Repurposing of SARS inhibitors against COVID 19

Priyanka Rajbhar, Dikshant Singh and Ruchi Yadav*
Amity Institute of Biotechnology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow-226028
email: ryadav@lko.amity.edu

Abstract

Coronaviruses are group of similar viruses which cause fatal infection and responsible for affecting the upper respiratory tract in many organisms. Throughout the time these viruses have been found to affect human life by causing major pandemic like SARS, MERS and COVID-19 due to their high rate of mutation and zoonotic transmission. There are various organisms which could act as a reservoir for the viruses but studies have shown it could be bat because of high genomic similarity with bat coronavirus. Till this date there is no vaccine or drug for coronavirus responsible for COVID-19 which has created a major challenge to stop the spread of disease. Repurposing of a drug could be a solution for this challenge, as many previously available drugs hold great potential to act as a drug molecule. The coronavirus mainly affects the host by binding to its specific receptor with the help of its surface glycoprotein or spike glycoprotein. Thus, interfering this interaction could be a potent mechanism to stop the viral infection and propagation. In this paper, we have discussed about the surface glycoprotein and how previously available drug could be repurposed to act as a potent inhibitor. Homology modelling and docking studies has been mainly done to identify the interaction and binding affinity of previous drugs and how they could act as a potential solution to stop the overall spread of the disease in case of pandemic like COVID-19 where unavailability of specific drug or vaccine is responsible for taking to the lives of many.

Keywords: Coronaviruses, SARS, MERS, COVID-19, zoonotic, spike glycoprotein, Homology modelling, docking

Introduction

The new coronavirus identified in Wuhan, the capital of China's Hubei province in December 2019 showed the similar symptoms as SARS-CoV and MERS-CoV and people infected with this virus suffered a severe inflammatory response. The World health organization has named the new coronavirus as 2019-nCoV which later got changed into SARS-CoV-2 or COVID 19(Gorbalenya AE, Baker SC 2020). COVID 19 is classified as zoonotic viral disease (Benvenuto D, Giovanetti M 2020) similar to SARS-CoV and MERS-CoV which means that the patient who were infected acquired these viruses directly from animals. And is mainly transmitted through air and infects the respiratory and gastrointestinal tract of mammals and birds.

The name of the coronavirus comes from its resemblance to solar crown or corona like appearance (Almeida JD, Berry DM, 1968). The viruses are enveloped non segmented positive sense RNA viruses, 27-32kb in size (Woo, Patrick C 2010). The virus belongs to the family **Coronaviridae** and the order **Nidovirale (**Groot RJ, Baker SC 2011) and divided into four Genera i.e. alpha, beta, gamma and delta. The virus responsible for COVID-19 belongs to beta coronavirus just like as SARS and MERS.

Talking about the molecular information of the virus, two third of the genome encodes for two large overlapping polyprotein **ORF1a** and **ORF1b** which are processed into viral polymerase and other non-structural protein. The other third of the genome encode for four **structural protein** i.e. **spike, envelope, membrane and nucleocapsid**. Out of these four proteins, spike protein plays main role in recognition and attachment to the cell surface receptor and entry into the host cell. The protein present in the outer surface of virus is in hetro-trimeric state. The extraordinary variations in host range and tissue tropism among coronaviruses are in large part attributable to variations in the spike glycoprotein. These spikes function to define viral tropism by their receptor specificity and perhaps also by their membrane fusion activity during virus entry into cells. The spike protein is mainly consisting of S1, S2 and S2' subunit which play important role in viral infection.

S1: It mainly help in the attachment of viral particle to the host cell receptor (ACE 2). Binding further led the viral particle into the endosome of host cell and induce conformational changes in spike glycoprotein.

S2: It act as a class I viral fusion protein and mediate the fusion of the virus and cell membrane. During the process of fusion, the coiled region of protein begins to form a trimer of hairpin structure, and start arranging the fusion peptide in close proximities to C-terminal region of the ectodomain. And due to the formation and positioning of this structure subsequent fusion of viral and target cell membranes take place.

