*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

# The Structure of the Membrane Protein of SARS-CoV-2 Resembles the Sugar Transporter semiSWEET

Sunil Thomas

Lankenau Institute for Medical Research, Wynnewood, PA-19096, USA.

E-mail: suntom2@gmail.com

## Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is responsible for the disease COVID-19 that has decimated the health and economy of our planet. The virus causes the disease not only in people but also in companion and wild animals. People with diabetes are at risk of the disease. As yet we do not know why the virus is highly successful in causing the pandemic within 3 months of its first report. The structural proteins of SARS include, membrane glycoprotein (M), envelope protein (E), nucleocapsid protein (N) and the spike protein (S). The structure and function of the most abundant structural protein of SARS-CoV-2, the membrane (M) glycoprotein is not fully understood. Using *in silico* analyses we determined the structure and potential function of the M protein. *In silico* analyses showed that the M protein of SARS-CoV-2 has a triple helix bundle, form a single 3-transmembrane domain (TM), and are homologous to the prokaryotic sugar transport protein semiSWEET. The advantage and role of sugar transporter like structures in viruses are unknown. If they are involved in energy metabolism, they may aid in the rapid proliferation and replication of the virus. Biological experiments should be performed to validate the presence and function of the semiSWEET sugar transporter.

## Key words

SARS-CoV-2, COVID-19, Virus, sugar transporter, SemiSWEET, Membrane glycoprotein.

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

## Introduction

The Covid-19 disease is currently responsible for the pandemic that has decimated the health and economy of every country. It is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a member of the betacoronavirus genus (Wang et al. 2020). Currently, the disease has resulted in a mortality of around 5-7 percent. As yet, there are no vaccines available commercially to protect against the virus.

The major structural proteins of SARS-CoV-2 are spike (S), membrane (M), envelop (E), and the nucleocapsid (N) proteins (Shereen et al. 2020; Chan et al. 2020). The spike protein of SARS-CoV-2 uses the host angiotensin-converting enzyme 2 (ACE2) as the entry receptor (Wrapp et al. 2020). Hence, the research community has an interest in studying the spike protein for drug and vaccine development.

The most abundant structural protein of coronaviruses is the M glycoprotein; it spans the membrane bilayer, leaving a short NH2-terminal domain outside the virus and a long COOH terminus (cytoplasmic domain) inside the virion (Mousavizadeh and Ghasemi, 2020). As the M proteins cooperates with the S protein, mutations may influence host cell attachment and entry of the viruses (Bianchi et al. 2020). The function of the M protein is also not fully understood. It is also not clearly understood how SARS-CoV-2 mediates sugar uptake and also the sugar transporters involved in the process.

Sugars will eventually be exported transporters (SWEETs) and SemiSWEETs are sugar transporters in eukaryotes and prokaryotes, respectively. SWEET proteins were first identified in plants as the novel family of sugar transporters that mediates the translocation of sugars across cell membranes (Chen et al. 2010, Feng and Frommer, 2015; Jia et al. 2018; Jeena et al. 2019). Sugar transporters are essential for the maintenance of blood glucose levels in animals, nectar production, phloem loading, seed and pollen development in plants, and also in pathogen nutrition (Chen et al. 2010; Jeena et al. 2019). Engineering of SWEET mutants using genomic editing tools mediated resistance to pathogens (Chen, 2014).

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

In eukaryotes, SWEET can discriminate and transport the uptake of mono and disaccharides across the plasma membrane by allowing solutes to permeate across biological membranes following a concentration gradient (Chen et al. 2010; Chen, 2014; Han et al. 2017). Eukaryotic SWEETs are composed of seven transmembrane helices (TMHs) that contain a pair of three transmembrane repeats, which are connected by an additional helix, while SemiSWEETs, the homologs of SWEETs in prokaryotes, contain three TMHs (Xuan et al. 2013; Feng and Frommer, 2015). The human genome contains only one *SWEET* gene and may be involved in glucose transport (Chen et al. 2010).

