 doi:10.20944/preprints202210.0158.v1

Mapping the Accessory Proteins (Orfs) of Severe Acute Respiratory Syndrome Coronavirus 2

Sunil Thomas

Lankenau Institute for Medical Research, 100 East Lancaster Avenue, Wynnewood, Pennsylvania -19096, USA.
*E-mail: thomass-02@mlhs.org; suntom2@gmail.com

ABSTRACT

The COVID-19 pandemic first observed in December 2019 caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has influenced every individual on the planet. The virus has influenced our lifestyle, education, economy and the environment. Though the vaccines against COVID-19 have provided protection against the disease; new strains of the virus have lowered the efficiency of the vaccines. There is still no effective therapy for the treatment of the disease. Understanding the protein structure of the virus may lead to the development of effective therapies for the disease. We recently mapped the structural proteins and non-structural proteins of SARS-CoV-2. The accessory proteins (Open reading frames, Orfs) of SARS-CoV-2 modulate the host environment to favor virus replication. This paper reports mapping the accessory proteins (Orfs) of SARS-CoV-2.

KEYWORDS: severe acute respiratory syndrome coronavirus 2, SARS-CoV-2, COVID-19, accessory protein, mapping, structure, Open reading frame, Orf.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the RNA virus responsible for the coronavirus-19 disease (COVID-19). Millions of people have been infected with COVID-19 since December 2019. As of October 2022, the pandemic has resulted in more than 620 million infections and 6.5 million deaths. The disease has impacted every country of our planet. The virus has infected humans, companions and wild animals and has led to huge economic loss (Thomas, 2020; 2021).

Though there are vaccines to protect against the disease, even after three years, there are no effective therapies for the treatment of the disease (Thomas, 2020). Current vaccines are based on the spike protein (S). There is a risk that the currently available vaccines may not be effective against mutant strains of the virus; hence, there is a need to develop new and potent vaccines to protect against multiple strains of SARS-CoV-2.

It is still not understood the functions of all the proteins of SARS-CoV-2. The structural proteins of SARS-CoV-2 include membrane glycoprotein (M), envelope protein (E), nucleocapsid protein (N), and the spike protein (S). In the host cell, the Orf1ab is cleaved to form 16 nonstructural proteins (nsps) that is involved in viral replication and inhibition of innate immunity (Thomas, 2021). We demonstrated recently that nsp3, nsp4 and nsp6 are transmembrane proteins (Thomas, 2021).

There are several accessory genes, including the Orf3a, Orf6, Orf7a, Orf7b, Orf8, and Orf10 that are interspaced among the structural proteins of SARS-CoV-2. The accessory proteins are involved in viral replication and inhibition of host immune defenses (Redondo et al. 2021). Understanding the structure of the accessory proteins may be useful to design new therapies and vaccines. Development of new strains of SARS-CoV-2 (Variants of Concern, VOC) warrants development of new and effective vaccines. This paper reports mapping the accessory proteins of SARS-CoV-2 that may aid in designing novel therapies, anti-viral and vaccines to protect against COVID-19.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

MATERIALS AND METHODS

2.1. SARS-CoV-2 accessory protein structure

The accessory protein sequences of the SARS-CoV-2 were downloaded from the NCBI (https://www.ncbi.nlm.nih.gov/protein/) protein database. The accessory proteins of SARS-CoV-2 include ORF3a (accession No. QNH88661), ORF6 (accession No. QNH88664), ORF7a (accession No. QWW27599), ORF7b (accession No. BDB04007), ORF8 (accession No. QNH88667) and ORF10 (accession No. QNH88669).

2.2. Protein modeling

To determine the snake diagram model of the SARS-CoV-2 accessory proteins we used Protter (http://wlab.ethz.ch/protter). For the three-dimensional homology modeling of the SARS-CoV-2 accessory proteins we employed the iterative threading assembly refinement (I-TASSER) (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) with default settings. The protein sequences of SARS-CoV-2 was entered in FASTA format.

We used multiple bioinformatics software for the prediction of transmembrane regions of membrane proteins. Transmembrane helix prediction (TMHMM) (transmembrane hidden Markov model) was used for prediction of transmembrane domains (www.cbs.dtu.dk/services/TMHMM/). Phobius was used for prediction of transmembrane topology and signal peptides from the amino acid sequence of proteins (https://phobius.sbc.su.se/) (Kall et al. 2007).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

RESULTS

The evolution of VOC among the SARS-CoV-2 virus warrants development of novel therapies and vaccines to protect against new strains of the virus. As yet there are no therapies for the treatment of COVID-19. Social distancing, wearing mask and vaccination are the only strategies to protect against COVID-19 (Thomas, 2020; Thomas, 2021; Thomas, 2022). There is a need to develop new therapies and vaccines to protect against the disease.

