Virtual screening of inhibitors against spike glycoprotein of 2019 novel corona virus: a drug repurposing approach

Kanishka S Senathilake[1], Sameera R Samarakoon[1], Kamani H Tennekoon[1]

Institute of Biochemistry, Molecular Biology and Biotechnology, University of Colombo, Colombo 03. Sri Lanka.

Author of correspondence - Sameera R Samarakoon   email -sam@ibmbb.cmb.ac.lk

Abstract

The novel coronavirus (2019-nCoV) is a human and animal pathogen recently emerged in the city of Wuhan in Hubei province of China, causing a spectrum of severe respiratory illnesses. Corona viruses makes entry in to human cells through its spike (S) protein that binds to cell surface receptors. Wide spread of 2019-nCoV has been attributed to relatively high affinity of S protein to its receptor. Although S protein is a highly importantdrug target, unavailability of a high-resolution crystal structure and solvent accessible binding surface has made it a tedious target for current rapid virtual screening. A homology model of the receptor binding domain (RBD) of 2019 -n CoV S protein that is reasonably acceptable for drug screening was prepared using a high resolution crystal structure of SARS corona virus (SARS CoV)S protein. Data obtained from RBD- receptor docking experiments and published molecular dynamics experiments were used to map a RBD-receptor interaction hotspot that can be used for designing small molecule inhibitors. The hot spot was then used for virtual screening of more than 3000 drugs approved by U.S Food and Drug Administration (FDA) and other authorities for human use. Two anthracycline class drugs (zorubicin and aclarubicin) and a food dye (E 155) were predicted to be potent inhibitors of RBD – receptor interaction. Results of present study provide evidence for the potential of these compounds asprophylactic medications or for use to reduce disease severity of COVID -19.

1. Introduction

The outbreak of coronavirus disease (COVID-19) which was first reported from Wuhan, China has been reported to spread around the world having potential to become a pandemic (Chan et al, 2020). According to the World Health Organization (WHO) on February 27th 2020, there had been more than 82,000 cases globally (WHO, 2020). The novel coronavirus has been recognized as a new member of betacoronavirus genus which is closely

related to bat coronavirus and severe acute respiratory syndrome coronavirus (SARS CoV) (Lu et al, 2020). WHO has declared the 2019-nCoV as a Public Health Emergency ofInternational Concern (PHEIC) on the 30th of January 2020 by considering its rapid transmission from human to human and spread into many of the continents (Sun et al,2020).

A group of Chinese scientists has recently sequenced the complete genome of 2019-nCoV and disclosed that the 2019-nCoV is a new human-infecting beta-coronavirus (Lu et al, 2020). It has been reported that the 2019-nCoV is adequately divergent from SARS-CoV. Further, research groups have published evidence that the 2019-nCoV binds to the human angiotensin converting enzyme 2 (ACE-2) receptor through densely glycosylated spike (S) protein as the initiation step of the entry mechanism to human cells. (Lu et al, 2020). The S protein binds ACE-2 through its receptor binding domains (RBD) and the RBD – up conformation of the S protein is a prerequisite for the formation of RBD-ACE-2 complex(Wrapp et al,2020)

Wide spread of 2019-nCoV has been attributed to relatively high affinity of S protein to its receptor. Although highly important, unavailability of a high-resolution crystal structure and solvent accessible binding surface has made it a tedious target for current rapid virtual screening research groups.(Gruber and Steinkellner2019) Cryo EM structure of 2019-nCoV S protein (pre fusion down conformation) has been published at 3.5 A resolution(Wrapp et al,2020).Availability of several high resolution X-ray crystal structures of SARS CoV S protein (UP conformation) in complexed with ACE-2, structural rigidity and high degree of sequence similarity of RBDs warrants us to generate reasonably accurate homology models for drug screening.In the current study more than 3000 compounds approved by various regulatory authorities including U.S. Food and Drug Administration (FDA) were screened against an optimized homology model of the 2019 -nCoV S protein RBD.