The prokaryotic semiSWEETs may be involved in the metabolism and transport of sugar synthesis. The semiSWEETs of prokaryotes are more diverse than SWEETs in plants; they seldom have homologs sharing >50% identity (Jia et al. 2018). The limited number of semiSWEET homologs suggest that they are not as important as the SWEETs in eukaryotes (Jia et al. 2018).

It is clearly not understood the function and role of the M proteins of the SARS-CoV-2 during host infection. Here, we report that the M proteins of SARS-CoV-2 are structurally similar to semiSWEET sugar transport proteins of prokaryotes based on *in silico* analyses.

## Materials and Methods

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

*SARS-CoV-2 protein structure*

The structural protein sequences of the SARS-CoV-2 were downloaded from Pubmed (https://www.ncbi.nlm.nih.gov/pubmed), protein database. The structural proteins include Membrane protein (Accession No. QJA17755), Envelope protein (Accession No. QJA17754), Spike protein (Accession No. QHR63290), Nucleocapsid protein (Accession No. QJC20758).

*Protein modeling*

Swiss model is a server that is used for 3D structure prediction. Homology modeling was constructed using Swiss model server (http//swissmodel.expasy.org/) with default settings. The M protein sequence of SARS-CoV-2 was entered in FASTA format.

Residue-based diagram of proteins, also called snake diagrams or protein plots, are 2D representations of a protein sequence that contains information about properties such as secondary structure (Skrabanek et al. 2003). To determine snake diagram model of protein we used Protter (http://wlab.ethz.ch/protter). Protter is an open-source tool that supports interactive protein data analysis and hypothesis generation by visualizing both annotated sequence features and experimental proteomic data in the context of protein topology.

*Sequence alignment*

Clustal W2 is a server for multiple sequence alignment which is also used for phylogenetic tree analysis. Multiple sequence alignment between M protein of SARS-CoV-2 and semiSWEET sequences from different microorganisms was performed using the clustalW2 server (http:/www.ebi.ac.uk/tools/msa/clustalW2/).

## Results

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

The S protein of SARS-CoV-2 binds to ACE2 receptors of the host for cell entry and may be a key target for drugs and vaccines. Hence the S protein of SARS-CoV-2 virus is well characterized. The SARS-CoV-2 is one of the most successful virus as it caused a pandemic within just two months of its first report in Wuhan, China. As yet, we do not yet know why the virus is successful in inducing a pandemic leading to millions of infection and thousands of death.

Three-dimensional (3D) protein structures provide valuable insights into the molecular basis of protein function (Schweede et al. 2003). Using *in silico* techniques the structure and potential function of the M protein of the SARS-CoV-2 virus is elucidated.

The structural protein sequence of the membrane protein (M) of SARS-CoV-2 was downloaded from NCBI protein database (Fig. 1). The FASTA sequence of the M protein was entered into the Swiss model server. Based on the sequence, the structure of the molecule was predicted as bidirectional sugar transporter SWEET2b. The ribbon representation, spacefill and surface models of the M protein is shown in Fig. 2.

The sugar transporter SWEET of eukaryotes are generally composed of seven transmembrane helices. Modeling proteins using residue-based diagram (snake diagrams) helps understand its function. Hence, we used Protter (http://wlab.ethz.ch/protter) to model the M protein.

The M glycoprotein is the most abundant envelope protein of SARS-CoV-2. *In silico* analyses of the M protein of SARS-CoV-2 using Protter demonstrated that it has a triple helix bundle, and formed a single 3-transmembrane domain (TM). In addition, the M glycoprotein has a short amino terminal outside the membrane and a long carboxy-terminal domain inside the membrane (Fig. 3A). The SWISS-MODEL predicted the M glycoprotein as SWEET2b. However, the M protein only has three transmembrane helices, not six or seven transmembrane helices observed in the SWEET sugar transporters of eukaryotes. Hence, the M glycoprotein structure of SARS-CoV-2 may be considered as semiSWEET. To confirm accuracy of the study, we also modeled the E, N

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

and S proteins of SARS-CoV-2. The modeling showed that the E protein has long outer amino terminal, a single helix and a short inner carboxy-terminal (Fig. 3B). The N protein had its entire structure inside the membrane (Fig. 3C). Whereas, the S protein had the majority of its structure outside the membrane with only a short carboxy-terminal inside the membrane (Fig. 3D).