S2': It also act as fusion peptide and mainly act after S2 cleavage occurring upon virus endocytosis

Human Transmembrane protease serine 2(TMPRSS2) help the viral particle in priming with human lung cells (Hoffman M, Kliene-Weber 2020) which is an essential step for viral entry while proteolysis by **cathepsin CTSL** may uncovered the fusion peptide of Spike protein 2(S2) and activate membranes fusion within endosome.

ACE2: ACE2 is a type 1 integral membrane protein mainly expressed in endothelium, lungs, kidney and heart. The extracellular domain of ACE2 enzyme contain a single catalytic metallopeptidase unit which is responsible for converting of Angiotensin 2 to Angiotensin 1-7 and thus play a crucial role in Renin-Angiotensin system (RAS). Apart from these ACE2 is also associated with integrin function.

How the virus enters into the cell

The virus enters into our body through contact with an infected person or due to direct contact with the viral particles and they mainly attack the respiratory system specifically speaking the alveoli. The alveoli consist of two types of cell i.e. pneumocytes I and pneumocytes II, the virus infect the later one because of the presence of ACE2 receptor, which has been found to have a higher affinity for spike protein (Letko M, Marzi A 2020).

The S1 protein mainly help in the viral attachment (Wrapp D, Wang N, Corbett KS 2020) while S2 and S2'act as a fusion peptide and help in the fusion of the virus and cell membrane. At this point the host protease mainly help in exposure of the fusion peptide by cleaving at the proteolytic site.

The coronavirus has (+) sense ssRNA, and it can enter in the cell by several methods and can proliferate i.e. either through direct translation or through replication and uses the

host machinery proteins in order to continue. Some studies have shown that SARS-CoV 2 secretes at least three virulence factor that is responsible for production of the new viral particle and suppressing the immune response (Wu C, Liu Y, Yang Y, Zhang P 2020).

Thus, the spike protein act as a main key for entering in the cell and helps in the viral attachment. It also helps in the fusion and allow infection to begin. So, the structural study of spike protein is an important aspect to understand the molecular mechanism of viral infection and could be very important for creating vaccines and for therapeutic drug discovery.

Materials and Methods

Structure Prediction by Homology Modelling

The complete genome sequence of SARS-CoV-2 was published in the NCBI database (www.ncbi.nlm.nih.gov) with Accession no**. MN908947** under the title -Severe acute respiratory syndrome coronavirus 2 Wuhan-Hu-1. Form this database the sequence of surface glycoprotein was retrieved with Accession no. **QHD43416**

Sequence alignment of surface glycoprotein for Homology modeling was done using protein BLAST and the highest aligned sequence was selected to act as a template. On the basis of the template the protein 3d structure was predicted using the Schrodinger software suite version10.4.018(Schrodinger 2011). Modelled protein structure was verified by Ramachandran plot analysis using PROCHECK software (Laskowski et al. 1993) which was further used for binding site prediction, grid generation and docking.

Selection of potential drug compound as ligand

The selection of ligands was done by going through various research papers in the PubChem database and the final list was created consisting of six compounds as shown in Table 1.

These compounds were prepared for docking, using ligand preparation method (Sastryetal.2013) of Schrodinger software suite and docking was done using Glide dock method as implemented in Schrodinger software suite (Friesner et al. 2006).

Table 1 :Ligand used against COVID 19 protein which can act as potential inhibitor

