <https://news.google.com/covid19/map?hl=en-PK&gl=PK&ceid=PK%3Aen>
 10. Dolja, V. V., & Koonin, E. V. (2018). Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Research*. <https://doi.org/10.1016/j.virusres.2017.10.020>
 11. Drosten, C., Günther, S., Preiser, W., Van der Werf, S., Brodt, H. R., Becker, S., ... Doerr, H. W. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *New England Journal of Medicine*, 348(20), 1967–1976. <https://doi.org/10.1056/NEJMoa030747>
 12. Enjuanes, L., Almazán, F., Sola, I., & Zuñiga, S. (2006). Biochemical Aspects of Coronavirus Replication and Virus-Host Interaction. *Annual Review of Microbiology*. <https://doi.org/10.1146/annurev.micro.60.080805.142157>
 13. Fehr, A. R., & Perlman, S. (2015). Coronaviruses: An overview of their replication and pathogenesis. In *Coronaviruses: Methods and Protocols*. https://doi.org/10.1007/978-1-4939-2438-7_1
 14. Guarner, J. (2020, March 9). Three Emerging Coronaviruses in Two Decades: The Story of SARS, MERS, and Now COVID-19. *American Journal of Clinical Pathology*, Vol. 153, pp. 420–421. <https://doi.org/10.1093/ajcp/aqaa029>
 15. Holmes, E. G., & Rambaut, A. (2004). Viral evolution and the emergence of SARS coronavirus. *Philosophical Transactions of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rstb.2004.1478>
 16. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*. <https://doi.org/10.1093/molbev/msy096>
 17. Kwak, S.-H., Song, S.-K., Lee, M. M., & Schiefelbein, J. (2015). ANGUSTIFOLIA mediates one of the multiple SCRAMBLED signaling pathways regulating cell growth pattern in *Arabidopsis thaliana*. *Biochemical and Biophysical Research Communications*, 465(3), 587–593. <https://doi.org/10.1016/J.BBRC.2015.08.067>

18. Lai, C. C., Shih, T. P., Ko, W. C., Tang, H. J., & Hsueh, P. R. (2020, March 1). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*, Vol. 55, p. 105924. <https://doi.org/10.1016/j.ijantimicag.2020.105924>
19. Lipinski, C. A. (2004, December 1). Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discovery Today: Technologies*, Vol. 1, pp. 337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007>
20. Mbadiko, C. M., Inkoto, C. L., Gbolo, B. Z., Lengbiye, E. M., Kilembe, J. T., Matondo, A., ... Mpiana, P. T. (2020). A Mini Review on the Phytochemistry, Toxicology and Antiviral Activity of Some Medically Interesting Zingiberaceae Species. *Journal of Complementary and Alternative Medical Research*, 9(4), 44–56. <https://doi.org/10.9734/jocamr/2020/v9i430150>
21. Millet, J. K., & Whittaker, G. R. (2015). Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Research*. <https://doi.org/10.1016/j.virusres.2014.11.021>
22. Mole, B. (2013). Deadly pig virus slips through US borders. *Nature*. <https://doi.org/10.1038/499388a>
23. Morris, G. M., Ruth, H., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791. <https://doi.org/10.1002/jcc.21256>
24. Reusken, C. B. E. M., Haagmans, B. L., Müller, M. A., Gutierrez, C., Godeke, G. J., Meyer, B., ... Koopmans, M. P. G. (2013). Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: A comparative serological study. *The Lancet Infectious Diseases*. [https://doi.org/10.1016/S1473-3099\(13\)70164-6](https://doi.org/10.1016/S1473-3099(13)70164-6)
25. Tahir ul Qamar, M., Kiran, S., Ashfaq, U. A., Javed, M. R., Anwar, F., Ali, M. A., & Gilani, A. ul H. (2016). Discovery of novel dengue NS2B/NS3 protease inhibitors using pharmacophore modeling and molecular docking based virtual screening of the ZINC database. *International Journal of Pharmacology*, 12(6), 621–632. <https://doi.org/10.3923/ijp.2016.621.632>
26. Thomas, V. H., Bhattachar, S., Hitchingham, L., Zocharski, P., Naath, M., Surendran, N., ... El-Kattan, A. (2006). The road map to oral bioavailability: An industrial perspective. *Expert Opinion on Drug Metabolism and Toxicology*, Vol. 2, pp. 591–608. <https://doi.org/10.1517/17425255.2.4.591>
27. Volkamer, A., Kuhn, D., Rippmann, F., & Rarey, M. (2012). Dogsitescorer: A web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*, 28(15), 2074–2075. <https://doi.org/10.1093/bioinformatics/bts310>
28. Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., & Veelsler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*, 181(2), 281-292.e6. <https://doi.org/10.1016/j.cell.2020.02.058>
29. Xu, Y., Liu, Y., Lou, Z., Qin, L., Li, X., Bai, Z., ... Rao, Z. (2004). Structural basis for coronavirus-mediated membrane fusion: Crystal structure of mouse hepatitis virus spike protein fusion core. *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.M403760200>