The SemiSWEET sugar transporter of prokaryotes are more diverse than SWEET counterpart in plants. In the prokaryotes the semiSWEET seldom share identity. We used Clustal W2 to determine sequence homology of the sugar transporters of multiple microorganisms. The sequence of semiSWEET of the M glycoprotein of SARS had an identity of 26% with the semiSWEET of *Rhizobiales* and 20% with *Streptococcus pneumoniae* demonstrating that the semiSWEET of the SARS-CoV-2 may be highly conserved (Fig. 4A, B).

## Discussion

The COVID-19 pandemic caused by the coronavirus SARS-CoV-2 is spreading at an alarming rate and has resulted in an unprecedented health emergency all over the world (Ghosh et al. 2020). The rapid spread of SARS-CoV-2 justifies the global effort to identify effective preventive strategies and optimal medical management (Castagnoli et al. 2020). As yet there is no effective vaccine to protect against COVID-19 or effective approved drugs to treat patients with the disease. The development of antivirals is an urgent priority to combat the disease (Ghosh et al. 2020). Understanding the biochemical events of the coronavirus replication cycle may provide a number of attractive targets for drug development (Ghosh et al. 2020). Current strategies involve developing drug and vaccine candidates against spike (S) protein of the virus. The rationale being that neutralizing antibodies against the S protein prevent uptake of the virus via the human ACE2 receptor (Le et al. 2020). Identifying drug targets that blunt the activity of the virus may lead to effective treatments for COVID-19.

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

Viruses are non-living entities, without any organelles devoid of their own metabolism, though they have the capability to dramatically modify the host cellular metabolism upon entry. Viruses upregulate consumption of glucose and converge on similar metabolic pathways for anabolism (Thaker et al. 2019). Virus-induced metabolism may provide free nucleotides for rapid viral genome replication, increased amino acid production for rapid virion assembly, and high amounts of ATP for the high energy costs of genome replication and packaging. The mechanism for increased glucose uptake by the virus is still not clearly understood.

Glucose is the energy source of cells and tissues. Cellular uptake of glucose is a fundamental process for metabolism, growth, and homeostasis. Glucose is a polar molecule that does not readily diffuse across the hydrophobic plasma membrane of the cells. Glucose molecules are transported through the glucose transporters that include, GLUTs, the sodium-driven glucose symporters SGLTs, STP, and SWEETs (Deng and Yan, 2016). SWEETs are seen in plants and animals. SWEET induction by plant pathogens leads to secretion of sucrose that is used by these microorganisms for nutrition/reproduction (Bezrutczyk et al. 2017).

The bacterial ancestors of SWEET, known as semiSWEET are the smallest of the sugar transporters and assemble into dimers (Xuan et al. 2013; Chen et al. 2015; Lee et al. 2015). In fact, eukaryotic SWEETs consist of two SemiSWEET-like units fused via an inversion linker transmembrane helix (Jia et al. 2018). The diverse gene neighbors of semiSWEETs suggest that semiSWEETs may transport diverse substrates and play diverse physiological roles in different organisms (Jia et al. 2018). As yet there are no reports of sugar transporters in viruses.

It is not known how SARS-CoV-2 has been successful to spread all over the world within three months of its first report in Wuhan, China. Identifying the mechanisms of how viruses alter cellular metabolism and where in the virus life cycle these metabolic changes are necessary will provide an understanding of virus replication needs and potentially provide cellular targets for inhibition of these viruses. In this paper using *in silico* data

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

analysis we demonstrate that the structure of the membrane (M) glycoprotein of SARS-CoV-2 resemble the semiSWEET sugar transporter of the prokaryotes.