S.No.	LIGAND NAME	PUBCHEM ID	Molecular formula	Molecular Wt.
1	Precose	444254	$C_{25}H_{43}NO_{18}$	645.6 g/mol
2	N-(1-Naphthyl)-2-(phenylthio)ethanethioamide	3246501	$C_{18}H_{15}NS_2$	309.5 g/mol
3	6MP-Arabinoside	3034423	$C_{10}H_{12}N_4O_4S$	284.29 g/mol
4	(2S)-2-(4-Methoxy-3,5-dimethylphenyl)-5-methyl-2-(3-pyrimidin-5-ylphenyl)-1,3-dihydroimidazol-4-amine	60202318	$C_{23}H_{25}N_5O$	387.5 g/mol
5	Verapamil	2520	$C_{27}H_{38}N_2O_4$	454.6

|				g/mol
6	2-(Phenethylthio)acetic acid	292540	$C_{10}H_{12}O_2S$	196.27 g/mol

Binding site prediction and docking

Binding site of modelled structure was predicted using sitemap tool (Halgren 2009) of Schrodinger software suit and the predicted binding site were then used for grid generation. The docking was done using Glide dock tool of Schrodinger software suit.

Ligand-protein interaction map was studied to identify binding properties and efficiency of the selected ligands against these proteins.

Result

Homology modelling is a method which uses template to predict the structure of query sequence and to build 3D model based on homology. The template identified for this procedure was 6acc_C (Spike glycoprotein of SARS-CoV) which is homologous with the SARS-CoV-2 spike glycoprotein with identity scores of 75%(Figure-1).
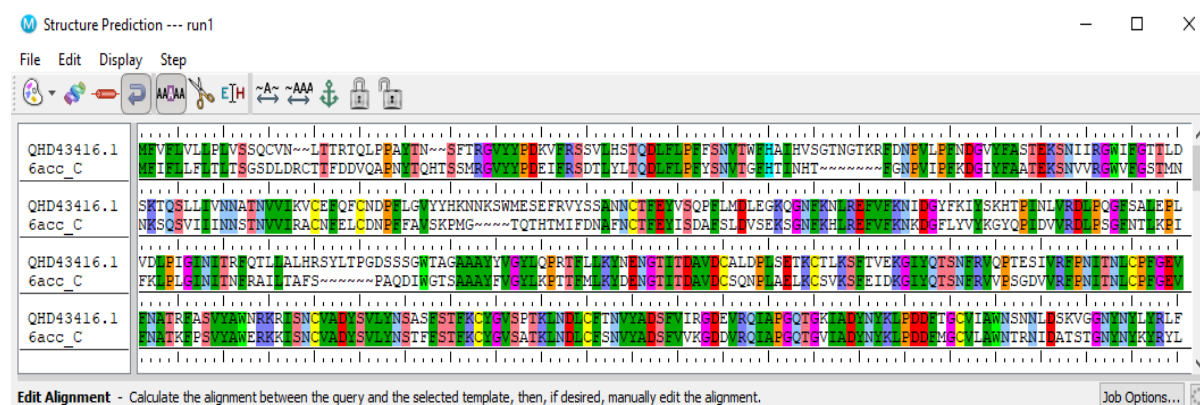

Figure 1 Template identification for Surface Glycoprotein of 2019-nCoV

Predicted 3D structure of spike glycoprotein (SARS-CoV-2) is shown in (Figure 2). Helices and sheets secondary structure can be seen in the modelled protein.
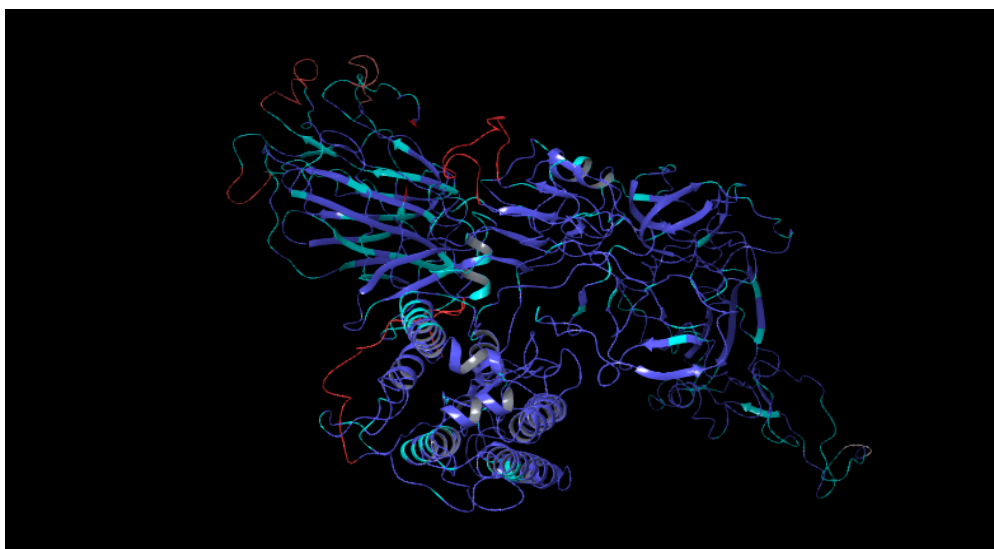
Figure 2 Modelled Structure of Surface Glycoprotein of 2019-nCoV