human coronavirus S-proteins. A total of 107, 153,647 (35,731,007,499 residues) were searched, among which, 0.0002% (22,067 sequences) passed filters. The top hit was bat SARS-like CoV S-protein (Uniref ID: A0A2R3SUW7). Average pairwise distance, 0.245654; lower bound, 0.106424; upper bound, 0.364375. The optimal evolutionary model was WAG.

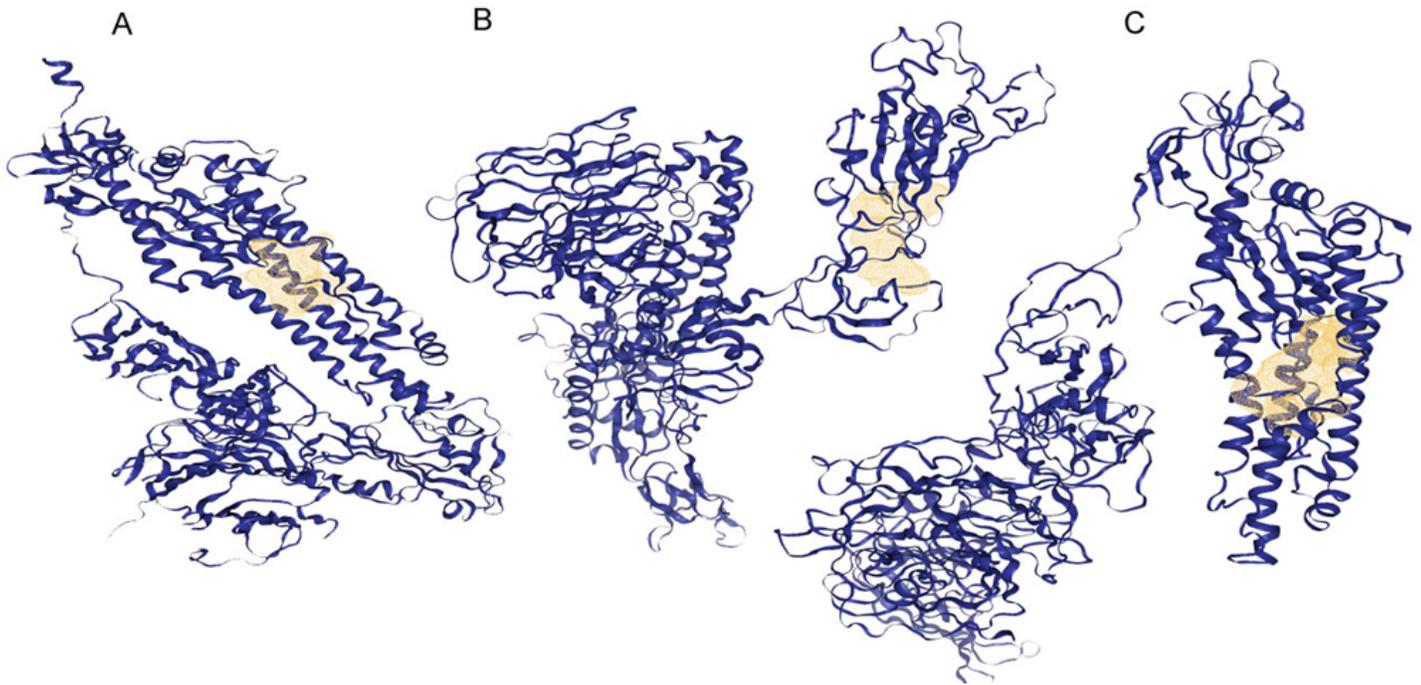


Figure 2

Different pocket position with highest score having different positions in structure. SARS-CoV-2 (A) and MERS-CoV have similar positions of drug binding pockets, while SARS-CoV-1 has completely different site. This SARS-CoV-1 is not continuous in single chain.

