

A plausible path to contain the spread of 2019 Novel Coronavirus; a computational analysis prompts to the use of ACE2 receptor blockers

Christopher Whitman (✉ chris@genetoo.com)

Genetoo Inc

Research Article

Keywords: 2019 Novel Coronavirus Spike Glycoprotein, ex-novo model, interaction with the human Angiotensin-converting Enzyme 2

Posted Date: April 7th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-21838/v1>

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

[Read Full License](#)

Abstract

Starting December 30th, 2019, a virus spread from Wuhan, in the Hubei Province of China. The virus had soon been recognized as part of the Coronavirus and temporarily named 2019 Novel Coronavirus. The dramatic increase of infections led to the death of over 400 people, by Feb 4th, 2020. By this day the virus had already crossed into 27 countries. March 11th, 2020 the World Health Organization declared the Novel Coronavirus a pandemic, pointing to over 118,000 cases of infections in over 110 countries. This public health threat drove the international community to real-time sharing of the genetic sequences isolated from the viruses. We used these freely accessible genetic data, while leveraging bioinformatic tools, with the intent to explore possible contributions to address this threat. Angiotensin-converting Enzyme 2 Inhibition has been proven to be a valuable strategy address the spread of SARS. After proving remarkable genetic similarities between SARS and the 2019 Novel Coronavirus, we computationally built the first known *ex-novo* model of the 2019 Novel Coronavirus Spike Glycoprotein entirely generated from its aminoacidic sequence, using I-TASEER. We then assessed the 2019 Novel Coronavirus interaction with the human Angiotensin-converting Enzyme 2. This research prompts at the potential use of Angiotensin-converting Enzyme 2 receptors blockers, as both clinical and prophylaxis measures to contain the spread of 2019 Novel Coronavirus.

Introduction

December 30th, 2019, the Wuhan Municipal Health Commission reported an increase in cases of pneumonia in Wuhan city, Hubei province of China¹. On December 31st, 2019 the WHO was alerted that the specific virus infection didn't reassemble any other known virus. One week later, January 7th, 2020, the Chinese Authorities confirmed the new virus to be a coronavirus, whose viral family includes common cold, SARS and MERS. The new virus was temporarily named 2019 Novel Coronavirus or 2019-n-CoV2 (COVID-2019). Coronaviruses belong to the family Coronaviridae and to the order of Nidovirales. These are enveloped, non-segmented, positive-sense RNA viruses. Coronaviruses are broadly distributed in mammals including humans³. The first sequence of COVID-2019 was shared by Dr. Yong-Zhen Zhang and colleagues at the Fudan University in Shanghai⁴. In the coming days, further sequences were deposited on the databases of the Global Initiative on Sharing all Influenza Data (GISAID) allowing scientist around the world to have access to data and be able to analyze genetic sequences⁵. By Jan 17th, 2020, new cases were reported in both Thailand and Japan.⁶ Given the global threat posed by this virus, several strategies were globally implemented by several governments, including repatriation of citizens and airlines halting their routes to China. By January 30th, multiple corporations, including Amazon, Starbuck and Apple had revised their operations in China due to COVID-2019, thus leading to important financial implications. By February 4th, over 20 thousand cases and 426 deaths were confirmed, in 27 countries. March 11th, 2020 the World Health Organization declared the Novel Coronavirus a pandemic, pointing to over 118,000 cases of infections in over 110 countries. Given the threats associated with the spread of this virus and its global implications, the international scientific community has joined forces to tackle this issue. Several portals were implemented allowing the real time sharing of information, including epidemiological⁷ and genetic data⁸. This study aims to leverage this

genetic information and use bioinformatic tools to analyze the nature of this virus, so do address possible contingency remedies. Genome analysis was performed to build a genealogy line of COVID–2019. Sequences alignments used to seek similarities to other known viral strains. Proteomic simulations were implemented to investigate the use of specific molecule as cure and/or prophylaxis measures to contain the spread of COVID- 2019.