2. Materials and method

2.1 Molecular modeling

Homology model of the RBD was constructed using modeler 9.23(Sali and Blundell, 1993, Webb and Sali, 2016) and using the crystal structure (2.9 A°) of BRD of SARS CoV (RBD domain complexed with ACE2 (PDB: 2AJF_E) as the template (76.47% sequence identity.). As previous studies have shown to preserve key interactions between RBD and ACE-2 of SARS CoV and 2019-nCoV, RBDs of 2019-nCoV (homology model) and SARS CoV

(2AJF_E) were superimposed on RBD of SARS CoV using match maker function of chimera (Peng et al,2020). Resulted BRD – ACE-2 complex was then energy minimized using chimera (400 steps of steepest descents, 50 steps of conjugated descents) while keeping the backbone of the ACE-2 fixed (Pettersen et al, 2004, Yang et al, 2012) using general Amber force field (James et al,2015) to obtain the final model (BRDACE2). Instead of using conformational ensembles, the single model generated by this approach was selected for virtual screening as it provides adequate evidence for a reliable model for virtual screening (see section 3.1).

2.2 Virtual screening

A library having geometry optimized 3D structures of 3350 approved compounds (world including FDA) was downloaded from Zinc library (http://zinc.docking.org/pdbqt/). Vina Grid Box was set to cover all the residues that form key interactions with ACE-2 (center X= 166.39, Y = 170..405, Z = 135.72, dimensions X = 28.91 Y = 28.59, Z = 28.55). Finally, all the ligands were docked using Auto Dock Vina (Trott and Olson, 2010) with the Lamarckian genetic algorithm as scoring function. TYR 449, GLN 493, GLN 498, TYR 501 and TYR 505 were made flexible and the other parameters were default values. Protein–ligand complexes having Vina energy values less than -7 were characterized using protein–ligand interaction profiler (PLIP) (Salentinet al., 2015) and PyMol. Compounds that do not form polar interactions with at least two critical residues were excluded.

3. Results and discussion

3.1. Selection of ligand binding site

Spike protein was selected for virtual screening instead of selecting its druggable ACE -2 receptor since compounds that block ACE -2 receptor is known to have modulatory effect on blood pressureand several other cardiovascular system related side effects (Mohammed et al, 2014). Cryo EM map of 2019-nCoVS protein has been published in receptor unbound prefusion conformation at a resolution of 3.5 Aº (Wrapp at al, 2020). The Cryo EM structure reflects the high degree of structural homology between 2019-nCoVRBD and SARS- CoV RBD having a RMSD value of 3.0 Å (Wrapp et al, 2020). Low resolution and the conformational considerations of Cryo EM structure makes it an unreliable structure for virtual screening. Therefore, a homology model was generated using high-resolution crystal structure of the SARS –CoV (PDB ID:2AJF) in ACE-2 bound up confirmation. Energy

minimized homology model used for the screening showed a Cα RMSD of 1.47 Å betweenthe RBD domains of 2019-nCoVand the SARS-CoV (Figure 1a).

Analysis of interaction surface of BRDACE2 in chimera showed strong H bonding at TYR 505 which is well exposed indicating an initial contact point with ACE-2. Further H bonding was observed at THR 500, GLY 502, TYR 449, GLN 493 and GLN 498. This approach gives interactions in consistence with pervious protein - protein docking data (Xu et al, 2020) and data arising from molecular dynamics (MD) experiments (Peng et al, 2020). Further surface analysis of the structure reveled Arg 439 > ASN 439 mutation on one side of the spike-like protrude (SpkP). This mutation seems to significantly weaken the H bond formed at ASN 439. There for the key interactions were restricted to one side of SpkP. (Figure 1b).