Clues to the viral metabolism can be understood from the patient population at risk of infection. It is known that people with diabetes are more prone to COVID-19 disease (Bornstein et al. 2020). It has been demonstrated that SARS coronavirus enters islets and damages islets causing acute diabetes (Yang et al. 2010). As people with diabetes have high glucose, the environment may favor proliferation of viruses.

Virus uses multiple mechanisms for the uptake of glucose. *Human cytomegalovirus* (HCMV), a herpesvirus, induces the sugar transporter, GLUT4 to increase glucose uptake during infection (Yu et al. 2011). Whereas, transmissible gastroenteritis virus (TGEV), a coronavirus induces multiple sugar transporters EGFR, SGLT1 and GLUT2 for glucose uptake (Dai et al. 2016). Rhinoviruses (RVs) are responsible for the majority of upper airway infections and they enhance the expression of the PI3K-regulated glucose transporter GLUT1; glucose deprivation from medium and via glycolysis inhibition by 2-deoxyglucose (2-DG) impairs viral replication (Gualdoni et al. 2018). Having a glucose transporter in the virus membrane may influence the energy metabolism thereby aiding in virus proliferation.

The membrane (M) glycoprotein is the most abundant envelope protein of coronaviruses (deHaan et al. 1999). *In silico* analysis showed that the M protein of SARS-CoV-2 resembles the sugar transporter, SWEET. Further analysis by residue-based structure demonstrated that the protein has the characteristic structure of semiSWEET, the sugar transporter of prokaryotes. To our knowledge this is the first report of the presence of a sugar transporter in a virus membrane. It is known that the prokaryotes have diverse sugar transporters. In our analysis, the SARS-CoV-2 sequence of semiSWEET has no homology to other prokaryotes.

One of the advantage of the virus having a sugar transporter in its membrane is that it may influence energy metabolism. How, the virus utilizes sugar molecules is unknown.

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

Presence of a semiSWEET glucose transporter in the M protein of the virus may be an efficient mechanism that may induce its rapid proliferation. People with diabetes are at risk of COVID-19 infection may be due to the high proliferation of the virus due to unmetabolized glucose. The current data are based on *in silico* analyses. However, further biological experiments are required to validate the presence and function of the virus sugar transporter.

**Figure Legends**

Fig. 1. The protein sequence of the M glycoprotein of SARS-CoV-2. The sequence was downloaded from NCBI protein database.

Fig. 2. Predicted M protein structure using the software SWISS-MODEL. The (A) ribbon representation, (B) spacefill and (C) surface models of the M protein of SARS-CoV-2.

Fig. 3. Membrane topology of proteins (snake diagrams) determined using Protter. (A) The membrane (M) glycoprotein of SARS-CoV-2 has a triple helix bundle, and formed a single 3-transmembrane domain. (B) Snake diagram of envelope (E) protein, (C) nucleocapsid (N) protein, and (D) spike protein (S).

Fig. 4. Protein sequences were aligned using ClustalW. (A) Comparison of protein sequence of the M protein of SARS-COV-2 with semiSWEET sugar transporter of Rhizobiales. (B) Comparison of protein sequence of the M protein of SARS-COV-2 with semiSWEET sugar transporter of *Streptococcus pneumoniae.*

**Reference**

Bezrutczyk M, Yang J, Eom JS, Prior M, Sosso D, Hartwig T, Szurek B, Oliva R, Vera-Cruz C, White FF, Yang B, Frommer WB. Sugar flux and signaling in plant-microbe interactions. Plant J. 2018; 93:675-685.

Bianchi M, Benvenuto D, Giovanetti M, Angeletti S, Ciccozzi M, Pascarella S. Sars-CoV-2 envelope and membrane proteins: Differences from closely related proteins linked to cross-species transmission? Preprints 2020; 2020040089 (doi: 10.20944/preprints202004.0089.v1).