Methods

Genomic sequences were retrieved from the GISAID database.⁸ The chosen COVID–2019 sequence was submitted Jan 24th, 2019 to the Division of Viral Disease, Centers for Disease Control and Prevention (Atlanta, GA) by Queen et al., title: Full genome sequence of first U.S. case of COVID–2019 (Accession MN985325)⁹. Using the National Center for Biotechnology Information (NCBI), Basic Local Alignment Search Tool (BLAST) the full 29882 pb RNA sequence was aligned with all sequences within the NCBI repository. A genealogic tree was generated. The analysis of the output data showed high conservancy of the sequence between gene 21563 and 25384. This sequence is known to be coding for a Surface Spike (S) Glycoprotein. Spike Glycoproteins (S-Protein) are a type I membrane-bound protein, which have been proven to be key in the spread of SARS by mediating to the viral attachment to the host cell receptor angiotensin-converting enzyme 2 (ACE2). In Human, the hACE2 receptor is located in chromosome X, between pb 15,56,033 and pb 15,602,148.¹⁰

Starting from the aminoacidic sequence of, I-TASSER was used to simulate the three-dimensional structure of COVID–2019 S-Protein. The tridimensional structure was then used to analyze both structural proprieties and possible interactions with target receptors. Five structures where generated by I-TASSER, the one with higher C-Score (–1) was then chosen to perform further analysis. Swiss-Dock was used to assess the interaction between the newly generated COVID- 2019 S-Protein and the hACE2 receptor (PDB ID: 4YAY). For comparison, a second simulation was performed to assess the docking properties of SARS Coronavirus Taiwan TC2 S-Protein (GenBank: AAQ01609.1) and the hACE2 receptor (PDB ID: 4YAY) the structure of the human ACE2 receptor used in this research was revealed by Zhang H et al using femtosecond crystallography in 2015.¹¹ The data generated through Swiss-Dock were then analyzed through Chimera. Data analysis showed remarkable similarities between the two docking models.

Results

The BLAST analysis of the entire COVID–2019 genome firstly isolated in the USA, showed E- value: 0.0, 89.12% identity and query cover of 95% with the Bat SARS-like coronavirus complete genome (Accession MG772933.1). E-value 0.0, identity of 82.34%, and Query Cover of 88% was further determined with the SARS coronavirus ZS-C, complete genome; (Accession AY395003.1). A 29647 pb RNA linear sequence submitted September 19th, 2003 to the SARS epidemiology consortium of Guangdong.

These data clearly prove little mutation and tight genetic correlation between SARS and COVID- 2019. The genealogy tree also points to a plausible common host animal ancestor to be the bat. (Figure 1).

Looking at the alignments between the COVID–2019 and the SARS genome, the region between gene 21563 and gene 25384 was determined to be highly conserved. This region is known to code for the COVID–2019 S-Protein (GenBank: QHO60594.1). A BLAST alignment of this aminoacidic sequence compared to the NCBI repository, showed E-Value 0.0, Identity 80.32% and Query Cover 99% with the S-Protein of Bat SARS-like coronavirus (Accession AVP78042.1) and E-Value 0.0, 76.27% Identity query cover 100%, with the SARS Coronavirus GZ02 S-Protein (GenBank: AAS00003.1). The aminoacidic sequence of the COVID–2019 S-Protein was then used to predict a tridimensional structure using I-TASSER; Five Models were generated with C-Scores respectively of –1.00, –1.24, –1.66 and –2.41. (Figure 2). C-score is typically in the range of –5,2, where a C-score of higher value signifies a model with a high confidence and vice-versa.

While generating the COVID–2019 S-Protein structure, I-TASSER also assessed matches with several templates. Comparison with Perfusion Structure of SARS-CoV spike glycoprotein, conformation 1 (PDB 5X58) showed 100% sequence identity of the template in the threading aligned region with the query sequence and 84% identity of the whole template chains with query sequence, coverage 84% (coverage of the threading alignment, which is equal to the number of aligned residues divided by the length of query protein) and Normalized Z-Score of 4.58, whereas alignment with a Normalized S-Score >1 mean a good alignment and vice versa.

Analysis of the predicted structure compared to PDB database showed identity of 98% and coverage of the alignment of 83% when compared to SARS-CoV complex with human neutralizing S230 antibody Fab fragment. (PDB 6nb6).

All together, these data show a reliable prediction, and close similarity between COVID–2019 S- Protein and SARS-CoV S-Protein.

The docking between the COVID–2019 S-Protein and the hACE2 receptor was simulated using the Swiss-Docking servers. Output data were further interpreter using Chimera. Analysis of the data showed similarities between the target-ligand interactions of the SARS-ACE2 and the COVID–2019-ACE2 docking models. (Figures 3a and 3b).