Using bioinformatics tools, we mapped the structure of structural proteins and nsps of SARS-CoV-2. We demonstrated that the membrane (M) protein of the virus could function as a glucose transporter (Thomas, 2020). We also demonstrated that nsp3, nsp4 and nsp6 are transmembrane proteins (Thomas, 2021). In this paper using bioinformatic tools we mapped the structure of the accessory proteins (Orf3a, Orf6, Orf7a, Orf7b, Orf8, Orf10) of SARS-CoV-2.

In our previous paper, we mapped the structure of nsps (Thomas, 2021). We used Orf3a as a reference protein. The Orf3a is a transmembrane protein with three transmembrane domains and a long and short luminal domain jutting into the ER lumen (**Fig. 1**). The bioinformatic tools Protter, TMHMM, and Phobius demonstrated that Orf3a has three transmembrane domains. The data confirmed that the largest accessory protein Orf3a, is a transmembrane protein.

The bioinformatic tools predicted that the accessory protein Orf6 is a cytosolic protein and not a transmembrane protein (**Fig. 2**).

Orf7a is predicted to be a transmembrane protein with a signal peptide at the N-terminus. Orf7a has a single transmembrane domain, a long N-terminal domain and a small C-terminal domain (**Fig. 3**).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

The accessory protein, Orf7b is the smallest transmembrane protein, as predicted by bioinformatics (**Fig. 4**). All the bioinformatics software has predicted Orf7b as a transmembrane protein.

Orf8 is not a transmembrane protein as predicted by different bioinformatics software. The protein has a signal peptide at the N-terminal domain (**Fig. 5**).

The final accessory protein, Orf10 is also not a transmembrane protein. It is the smallest accessory protein (**Fig. 6**).

Our study demonstrates that the accessory proteins Orf3a, Orf7a, and Orf7b are transmembrane proteins.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

DISCUSSION

The COVID-19 disease is the first pandemic disease of the twenty first century that has impacted every family of our planet. The disease is also observed in companion animals and wild animals (Thomas, 2022). The pandemic has resulted in travel bans, change in lifestyle, ban on crowding, maintaining safe distance, and wearing mask while travelling outside home. The disease has negatively impacted the economy of all the countries (Thomas, 2020; Thomas, 2022). Development of vaccines against the disease has provided relief; however, arise of VOC of SARS-SoV-2 warrants development of new therapies and vaccines to protect against the virus. It is still not understood why SARS-CoV-2 turned out to be a successful virus causing pandemic.

The bioinformatics software available online are important tools that aid in protein modelling and to a limited extend unravel the function. We used the tools to determine the protein structure of the structural proteins S, M, E and N (Thomas, 2020). We demonstrated that the M protein resembles a glucose transporter (Thomas, 2020). Later, we modeled the nsps of SARS-CoV-2. We showed that nsp3, nsp4 and nsp6 are transmembrane proteins (Thomas, 2021). In this paper, using bioinformatics tools we model the accessory proteins (Orfs) of SARS-CoV-2. We demonstrate that the SARS-CoV-2 accessory proteins Orf3a, Orf7a, and Orf7b are transmembrane proteins.

The role of SARS-CoV-2 accessory proteins in viral pathogenesis and viral replication is not completely understood (Redondo et al. 2021). The accessory proteins work as antagonists of IFN-signaling during viral infection, viral pathogenesis, apoptotic inducers, or antiviral suppressors (Redondo et al. 2021).

The accessory protein ORF3a induces cell death through apoptosis, necrosis, and pyroptosis leading to tissue damage of the host (Thomas, 2021; Zhang et al. 2022). In addition, ORF3a could trigger cytokine storm to promote pro-inflammatory cytokines and chemokines (Bianchi et al. 2021; Zhang et al. 2022). Our study demonstrated that Orf3a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

is a transmembrane protein. ORF3a is a viroporin that activates the NLRP3-inflammasome that contributes to virus release (Xu et al. 2022).

ORF6 proteins can block interferon signaling (Lee et al. 2021; Miorin et al. 2020; Miyamoto et al. 2022), can act as a virulence factor that modulates nucleocytoplasmic trafficking to accelerate viral replication, progressing to disease (Miyamoto et al. 2022). Lee et al. (2021) reported that Orf6 is localized to the endoplasmic reticulum, autophagosome and lysosomal membranes. Wong et al. (2022) were of the opinion that ORF6 is a peripheral membrane protein, as opposed to being a transmembrane protein. Our study also confirms that that Orf6 is not a transmembrane protein.