We took the advantage of ARG439> ASN 439 mutation in 2019-nCoV to define a potent inhibitor binding site on the side opposite to ASN 439. Hot spot spanned from GLN 498 through GLN 493 and Tyr 505 and ending in a hydrophobic pocket consisting LEU 555, ILE 418 and hydrophobic regions of LYS 417, TYR 421 and TYR 453. (Figure 1b)
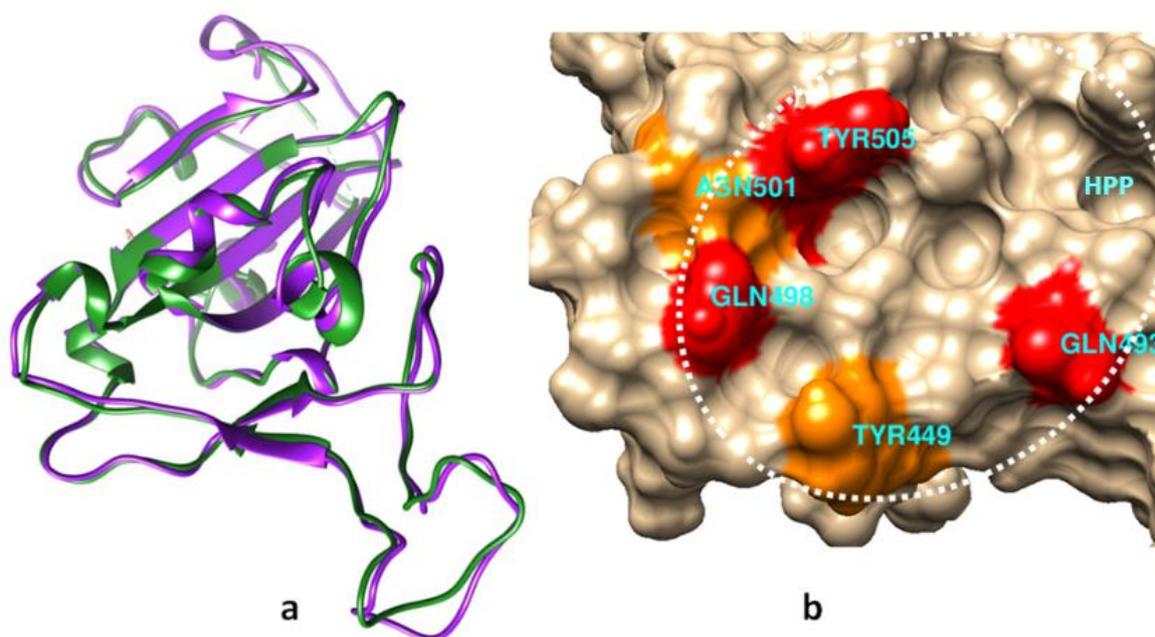

Figure 1- a. Superposed model of 2019-nCoV s protein RBD domain used for virtual screening (magenta) and RBD domain of SARS CoV (green) b. Hotspot of 2019-nCoVRBD selected for virtual screening. Residues involved in forming high energy interactions with ACE -2 are shown in orange and red colors HPP = hydrophobic pocket.

3.2. Binding interaction analysis

Two anthracycline compounds zorubicin, aclarubicin and an approved food dye E 115 were identified as potent inhibitors of BRD- ACE-2 interaction (Table 1). Highest energy biding pose of zorubicin (zinc accession - 03831623) formed two H bonds with the residue most critical for ACE-2 binding while forming one H bond with THR 501 which provide significand energy landscape for ACE-2 binding (Figure 3). Notably this populated binding pose (based on RMSD cutoff of 2) greater than 0.5 energy difference when compared with other poses.Attributing mainly to the ring structure, this highest energy binding pose had only a 1.2 A° RMSD deviation from its geometry optimized free form indicating no major conformational changes are required for the interaction. Thus the compound picks two most critical residues through H bonds while causing steric hindrance at Tyr 449. When bound to RBD, zorubicin stands in the binding site providing visible steric hindrance at other critical residues GLU 498 and GLU 493. Interestingly, carbonyl groups of quinine ring of zorubicin were not involved in forming interactions with RBD. This provides insights to further ligand optimization to improve binding and/or to reduce toxicity (Edwardso, 2015). Thus we propose strong evidence of zorubicin as a potent candidate for the inhibition RBD -ACE-2 interaction. (Table 1)

Table 1. Binding energy values and interactions between ligands and key residues of RBD. (HB= hydrogen bonding, HP = Hydrophobic interaction, CΠ= Cation- Π interaction)