For the SARS-ACE2, the values ΔG ranged from –11.381753 to –5.243281, Total H bonds ranged from 14 to 0, with max number of H bond respectively 7 for Ligand Atoms and 12 for Receptor Atoms. Full fitness ranged from –773.06036 to –891.9281. While for the COVID–2019, the values ΔG ranged from –11.013132 to –5.3335133, H Bonds from 10 to 0, with max number of H bond respectively 5 for Ligand Atoms and 10 for Receptor Atoms. Full fitness ranged from –777.1416 to –895.999.

Despite a more detailed analysis of the generated docking models might be appropriate within another publication, these data strongly prove how the interaction between hACE2 and both SARS and COVID–2019 S-Protein are of the same nature. Both S-Proteins show affinity to the hACE receptor. Despite a lower affinity is reported for the hACE2-nCoV S-Protein complex, it is still reasonable to conclude that S-Proteins interactions with the hACE2 receptor play a key role also for the replication and spread of COVID–2019.

These data are in accordance also with the conclusions of Xintian Xu et al., that, constructing the Wuhan CoV S-protein by Swiss-model, and using SARS Coronavirus as a template, determined that despite the loss of hydrogen bonds due to the replacing of Arg426 and Asn426 in the COVID-2019 S-Protein, compared to the SARS S-Protein (increase by 28 Kcal mol⁻¹ when compared to the SARS S-Protein binding) the Wuhan CoV S-Protein, was found to have significant binding affinity to human ACE2, thus concluding that Wuhan COVID-2019 poses a significant public health risk for human-to-human transmission of COVID-2019 via the S-protein-ACE2 pathway.¹²

Discussion

Discovered in 2000, the zinc metallopeptidase Angiotensin-converter Enzyme 2 has been proven to be involved in hearth function, hypertension and diabetes. Inhibition of this receptor has been a valuable cure for hypertension, diabetic nephropathy and congestive heart failure. Angiotensin II blockers, also known as Sartans or Angiotensin II Antagonist, are pharmaceuticals that modulate the renin-angiotensin system. These include molecule such as Losartan, Irbersartan, Olmesartan, Valsartan and Azilsartan.

During the outbreak of Severe Acute Respiratory Syndrome (SARS) a Corona Virus disease that in 2003 rapidly spread from China to other countries, leading to almost 800 deaths, several researches identified ACE2 as a receptor involved in the spread of this virus. *In vitro* assessments, such as the one done by Wenhi Li et al., demonstrated ACE2 to be a functional receptor of SARS13. The research completed by M. Alejandra Tortorici et al. further concluded that coronaviruses Infections are initiated by the transmembrane viral S-Protein, which binds to host receptors and fuses the viral and cellular membranes.¹⁴

The capsid of Coronaviruses is enhanced by S-Protein that project outwards to give a “corona”-like appearance to the virus, hence its name.¹⁵

S-Proteins mediate the attachment of the virion to the cell membrane by binding to human ACE2 (and CLEC4m/DC-SIGNR) receptors and mediating the internalization of the virus into the endosome of host cells. It is also believed that proteolysis by cathepsin CTSL may activate the membrane fusion with endosomes.¹⁶

Further *in vivo* assessment proved again the critical role, in the replication and spread of SARS, given by the interaction of ACE2 receptors and S-Proteins. As an example, Yang, Xiu-hong et al., developed a mouse model of SARS by introducing the human gene of hACE2, driven by the mouse ACE2 promoter, into the mouse genome. The authors proved that SARS-CoV replicated more efficiently in the lungs of transgenic mice than those with wild-type. The authors further outlined how the pulmonary lesions of transgenic mice reassembled human SARS infections. They proved these mice to be valuable for testing potential vaccine and drug therapies for corona virus infections.¹⁷

This research proves a strict genetic correlation between SARS and the COVID-2019. The conserved S-Protein structure outlines a common cell biology pathway involved in the spread of both viruses. A

polygeny generously shared by GISAID further shows the highly conserved nature of COVID–2019 with at most seven mutations relative to a common ancestor reported by Feb 4th,2020, suggesting a shared common ancestor in Nov-Dec 2019 followed by subsequent human-to- human transmission.⁸