ORF7a is an immunomodulating factor for immune cell binding and triggers dramatic inflammatory responses (Zhou et al. 2021). ORF7a efficiently binds to CD14+ monocytes and contributes to the recruitment of monocytes to infected lungs during COVID-19. ORF7a may suppress the antigen-presenting ability of these monocytes (Zhou et al. 2021). Our study demonstrate that Orf7a is a transmembrane protein.

The accessory protein ORF7b promotes expression of inflammatory cytokines that may induce apoptosis (Yang et al. 2021). ORF7b interfere with important cellular processes that involve leucine-zipper formation and contribute to heart arrythmias, odor loss, impaired oxygen uptake and intestinal dysfunction (Fogeron et al. 2021). Our study using bioinformatics tools demonstrate that Orf7b is a transmembrane protein.

Orf8 is an accessory protein that has been proposed to interfere with immune responses (Flower et al. 2020). Orf8 promotes the expression of pro-inflammatory factors thereby acting as a contributing factor to cytokine storm during COVID-19 infection. Orf8 is also not a transmembrane protein as determined by bioinformatics analysis.

The accessory protein, ORF10 suppress the expression of type I interferon (IFN-I) genes and IFN-stimulated genes (Li et al. 2022). ORF10 is known to impair cilia function thereby leading to loss of smell and taste that are symptoms of COVID-19 (Wang et al. 2022).

 doi:10.20944/preprints202210.0158.v1

Hassan et al. (2022) predicted that Orf10 is a non-cytoplasmic protein. Our bioinformatics analysis also confirm that Orf10 is a non-cytoplasmic protein.

Overall, our bioinformatic analyses demonstrate that the accessory protein Orf3a, Orf7a and Orf7b are transmembrane proteins. Our protein modeling could lead to the development of better therapies and vaccines.

 doi:10.20944/preprints202210.0158.v1

REFERENCES

Bianchi M, Borsetti A, Ciccozzi M, Pascarella S. SARS-Cov-2 ORF3a: Mutability and function. Int J Biol Macromol. 2021; 170:820-826.

Flower TG, Buffalo CZ, Hooy RM, Allaire M, Ren X, Hurley JH. Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. Proc Natl Acad Sci U S A. 2021; 118(2): e2021785118.

Fogeron M-L, Montserret R, Zehnder J, Nguyen M-H, Dujardin M, Brigandat L, et al. SARS-CoV-2 ORF7b: Is a Bat Virus Protein Homologue a Major Cause of COVID-19 Symptoms? bioRxiv [Preprint] (2021) 2021.02.05.428650.

Hassan SS, Lundstrom K, Serrano-Aroca Á, Adadi P, Aljabali AAA, Redwan EM, Lal A, Kandimalla R, El-Aziz TMA, Pal Choudhury P, Azad GK, Sherchan SP, Chauhan G, Tambuwala M, Takayama K, Barh D, Palu G, Basu P, Uversky VN. Emergence of unique SARS-CoV-2 ORF10 variants and their impact on protein structure and function. Int J Biol Macromol. 2022 Jan 1;194:128-143.

Käll L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. Nucleic Acids Res. 2007; 35(Web Server issue):W429-32.

Lee, JG., Huang, W., Lee, H. et al. Characterization of SARS-CoV-2 proteins reveals Orf6 pathogenicity, subcellular localization, host interactions and attenuation by Selinexor. Cell Biosci 11, 58 (2021).

Li, X., Hou, P., Ma, W. *et al.* SARS-CoV-2 ORF10 suppresses the antiviral innate immune response by degrading MAVS through mitophagy. Cell Mol Immunol**.** 2022; 19: 67–78.

Miorin L, Kehrer T, Sanchez-Aparicio MT, Zhang K, Cohen P, Patel RS, Cupic A, Makio T, Mei M, Moreno E, Danziger O, White KM, Rathnasinghe R, Uccellini M, Gao S, Aydillo T, Mena I, Yin X, Martin-Sancho L, Krogan NJ, Chanda SK, Schotsaert M, Wozniak RW, Ren Y, Rosenberg BR, Fontoura BMA, García-Sastre A. SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. Proc Natl Acad Sci U S A. 2020 Nov 10;117(45):28344-28354.