Bornstein SR, Rubino F. et al. Practical recommendations for the management of diabetes in patients with COVID-19. Lancet Diabetes Endocrinol. 2020; DOI:https://doi.org/10.1016/S2213-8587(20)30152-2

Chan JF, Kok KH, Zhu Z, Chu H, To KK, Yuan S, Yuen KY. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg Microbes Infect. 2020; 9:221-236.

Chen LQ, Cheung LS, Feng L, Tanner W, Frommer WB. Transport of sugars. Annu. Rev. Biochem. 2015; 84: 865–894.

Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, Qu XQ, Guo WJ, Kim JG, Underwood W, Chaudhuri B, Chermak D, Antony G, White FF, Somerville SC, Mudgett MB, Frommer WB. Sugar transporters for intercellular exchange and nutrition of pathogens. Nature 2010; 468:527-532.

Chen LQ. SWEET sugar transporters for phloem transport and pathogen nutrition. New Phytol. 2014; 201:1150-1155.

de Haan CA, Smeets M, Vernooij F, Vennema H, Rottier PJ. Mapping of the coronavirus membrane protein domains involved in interaction with the spike protein. J Virol. 1999; 73: 7441-7452.

Dai L, Hu WW, Xia L, Xia M, Yang Q. Transmissible gastroentitis virus infection enhances SGLT1 and GLUT2 expression to increase glucose uptake. PLoS One. 2016; 11(11):e0165585

Deng D, Yan N. GLUT, SGLT, and SWEET: Structural and mechanistic investigations of the glucose transporters. Protein Sci. 2016; 25: 546-558.

Feng L, Frommer WB. Structure and function of semiSWEET and SWEET sugar transporters. Trends Biochem Sci. 2015; 40: 480-486.

Gualdoni GA, Mayer KA, Kapsch AM, Kreuzberg K, Puck A, Kienzl P, Oberndorfer F, Frühwirth K, Winkler S, Blaas D, Zlabinger GJ, Stöckl J. Rhinovirus induces an anabolic reprogramming in host cell metabolism essential for viral replication. Proc Natl Acad Sci U S A. 2018; 115: E7158-E7165.

Ghosh AK, Brindisi M, Shahabi D, Chapman ME, Mesecar AD. Drug development and medicinal chemistry efforts toward SARS-Coronavirus and Covid-19 therapeutics. Chem. Med. Chem. 2020, doi: 10.1002/cmdc.202000223.

Han L, Zhu Y, Liu M, Zhou Y, Lu G, Lan L, Wang X, Zhao Y, Zhang XC. Molecular mechanisms of substrate recognition and transport by the AtSWEET 13 sugar transporter. Proc Natl Acad Sci U S A. 2017; 114:10089-10094.

Jeena GS, Kumar S, Shukla RK. Structure, evolution and diverse physiological roles of SWEET sugar transporters in plants. Plant Mol Biol. 2019; 100:351-365.

Jia B, Hao L, Xuan YH, Jeon CO. New Insight into the Diversity of SemiSWEET Sugar Transporters and the Homologs in Prokaryotes. Front Genet. 2018; 9:180.

Le T, Andreadakis Z, Kumar A, Gómez Román R, Tollefsen S, Saville M, Mayhew S. The COVID-19 vaccine development landscape. Nat Rev Drug Discov. 2020 Apr 9. doi: 10.1038/d41573-020-00073-5.

Lee Y, Nishizawa T, Yamashita K, Ishitani R, Nureki O. Structural basis for the facilitative diffusion mechanism by SemiSWEET transporter. Nat Commun. 2015; 6:6112.

Mousavizadeh L, Ghasemi S. Genotype and phenotype of COVID-19: Their roles in pathogenesis. J Microbiol Immunol Infect. 2020; pii: S1684-1182(20)30082-7.

Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. 2003; 31: 3381-3385.

Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R. COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. J Adv Res. 2020; 24:91-98.

Skrabanek L, Campagne F, Weinstein H. Building protein diagrams on the web with the residue-based diagram editor RbDe. Nucleic Acids Res. 2003; 31: 3856-3858.