Conclusions

This research proves strict genetic correlations between SARS and COVID–2019. During the SARS emergency, several researches have proven the inhibition of ACE2 to be a promising strategy to address the spread of this virus. While building what, to the best of our knowledge, is the firstly known *ex novo* model of the COVID–2019 S-Spike entirely generated starting from its aminoacidic structure, using I-TASSER. This research proves the interaction between ACE2 and COVID–2019 S-Spike to be the same of if compared to the ACE2 and SARS S-Spike complex. In light of these findings it is reasonable to conclude that the use of ACE2 antagonist could be further explored both in terms of treatment and prophylaxis measure for the treatment of COVID–2019. If on the one hand Sartans represents good candidates, since their pharmacodynamics and pharmacokinetics are well known, on the other hand, we would like to point out that several other molecules, including some isolated by Chinese medical herbs, such as the root tubers of *Polygonum multiflorum*, and the vines of *P. multiflorum* have been proven to significantly block the S-protein and ACE2 interaction in a dose-dependent manner.¹⁸ The suggested approach indeed requires further clinical assessments, especially in light of the possible implications that lowering blood pressure could have on infected patients. But it does surely represent a valuable contribute to the international community for the development of active molecules able to address this global issue.

Declarations

Acknowledgments:

Data sequences were accessed through the Global Initiative on Sharing All Influenza Data (GISAID) database. Further sequences were accessed via the National Center for Biotechnology Information (NCBI) repositories. The open source Basic Local Alignment Search Tool (BLAST) was accessed via the NCBI portal. The I-TASSER Server was kindly offered by the Department of Computational Medicine and Bioinformatics at the University of Michigan. The Swiss-Dock used to simulate the docking was freely offered by the Swiss institute of Bioinformatics. Chimera, the open source software was used for both data visualization and interpretation. I further would like to acknowledge my wife's support, that allowed me to work on this project full time, although we have a small baby in the house. I did this research independently, while affiliated with my own Corporation: Genetoo Inc.

References

1. europa.eu. (2020). *Outbreak of acute respiratory syndrome associated with a novel coronavirus, China; First cases imported in the EU/EEA; second update*. [online] Available at:

- https://www.ecdc.europa.eu/sites/default/files/documents/Risk-assessment-pneumonia-Wuhan-China-26-Jan-2020_0.pdf [Accessed 4 Feb. 2020].
2. int. (2020). *Novel coronavirus (2019-nCoV)*. [online] Available at: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> [Accessed 4 Feb. 2020].
 3. Huang, , Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J. and Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*.
 4. org. Novel 2019 Coronavirus Genome. [Online] Available at: <http://virological.org/t/novel-2019-coronavirus-genome/319> [accessed on 19 Jan 2020].
 5. GSAID 2020 Coronavirus. [Online] Available at: <https://www.gisaid.org/CoV2020/> [accessed on 19 Jan 2020]
 6. World Health Novel Coronavirus—Japan (ex-China). [Online] Available at: <https://www.who.int/csr/don/16-january-2020-novel-coronavirus-japan-ex-china/en/> [accessed on 19 Jan 2020].
 7. maps.arcgis.com. (2020). *Operations Dashboard for ArcGIS*. [online] Available at: <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6> [Accessed 4 Feb. 2020].
 8. org. (2020). *GISAID - Next betaCoV App*. [online] Available at: <https://www.gisaid.org/epiflu-applications/next-betacov-app/> [Accessed 4 Feb. 2020].
 9. nlm.nih.gov. (2020). *Wuhan seafood market pneumonia virus isolate 2019-nCoV/USA- WA1/2020, c - Nucleotide - NCBI*. [online] Available at: <https://www.ncbi.nlm.nih.gov/nucore/MN985325>. [Accessed 4 Feb. 2020].
 10. org. (2020). *Gene: ACE2 (ENSG00000130234) - Summary - Homo sapiens - Ensembl genome browser 99*. [online] Available at: https://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000130234;r=X:15561033-15602148 [Accessed 4 Feb. 2020].
 11. Zhang, H., Unal, H., Gati, C., Han, G., Liu, W., Zatsopin, N., James, D., Wang, D., Nelson, G., Weierstall, , Sawaya, M., Xu, Q., Messerschmidt, M., Williams, G., Boutet, S., Yefanov, O., White, T., Wang, C., Ishchenko, A., Tirupula, K., Desnoyer, R., Coe, J., Conrad, C., Fromme, P., Stevens, R., Katritch, V., Karnik, S. and Cherezov, V. (2015). Structure of the Angiotensin Receptor Revealed by Serial Femtosecond Crystallography. *Cell*, 161(4), pp.833-844.
 12. Xu, X., Chen, P., Wang, J., Feng, J., Zhou, H., Li, X., Zhong, W. and Hao, P. (2020). Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Science China Life Sciences*. [online] Available at: <http://engine.scichina.com/doi/10.1007/s11427-020-1637-5>.
 13. Li, W., Moore, M., Vasilieva, N., Sui, J., Wong, S., Berne, M., Somasundaran, M., Sullivan, J., Luzuriaga, K., Greenough, T., Choe, H. and Farzan, M. (2003). Angiotensin-converting enzyme 2 is a functional