Miyamoto, Y., Itoh, Y., Suzuki, T. et al. SARS-CoV-2 ORF6 disrupts nucleocytoplasmic trafficking to advance viral replication. Commun Biol. 2022; 5: 483.

Redondo N, Zaldívar-López S, Garrido JJ, Montoya M. SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns. Front Immunol. 2021; 12:708264.

Thomas S. The Structure of the Membrane Protein of SARS-CoV-2 Resembles the Sugar Transporter SemiSWEET. Pathog Immun. 2020; 5(1):342-363.

Thomas S. Mapping the Nonstructural Transmembrane Proteins of Severe Acute Respiratory Syndrome Coronavirus 2. J Comput Biol. 2021; 28(9):909-921

Thomas S. Status of COVID-19 Pandemic Before the Administration of Vaccine. Methods Mol Biol. 2022; 2410:93-108.

Wang L, Liu C, Yang B, Zhang H, Jiao J, Zhang R, Liu S, Xiao S, Chen Y, Liu B, Ma Y, Duan X, Guo Y, Guo M, Wu B, Wang X, Huang X, Yang H, Gui Y, Fang M, Zhang L, Duo S, Guo X, Li W. SARS-CoV-2 ORF10 impairs cilia by enhancing CUL2ZYG11B activity. J Cell Biol. 2022; 221(7):e202108015.

Wong HT, Cheung V, Salamango DJ. Decoupling SARS-CoV-2 ORF6 localization and interferon antagonism. J Cell Sci. 2022; 135(6):jcs259666.

Xu H, Akinyemi IA, Chitre SA, Loeb JC, Lednicky JA, McIntosh MT, Bhaduri-McIntosh S. SARS-CoV-2 viroporin encoded by ORF3a triggers the NLRP3 inflammatory pathway. Virology 2022; 568:13-22.

Zhang J, Ejikemeuwa A, Gerzanich V, Nasr M, Tang Q, Simard JM, Zhao RY. Understanding the Role of SARS-CoV-2 ORF3a in Viral Pathogenesis and COVID-19. Front Microbiol. 2022; 13:854567.

Zhou Z, Huang C, Zhou Z, Huang Z, Su L, Kang S, Chen X, Chen Q, He S, Rong X, Xiao F, Chen J, Chen S. Structural insight reveals SARS-CoV-2 ORF7a as an immunomodulating factor for human CD14+ monocytes. iScience 2021; 24:102187.

ACKNOWEDGEMENTS

The author acknowledges Abraham Thomas Foundation and Women's Board of Lankenau for providing the resources for this work.

CONFLICT OF INTEREST

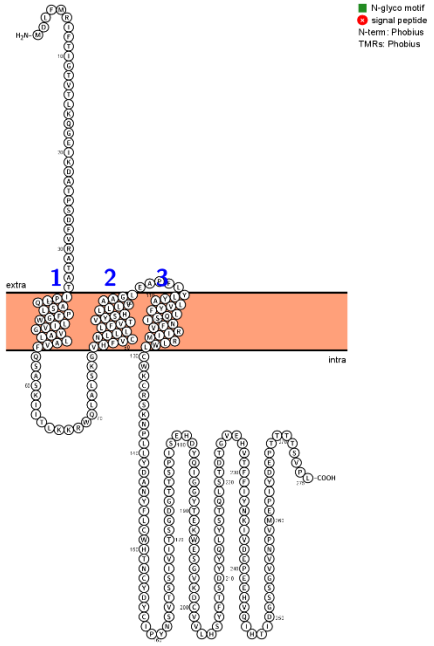
The author declares no conflict of interest.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