A 2D Plot protein ligand interaction with 4'-Me ether, 3,5-di-Ac 3,5-di-Gingerdiols (A) top hit, drug having highest S-value followed by 1-(5-Butyltetrahydro-2-furanyl)-2-hexacosanone (B) as second and Ginsenoyne N was third to have highest S-value.

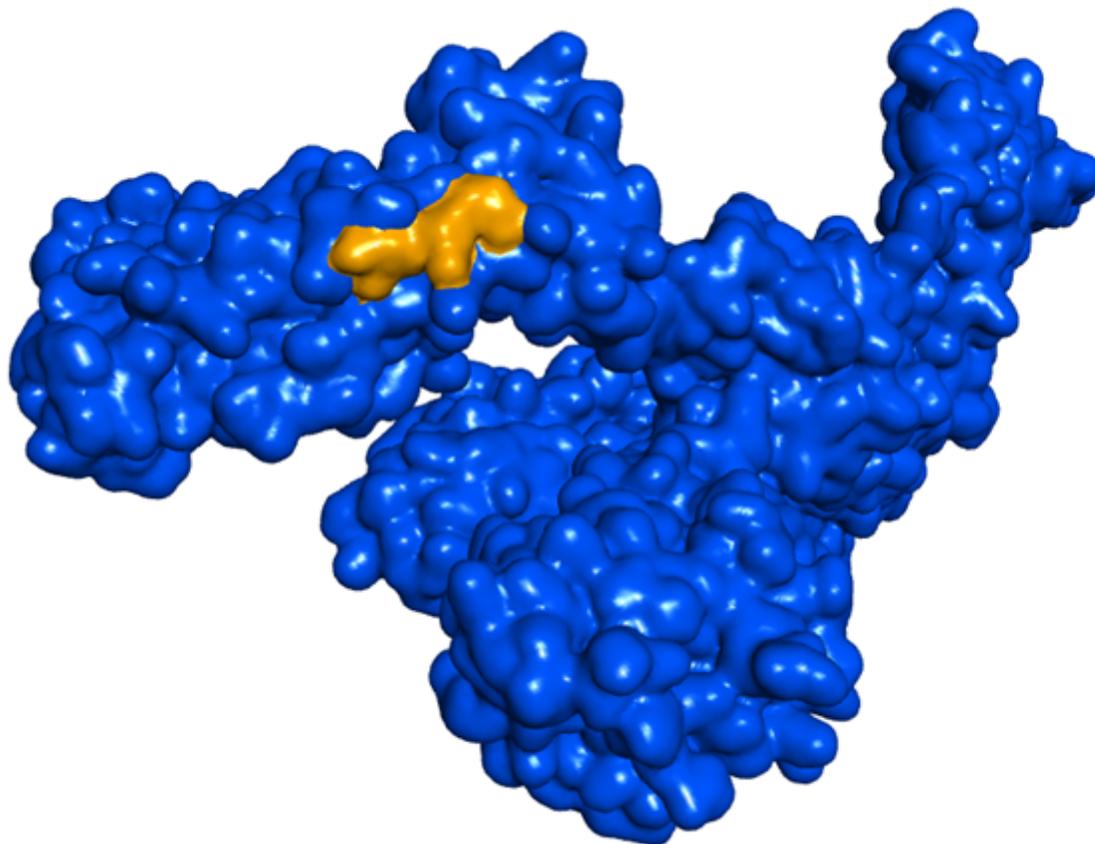


Figure 5

Protein drug interaction shows that protein (purple) has highly embedded drug (orange). Clearly showing a good interaction between protein and ligand with proper fitting in binding pocket.

Supplementary Files

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

- [Supplementaryfile1.txt](#)