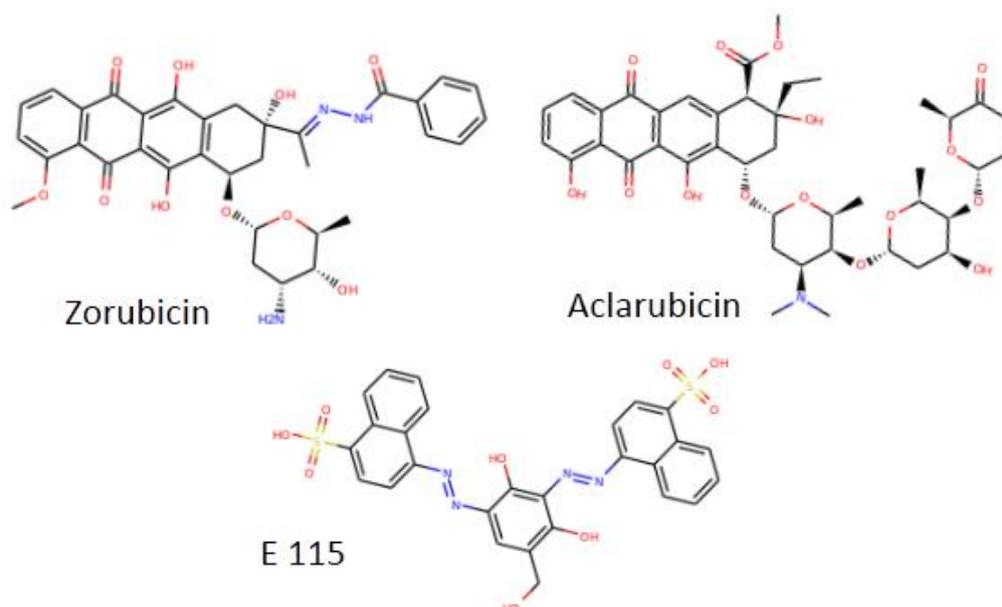| Zinc accession number	Name as in drug bank	Vina energy (kcal/mol)	Key interactions		
			Tyr 449	Thr 501	Tyr 505
03831623	Zorubicin	-7.4	HP (3.8)	HB (3.2)	HB (2.9) HB (3.2)
85537142	Aclarubicin	- 6.8	HB (2.98) CΠ (4.22)	HP (3.73) HP (3.97)	HB (3.23)
03830332	E155	- 7.3	HB(2.82) HB (3.01)	-	HP (3.4)

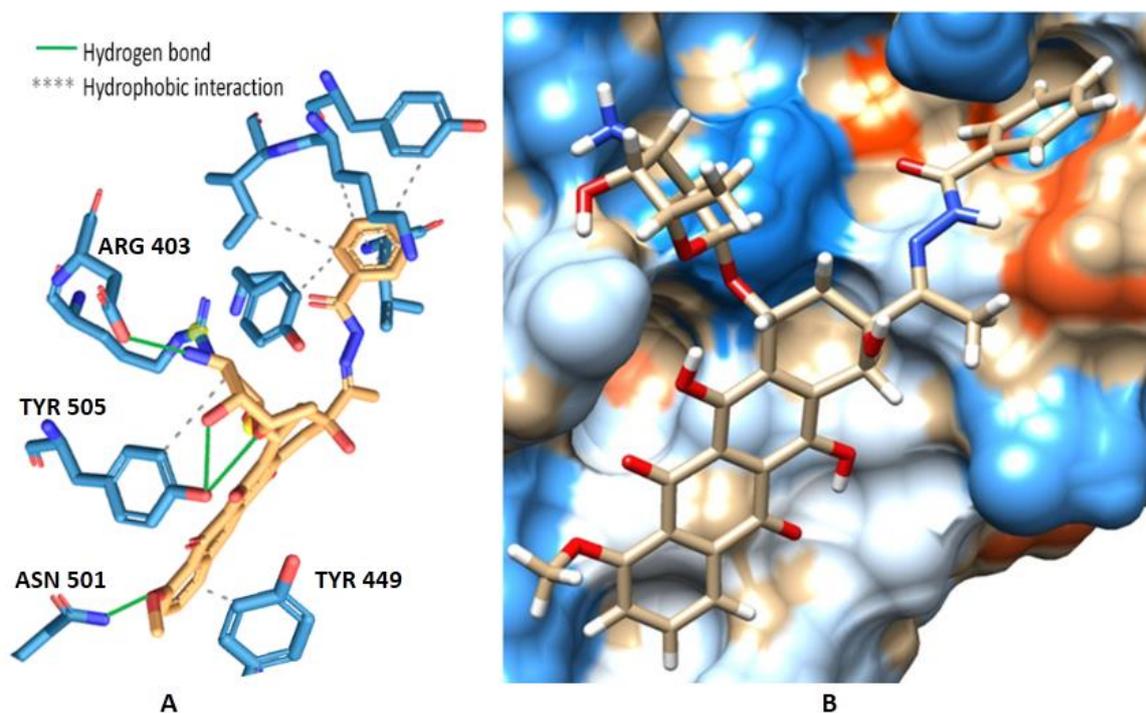Figure 2. Structures of inhibitors showing chirality.