receptor for the SARS coronavirus. *Nature*, 426(6965), 450-454.

14. Tortorici, M., Walls, A., Lang, Y., Wang, C., Li, Z., Koerhuis, D., Boons, G., Bosch, B., Rey, F., de Groot, R. and Veerle, D. (2019). Structural basis for human coronavirus attachment to sialic acid receptors. *Nature Structural & Molecular Biology*, 26(6), 481-489.
15. Turner, A., Hiscox, J. and Hooper, N. (2004). ACE2: from vasopeptidase to SARS virus. *Trends in Pharmacological Sciences*, 25(6), pp.291-294.
16. org. (2020). *S - Spike glycoprotein precursor - Human SARS coronavirus (SARS-CoV) - S gene & protein*. [online] Available at: <https://www.uniprot.org/uniprot/P59594> [Accessed 4 Feb. 2020].
17. Yang, X., Deng, W., Tong, Z., Liu, Y., Zhang, L., Zhu, H., Gao, H., Huang, L., Liu, Y., Ma, C., Xu, Y., Ding, M., Deng, H. and Qin, C. (2020). Mice Transgenic for Human Angiotensin-converting Enzyme 2 Provide a Model for SARS Coronavirus. *Comparative Medicine*, [online] 57(5), pp.450-459. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/17974127> [Accessed 4 Feb. 2020].
18. HO, T., WU, S., CHEN, J., LI, C. and HSIANG, C. (2007). Emodin blocks the SARS coronavirus spike protein and angiotensin-converting enzyme 2 interaction. *Antiviral Research*, 74(2), pp.92- 101.