Thaker SK, Ch'ng J, Christofk HR. Viral hijacking of cellular metabolism. BMC Biol. 2019; 17:59.

Wang X, Xu W, Hu G, Xia S, Sun Z, Liu Z, Xie Y, Zhang R, Jiang S, Lu L. SARS-CoV-2 infects T lymphocytes through its spike protein-mediated membrane fusion. Cell Mol Immunol. 2020 Apr 7. doi: 10.1038/s41423-020-0424-9.

Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, Graham BS, McLellan JS. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 2020; 367): 1260–1263.

Xuan YH, Hu YB, Chen LQ, Sosso D, Ducat DC, Hou BH, Frommer WB. Functional role of oligomerization for bacterial and plant SWEET sugar transporter family. Proc Natl Acad Sci U S A. 2013; 110: E3685-94.

Yang JK, Lin SS, Ji XJ, Guo LM. Binding of SARS coronavirus to its receptor damages islets and causes acute diabetes. Acta Diabetol. 2010; 47:193-199.

Yu Y, Clippinger AJ, Pierciey FJ Jr, Alwine JC. Viruses and metabolism: alterations of glucose and glutamine metabolism mediated by human cytomegalovirus. Adv Virus Res. 2011; 80:49-67.

## Acknowledgements

The author would like to thank the Women's Board of Lankenau for support. The author also acknowledges Abraham Thomas Foundation for providing the resources.

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

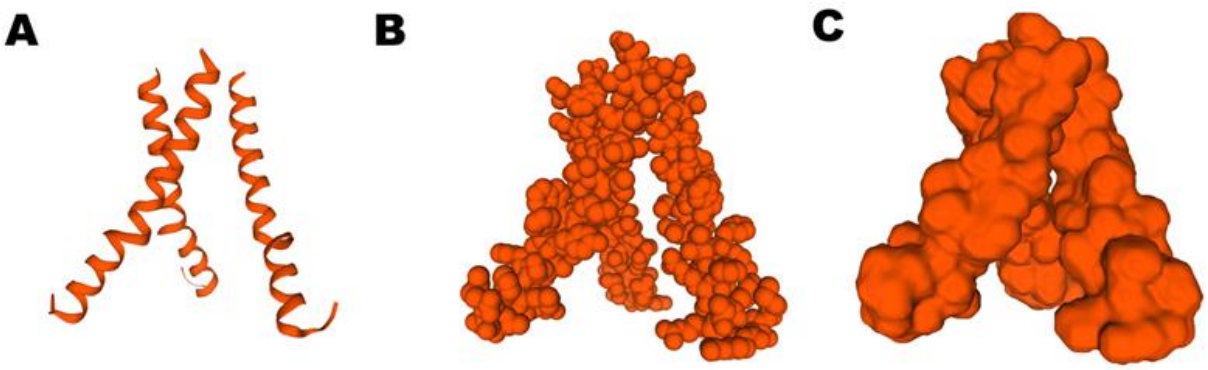
**Fig.1**

**QJA17755.1 membrane glycoprotein**
**[Severe acute respiratory syndrome coronavirus 2]**

1 madsngtitv eelkklleqw nlvigflflt wicllqfaya nrnrflyiik liflwllwpv
61 tlacfvlaav yrinwitggi aiamaclvgl mwlsyfiasf rlfartrsmw sfnpetnill
121 nvplhgtilt rplleselvi gavilrghlr iaghhlgrcd ikdlpkeitv atsrtlsyyk
181 lgasqrvagd sgfaaysryr ignyklntdh ssssdniall vq