Figure 3. Binding pose of zorubicin, with the selected ligand binding site of RBD. a) Predicted interactions. b.) Binding surface colored based on hydropathycity.

Potent second compound (aclarubicin) formed H bonds with TYR 505 and TYR 449 giving it the second priority (Figure 4). Binding pose was not the least energy conformation. However, the conformation was densely populated when clustered (manually based on < 2 A criteria).

Aclarubicin is anthracycline class drug similar to zorubicin. When bound aclarubicin stands in the binding site providing visible steric hindrance at other critical residues GLU 498 and GLU 493. Thus we propose evidence that the aclarubicin as a potent candidate for the inhibition RBD -ACE-2 interaction. (Table 1)
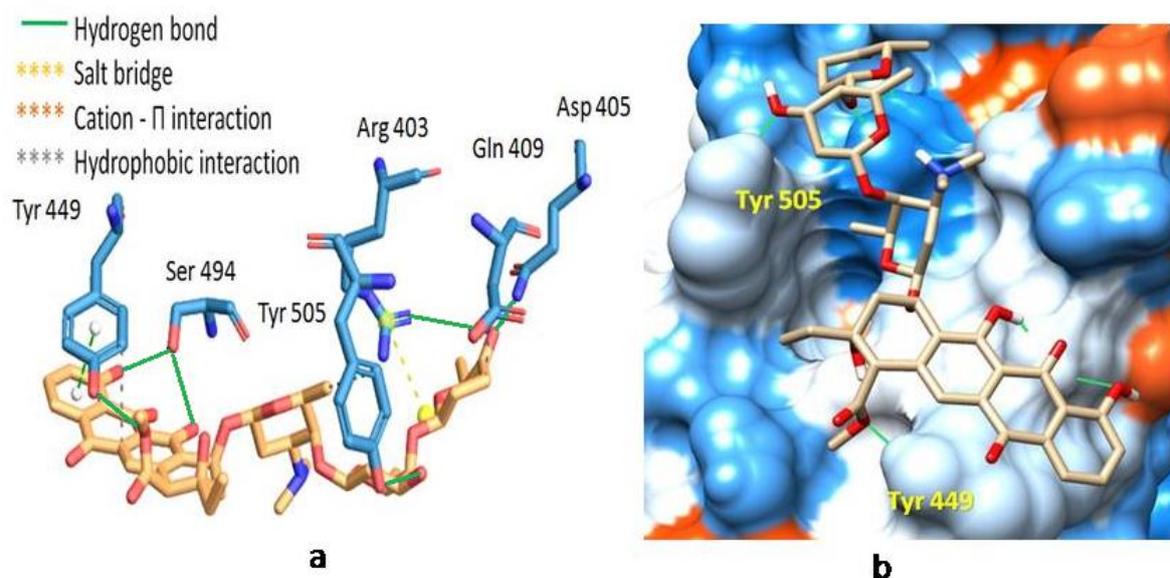

Figure 4. Binding pose of Aclarubicinwith the selected ligand binding site of RBD. a) Predicted interactions. b.) Binding surface colored based on hydropathycity.