Figures

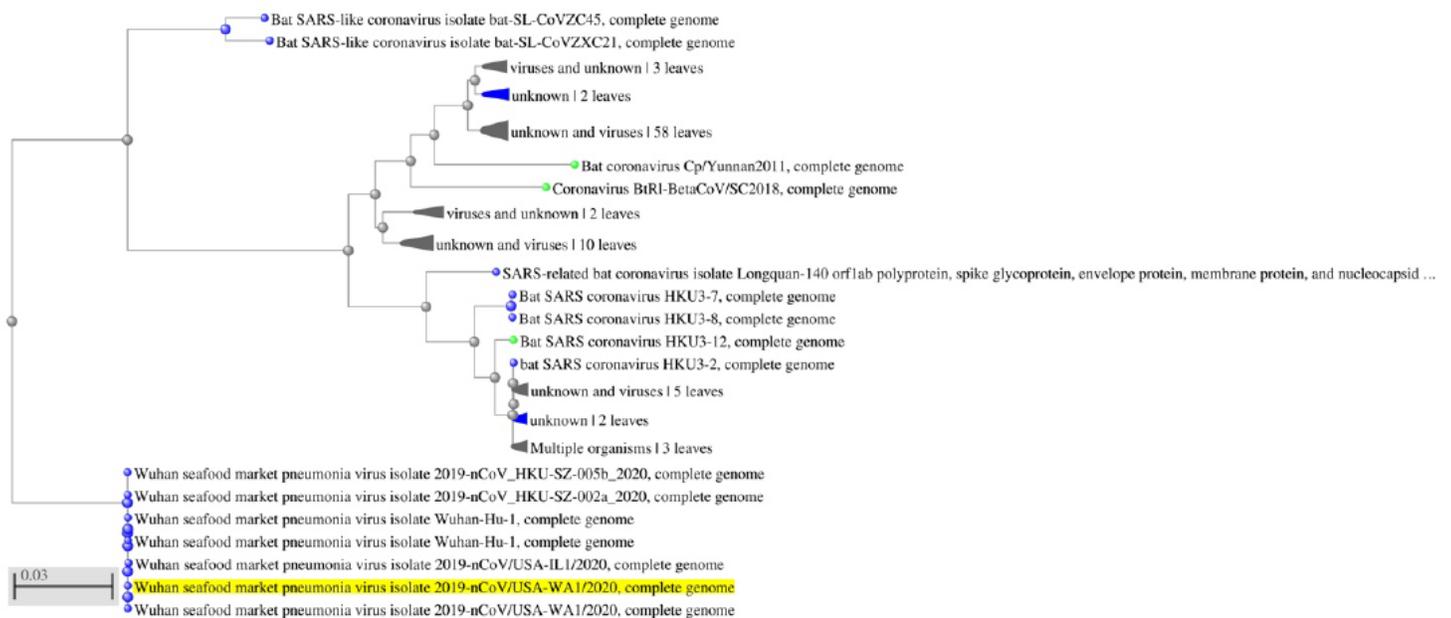


Figure 1

Genealogy tree built starting from BLAST search of the entire COVID-2019 genome showing tight genetic correlations with the Bat SARS coronavirus.