Structure verification of the modelled protein was done by Ramachandran plot analysis using ProCheck software which showed 86.0% of residues in favoured region, 11.8% Number of residues in allowed region and 2.2% residues in outlier region (Figure 3). The result show that modelled structure of surface glycoprotein of SARS 2-CoV was accurately modelled and can be further used for binding site prediction and docking studies.
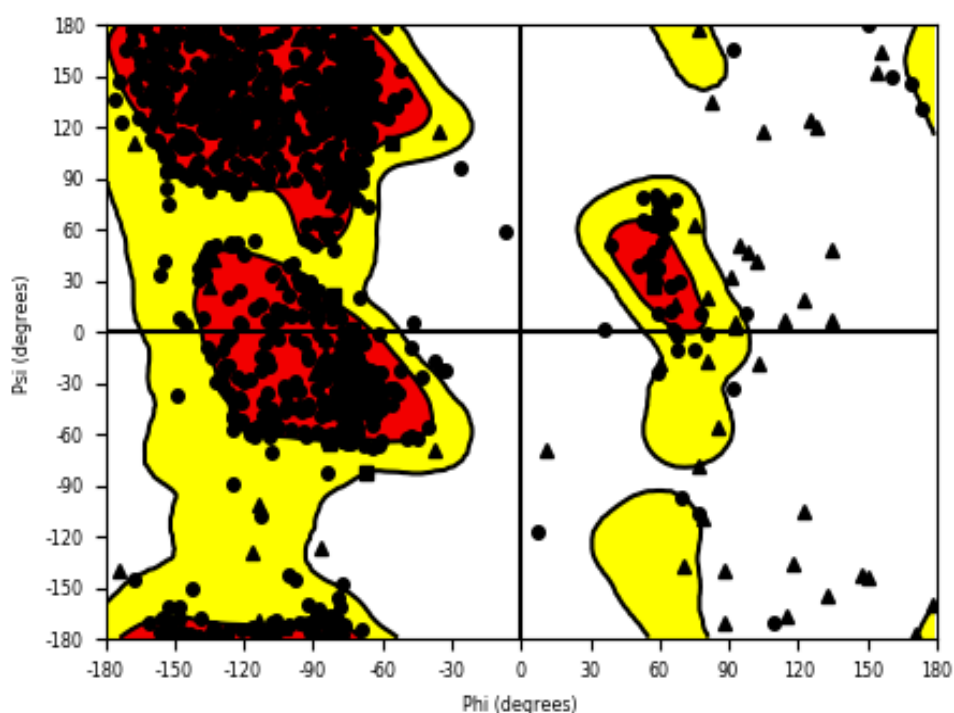
*Figure 3 Structure verification using Ramachandran Plot Analysis*

Binding site prediction and grid generation

Docking between ligands and modelled spike glycoprotein structure was done using slide dock method. The first step in the process of docking was the identification of the binding site; it was done using Sitemap tool of Schrodinger software (Figure 4a).

Binding site mainly shows the site score size, D score, volume, phobic-philic nature and residue position for each site predicted and the best binding site with the highest score was selected.