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

**Fig. 2**



*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

**Fig. 3A**

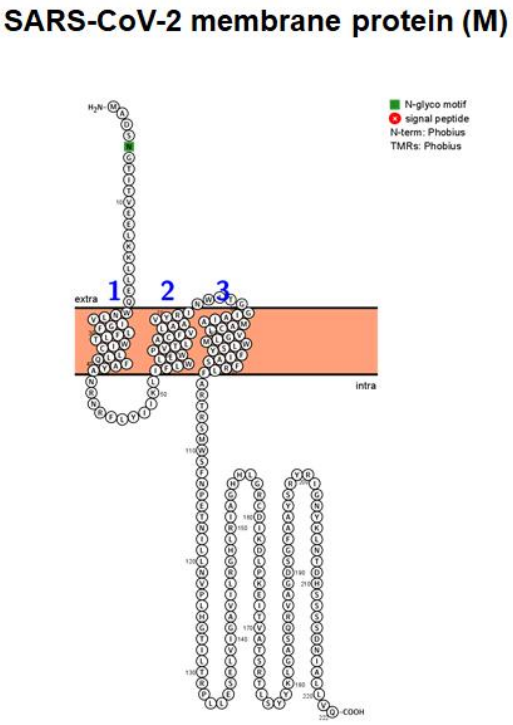
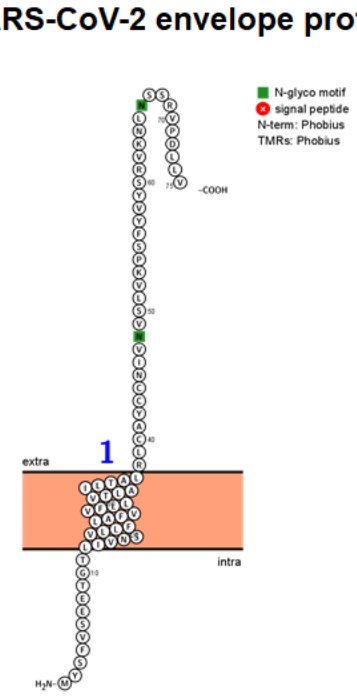



**Fig. 3B**



*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

**Fig. 3C**

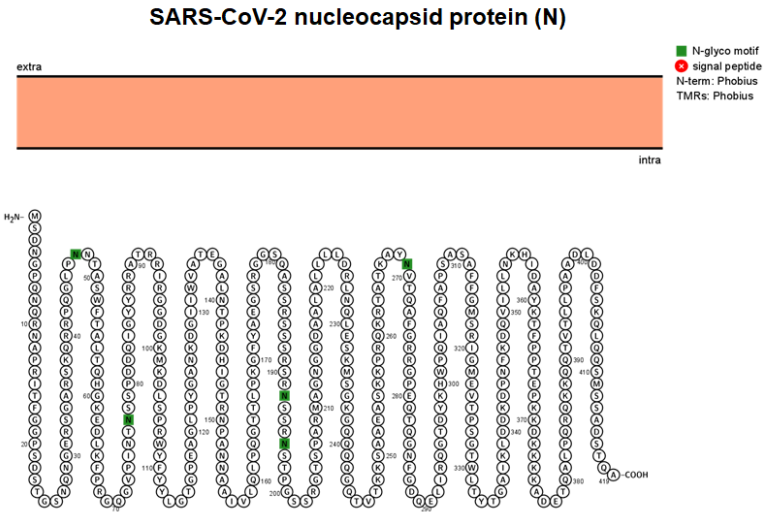
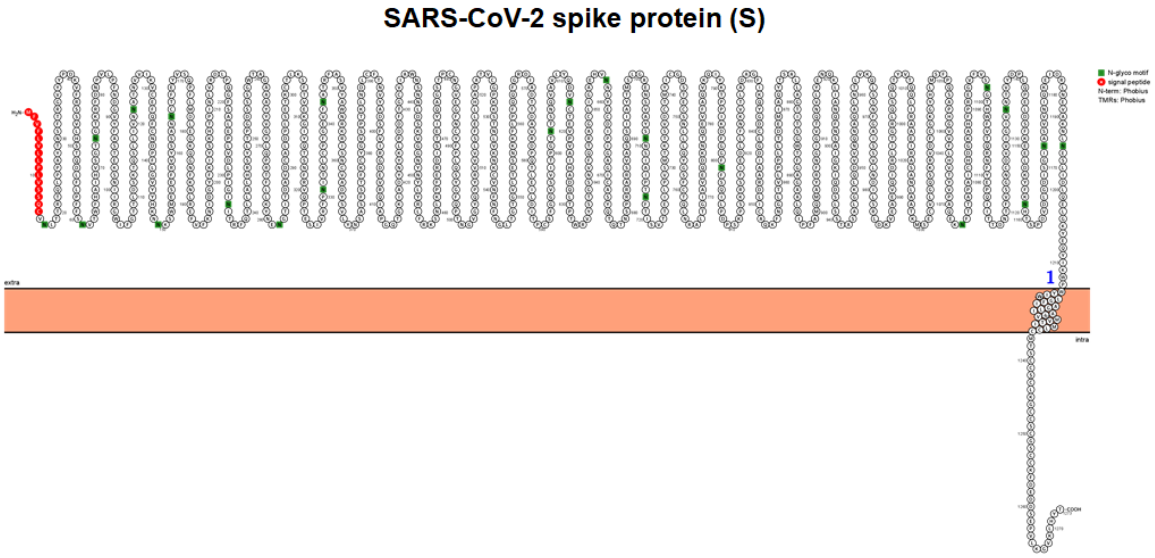