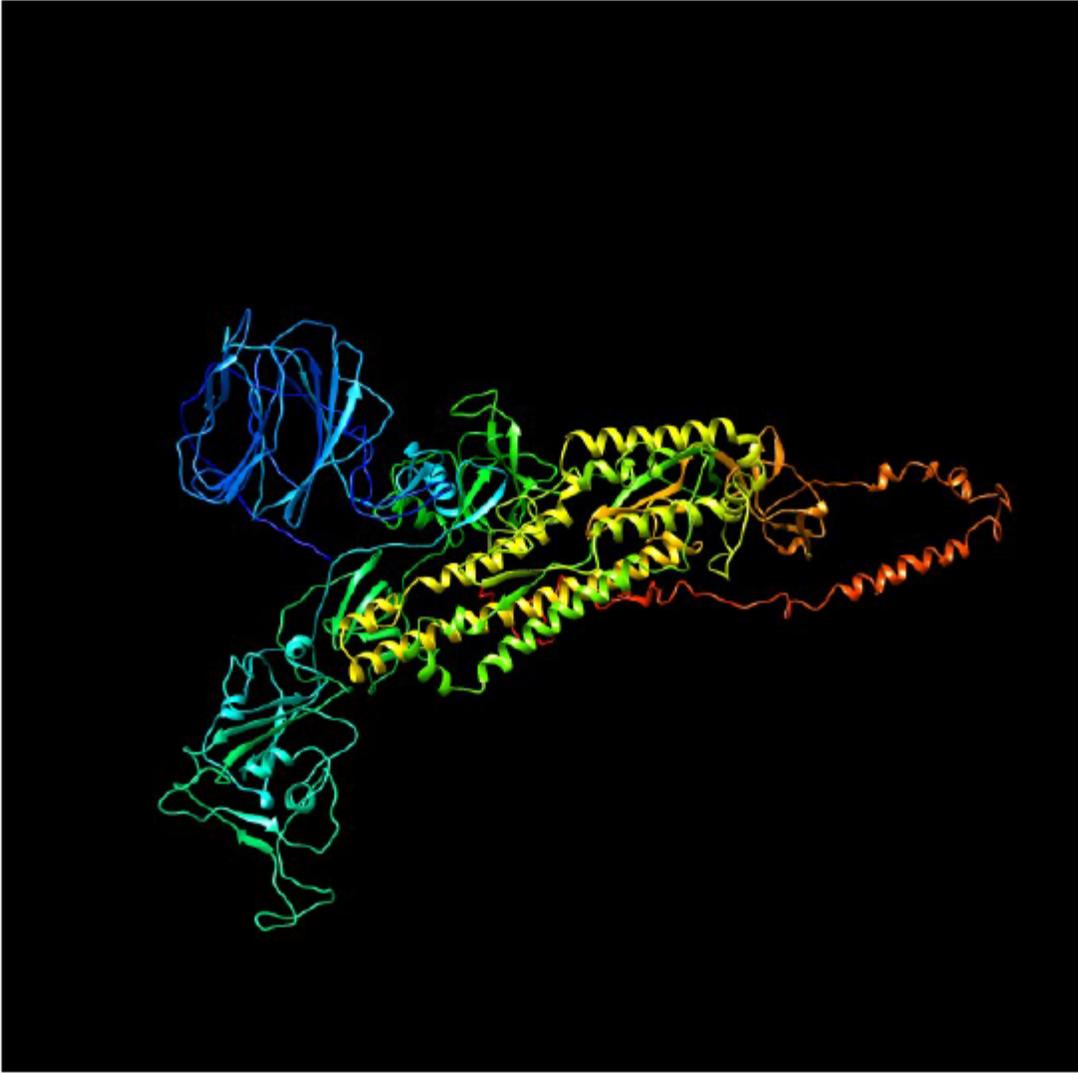


Figure 2

COVID-2019 S-Protein generated using I-Tasser – Model 1 – C-Score -1.00

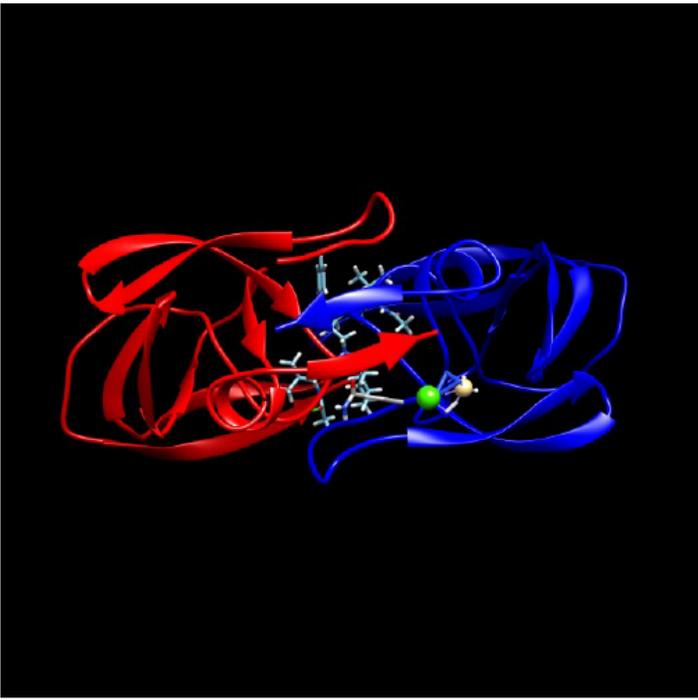


Fig 3a: SARS S-Protein in complex with ACE2.

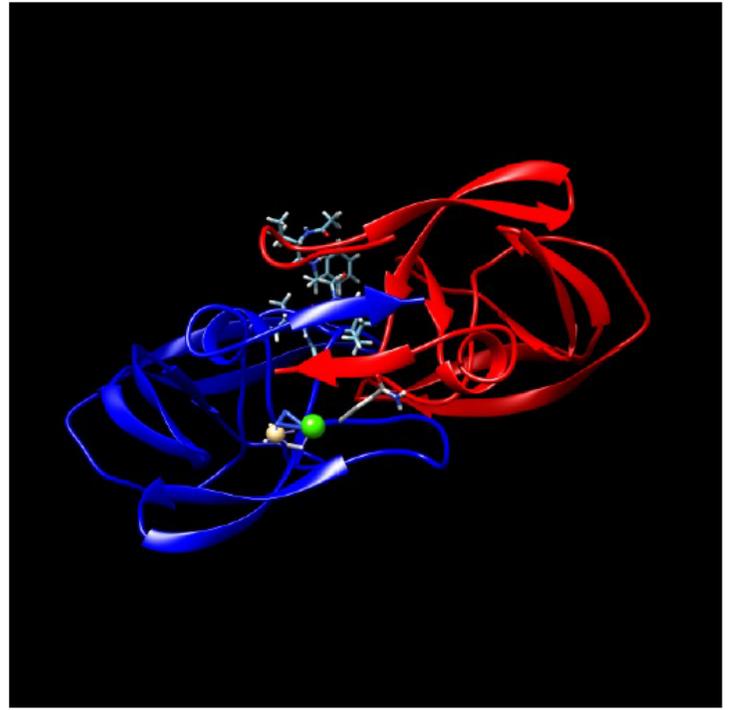


Fig 3b: COVID-2019 S-Protein in complex with ACE2

Figure 3

a: SARS S-Protein in complex with ACE2. b: COVID-2019 S-Protein in complex with ACE2