The next important step in docking was grid generation; it basically defines the binding positions on the target protein. The site with highest site score predicted using site map tool is used for grid generation and the grid are made using the grid generation program of glide dock as given in Schrodinger software. Grid map of modelled spike glycoprotein is shown in (Figure 4b).

The last main step was the docking which is done between the ligand and the modelled protein which mainly uses glide dock method as a tool.

(a)   (b)
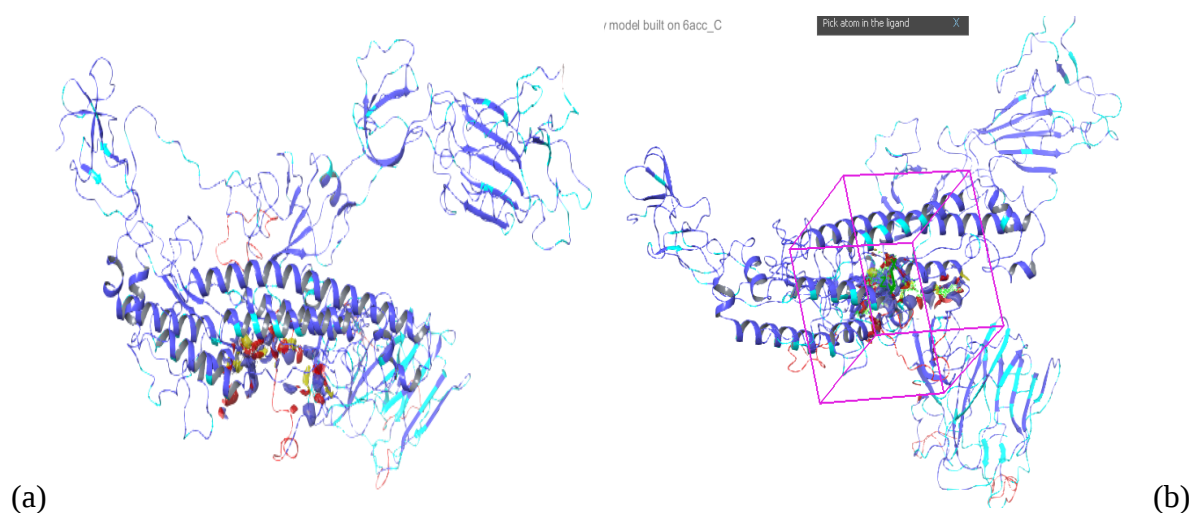
Figure 4(a) Binding site prediction (b) Grid generation of spike glycoprotein.

Glide docking result

All the six-ligand were docked against the predicted active site (grid) of the protein in order to identify the best ligand which could act as inhibitor for the selected protein. Table-2 shows docking result of all six ligands with their respective glide score.

S.No	LIGAND NAME	GLIDE SCORE
1.	Precose	-8.372
2.	N-(1-Naphthyl)-2-(phenylthio)ethanethioamide	-5.695
3.	6MP-Arabinoside	-5.383
4.	(2S)-2-(4-Methoxy-3,5-dimethylphenyl)-5-methyl-2-(3-pyrimidin-5-ylphenyl)-1,3-dihydroimidazol-4-amine	-4.605
5.	Verapamil	-4.119
6.	2-(Phenethylthio)acetic acid	-3.740

From the docking result we concluded that the highest glide score was of ligand with PubChem Id 444254 and is commonly Known as Precose having glide score or docking score of -8.372 which shows that it has stable and strong interaction with Spike glycoprotein.