Third compound E115 which is a food dye formed H bondswith key residues TYR 505 and TYR 449.When bound E115 stands in the binding site providing visible steric hindrance at other critical residues GLU 493 (Figure 5). Thus we propose evidence for E 155 as a potent candidate for the inhibition RBD -ACE-2 interaction. (Table 1)
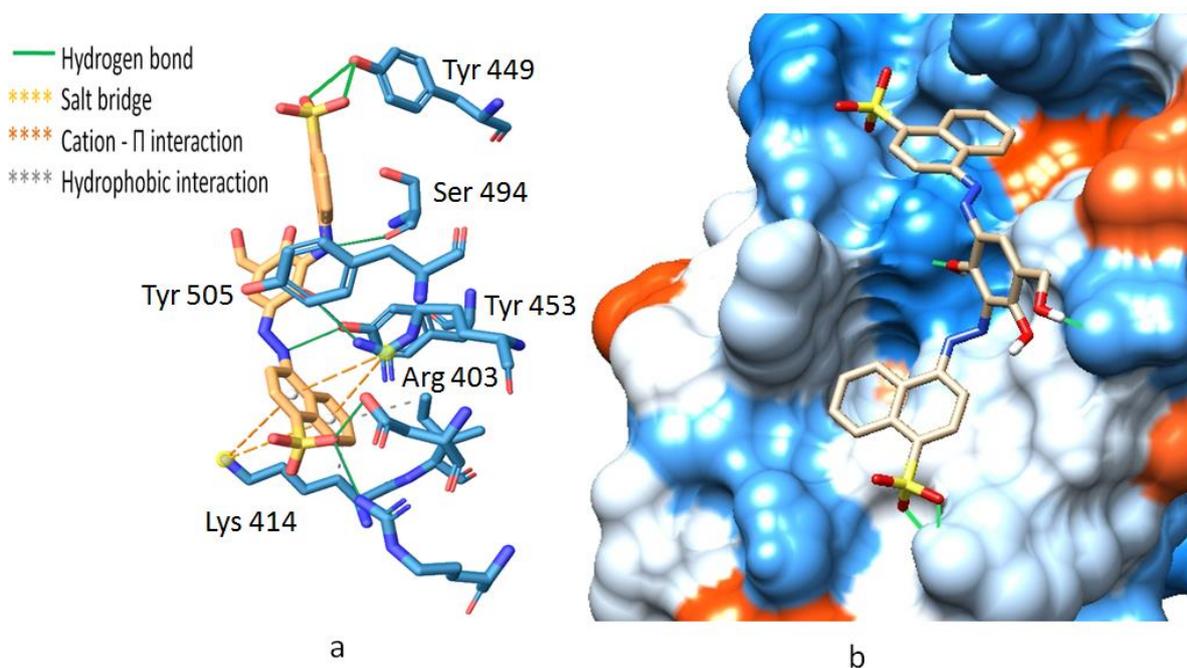
Figure 5. Binding pose of E 155, with the selected ligand binding site of RBD. a) Predicted interactions. b.) Binding surface colored based on hydropathycity.

According to our previous experience with protein-protein interaction inhibitors, a two hit hypothesis (targeting at least two critical protein – proten interactions) was employed to select potent inhibitors of S protein and ACE-2. This approach when coupled with high energy binding (vina energy over -7 kcal/mol) provides a better tool for inhibiting relatively strong protein interactions that span through large surface area.

In conclusion, Anthracycline class drug zorubicin and aclarubicin and food dye E 115 were predicted to be potent inhibitors of RBD–Receptor interaction. Results of present study suggest the potential of these compounds as a asprophylactic medication or use in preventive measures.

4. Compliance and ethics

The authors declare that they have no conflict of interest.