**Fig. 3D**



*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2020 doi:10.20944/preprints202004.0512.v1

**Fig. 4A**

```
CLUSTAL O(1.2.4) multiple sequence alignment

WP_113585511.1      -----------------MNWVTVIGFGAALCSTVSFMPQAWRIVKTRDTSSLSAPMYAIN      43
QJA17755.1          MADSNGTITVEELKKLLEQWNLVIGFLFL----------TWICLLQFAYANRNRFLYIIK      50
                                      :*  ****            :*  :     :. .  :* *:

WP_113585511.1      TIGFMLWLIYGVMLGQWPLI-------LTNGICLVLAAF---------ILTMTLASSK--      85
QJA17755.1          --LIFLWLLWPVTLACFVLAAVYRINWITGGIAIAMACLVGLMWLSYFIASFRLFARTRS      108
                      ::***:: * *. : *        :*.**.:.:*.:         * :: * : .

WP_113585511.1      -----QKAK-------------------------------------------ITDALE--      95
QJA17755.1          MWSFNPETNILLNVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEI      168
                          :::                                           *.*  :

WP_113585511.1      ---------------------------------------------------          95
QJA17755.1          TVATSRTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIALLVQ      222
```

WP_113585511.1: semiSWEET Rhizhobiales
QJA17755.1: Membrane protein SARS-CoV-2
Percent identity: 26.51

**Fig. 4B**

```
CLUSTAL O(1.2.4) multiple sequence alignment

WP_000580383.1      --------MIGSIAAILT------TFAFLPQVFR-------------VVK---TKDTGSI      30
QJA17755.1          MADSNGTITVEELKKLLEQWNLVIGFLFLTWICLLQFAYANRNRFLYIIKLIFLWLLWPV      60
                              : .:  :*        * **  :               ::*         :

WP_000580383.1      ALGMYVMQVIGIALWLDHGIRIGDLPLILANSVSFLLSGI--------------------      70
QJA17755.1          TLACFVLAAVYRINWITGGIAIAMACLVGLMWLSYFIASFRLFARTRSMWSFNPETNILL      120
                    :*. :*: .:    *:  ** *.   *:    :*:::::.:

WP_000580383.1      ------------------------------------------------------ILFYK      75
QJA17755.1          NVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTLSYYK      180
                                                                         : :**

WP_000580383.1      LKYK---------------------------------------  79
QJA17755.1          LGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIALLVQ  222
                    *  .
```

WP_000580383.1: semiSWEET *Streptococcus pneumoniae*
QJA17755.1: Membrane protein SARS-CoV-2
Percent identity: 20.25