Except, Precose other five ligand which were used out of the total six has docking score in between -5.69 to -3.74 which could be considered less stable as well as weak in interaction with the modelled protein structure Precose (Figure5). Precose is also commonly known as Acarbose and it's a pseudotetrasaccharide which is mainly responsible for the inhibition of the alpha- glucosidase and alpha- amylase with anti-hyperglycemic activity. Structure and molecular analysis of compound (Precose) shows that it has exact mass of 645.2480g/mol and 14-hydrogen bond donors and 19-hydrogen bond acceptors are present which help in the interaction.
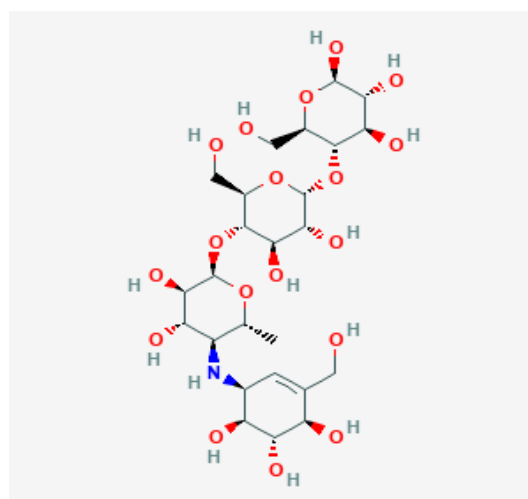

Figure 5-Precose($C_{25}H_{43}NO_{18}$) chemical structure

Protein-ligand interaction map of Precose was studied in order to find out the type of bond formed and amino-acid involved in the binding and is shown in Figure-6 which shows that compound makes hydrogen bond with PHE 823, VAL826,PRO 863, ASP 867 and HIS 1058 amino acid residue of the active site of spike glycoprotein of SARS-CoV-2 which potentially suggested that it could act as inhibitor for the spike glycoprotein and further interfere the binding between the spike protein and host receptor.
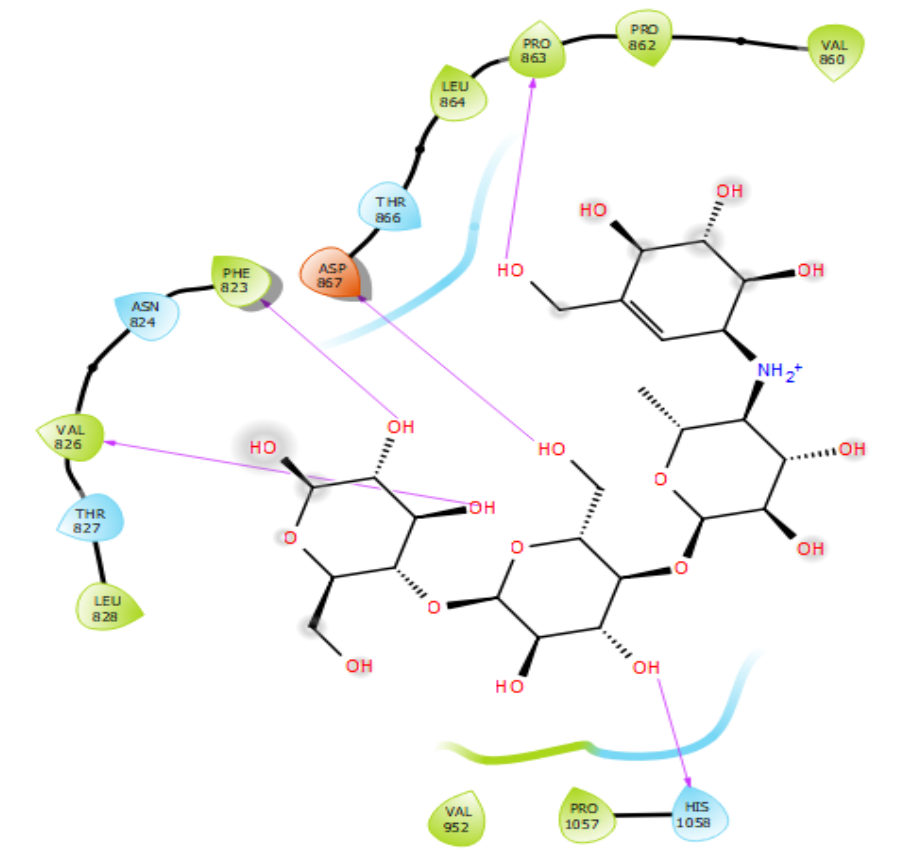
Figure 6 *Interaction of Surface Glycoprotein of 2019-nCoV with* **Precose ligand**. *It shows four hydrogen bond interaction with the surface glycoprotein, with glide score of -8.372.*