5.Acknowledgement

6. References

Chan JF, Yuan S, Kok KH**,** To KK, Chu H, Yang J, Xing F, Liu J, Yip CC, Poon RW, Tsoi HW, Lo SK, Chan KH, Poon VK, Chan WM, Ip JD, Cai JP, Cheng VC, Chen H, Hui CK, Yuen KY. (2020) A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 395:514-523

Edwardson DW, Narendrula R, Chewchuk S, Mispel-Beyer K, Mapletoft JP, Parissenti AM. Role of Drug Metabolism in the Cytotoxicity and Clinical Efficacy of Anthracyclines. Curr Drug Metab. 2015;16(6):412–426. doi:10.2174/1389200216888150915112039

Gruber C, Steinkellner G: Coronavirus COVID-19 (formerly known as Wuhan coronavirus and 2019-nCoV) - what we can find out on a structural bioinformatics level. 2020.

James A. Maier, Carmenza Martinez, KoushikKasavajhala, Lauren Wickstrom, Kevin E. Hauser, and Carlos Simmerling. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. Journal of Chemical Theory and Computation 2015 11 (8), 3696-3713DOI: 10.1021/acs.jctc.5b00255

Liu, X., Zhang, B., Jin, Z., Yang, H., Rao, Z.The crystal structure ofCOVID-19mainprotease in complex with an inhibitor N3 (2020) doi:10.2210/pdb6lu7/pdb

Lu R, Zhao X, Li J, Niu P, Yang B, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 2020; 395: 565–74.

Mohammed A.R ,Chamsi-Pasha MA, Shao Z, Tang WH. Angiotensin-converting enzyme 2 as a therapeutic target for heart failure. *Curr Heart Fail Rep*. 2014;11(1):58–63. doi:10.1007/s11897-013-0178-0.

O. Trott, A. J. Olson, AutoDockVina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, Journal of Computational Chemistry 31 (2010) 455-461

Peng, Cheng; Zhu, Zhengdan; Shi, Yulong; Wang, Xiaoyu; Mu, Kaijie; Yang, Yanqing; et al. (2020): Exploring the Binding Mechanism and Accessible Angle of SARS-CoV-2 Spike and ACE2 by Molecular Dynamics Simulation and Free Energy Calculation. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.11877492.v1

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. J. UCSF Chimera--a visualization system for exploratory research and analysis. Comput Chem. 2004 Oct;25(13):1605-12.

Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M (2015) PLIP: Fully automated protein-ligand interaction profiler. Nucleic Acids Res 43(W1):W443–W447

Sali A& Blundell T.L. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779-815, 1993.

Song, W., Gui, M., Wang, X., & Xiang, Y. (2018). Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. PLOS Pathogens, 14(8), e1007236. doi:10.1371/journal.ppat.1007236

Sun K, Chen J, Viboud C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a populationlevel observational study. Lancet Digital Health 2020 Published Online February 20, 2020 https://doi.org/10.1016/ S2589-7500(20)30026-1.

Webb B, Sali A. Comparative Protein Structure Modeling Using Modeller. Current Protocols in Bioinformatics 54, John Wiley & Sons, Inc., 5.6.1-5.6.37, 2016.

World Health Organization, Coronavirus disease 2019 (COVID-19) Situation Report – 38. 27.02.2020. Accessed online on 28th February 2020.https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200227-sitrep-38-covid-19.pdf?sfvrsn=9f98940c_2

Wrapp D, Wang N, Corbett K S., Goldsmith J A., Hsieh C, Abiona O, Graham B S, McLellan J S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation *Science* 19 Feb 2020:DOI: 10.1126/science.abb2507

Xu, X Chen v, Wang J, Feng J, Zhou H, Hao L X, W Z P Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission, 2020. Sci China Life Sci. 2020 Jan 21. doi: 10.1007/s11427-020-1637-5.

Yang Z, Lasker K, Schneidman-Duhovny D, Webb B, Huang CC, Pettersen EF, Goddard TD, Meng EC, Sali A, Ferrin TE.UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *J Struct Biol.* 2012 Sep;179(3):269-78