Conclusion and discussion

Viruses are sub-microscopic agents that mainly replicate when get inside a host and causes numerous infectious diseases in all life- forms (Koonin EV, Senkevich TG 2020). Throughout the time viruses has caused many deadly diseases influenza, chickenpox, AIDS, Ebola, SARS, MERS etc. Due to this virulence property the infection become much more severe and fatal in most case.

Coronavirus has diverse class of viruses which is responsible of many diseases in the past like SARS, MERS and recently SARS-2 (Bedford T, Neher R 2020), which is responsible for taking many lives and has caused major economic loss to many countries. Thus, developing a potential drug or vaccine is main task in order to stop the infection as well as to safe many lives.

Recent development in the bioinformatics allow the use of complex computational algorithm in order to solve the problems of biology, computers are now being used to store large amount of database which are produced every day and development of different tools

allow to process these data in a meaningful result. Prediction of protein structure by Homology modelling and finding a potent drug by molecular docking is one of the recent innovations in the field of bioinformatics and led to the discovery of many protein structure and ligand molecule which could be act as potent drug.

Repurposing of a drug basically means the use of previously available drug in order to find a cure for a novel disease, so throughout our work we just tried to repurpose the previously available drug by using bioinformatics tool and technique in order to find a potent drug molecule which act as a inhibitor for the spike glycoprotein thus inhibit the viral infection.

Structural properties of **Surface Glycoprotein of SARS 2-CoV**was studied and protein sequence of surface glycoprotein was retrieved from NCBI database. Homology modelling was done and further binding site and grid generation was done for docking studies. Docking result mainly helps in finding the insight of bond formation, ligand efficiency, binding affinity and stability of protein-ligand interaction. The result mainly showed Precose which is commonly known as Acarbose can act as a potential inhibitor for the spike glycoprotein, while the protein- ligand interaction map also showed the important amino acid with their position. This paper described and highlighted the importance of repurposing of the previously available drug to act as potent inhibitor in the newly discovered or novel diseases.

References

Gallagher TM, Buchmeier MJ. Coronavirus spike proteins in viral entry and pathogenesis. Virology. 2001 Jan 20.

Letko M, Marzi A, Munster V (February 2020). "Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses".

Novel Coronavirus (2019-nCoV): situation report, 22 (Report). World Health Organization. 11 February 2020.

Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, et al. (February 2020) A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster.

Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, Decroly E (February 2020). The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade.

Letko M, Munster V (January 2020). "Functional assessment of cell entry and receptor usage for lineage B β-coronaviruses, including 2019-nCoV.

Wang K, Chen W, Zhou YS, Lian JQ, Zhang Z, Du P, et al. (14 March 2020). SARS-CoV-2 invades host cells via a novel route.

"Novel coronavirus structure reveals targets for vaccines and treatments". *National Institutes of Health (NIH)*. 2 March 2020.

Oberholzer M, Febbo P (19 February 2020). What We Know Today about Coronavirus hbSARS-CoV-2 and Where Do We Go from Here.

Hoffman M, Kliene-Weber H, Krüger N, Herrler T, Erichsen S, Schiergens TS, et al. (16 April 2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor.

Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, et al. (February 2020). Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods.

Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. (February 2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation.

Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. (March 2020). "The species Severe acute respiratory syndrome-related coronavirus classifying 2019-nCoV and naming it SARS-CoV-2.

How COVID-19 Spreads. U.S. Centers for Disease Control and Prevention (CDC). 27 January 2020.

CDC (11 February 2020). "Coronavirus Disease 2019 (COVID-19)". Centers for Disease Control and Prevention.

Cyranoski D (26 February 2020). "Mystery deepens over animal source of coronavirus.