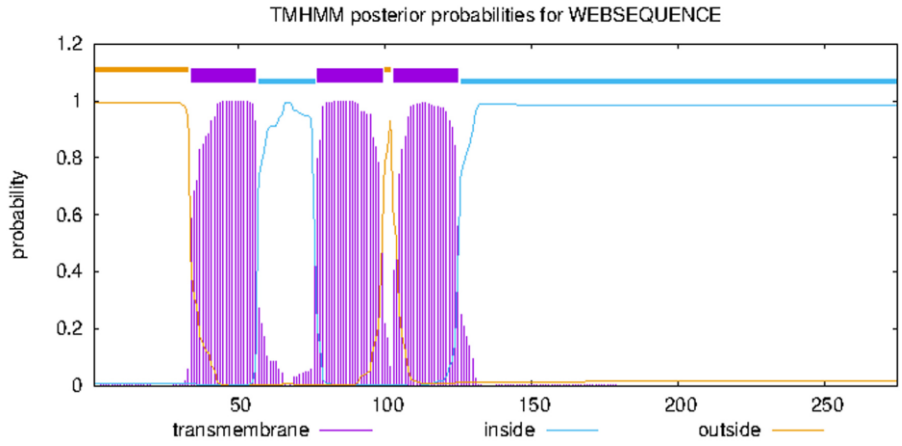
FIGURES

Figure 1. The topology of the accessory protein Orf3a of SARS-CoV-2. (A) The topology of Orf3a of SARS-CoV-2 determined using Protter. (B) TMHMM of Orf3a. (C) The topology of Orf3a as determined by Phobius database (D) The domains (cytoplasmic, transmembrane, and luminal) of Orf3a of SARS-CoV-2. (E) The predicted Orf3a protein structure of SARS-CoV-2 (ribbon diagram) determined using the software I-TASSER.

A

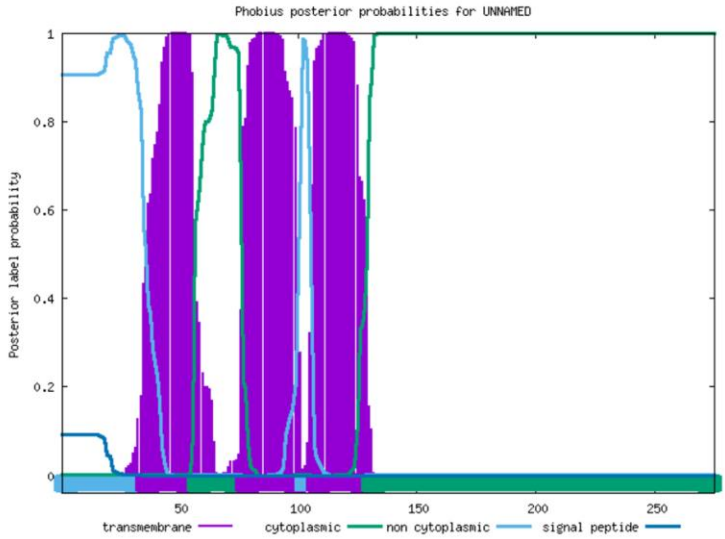


B



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

C

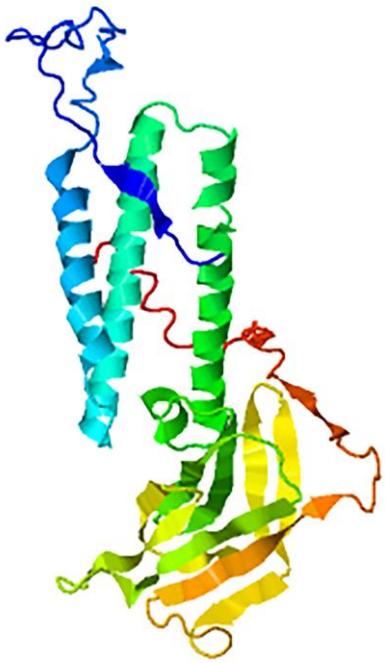


D

MDLFMRIFTIGTVTLKQGEIKDATPSDFVRATAT IPIQASLPFGWLIVGVALLAVF QSASKIITLKKRWQ LALSKG VHFVCNLLLLFVTVYSHLLLVAAGL EAPFL YLYALVYFLQSINFVRIIMRLWL CWKCRSKNPLL YDANYFLCWHTNCYDYCIPYNSVTSSIVITSGDGTTSPISEHDYQIGGYTEKWESGVKDCVVLHSYFTSD YYQLYSTQLSTDTGVEHVTFFIYNKIVDEPEEHVQIHTIDGSSGVVNPVMEPIYDEPTTTTSVPL

Yellow: Luminal domain
Blue: Transmembrane domain
Green: Cytoplasmic domain

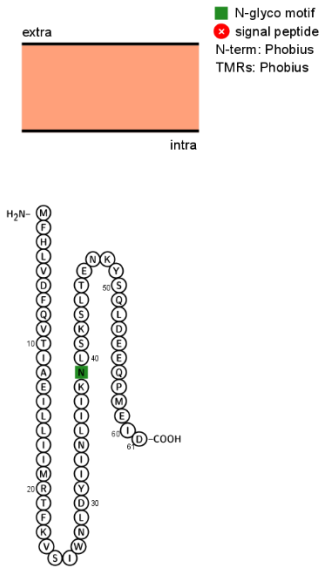
E



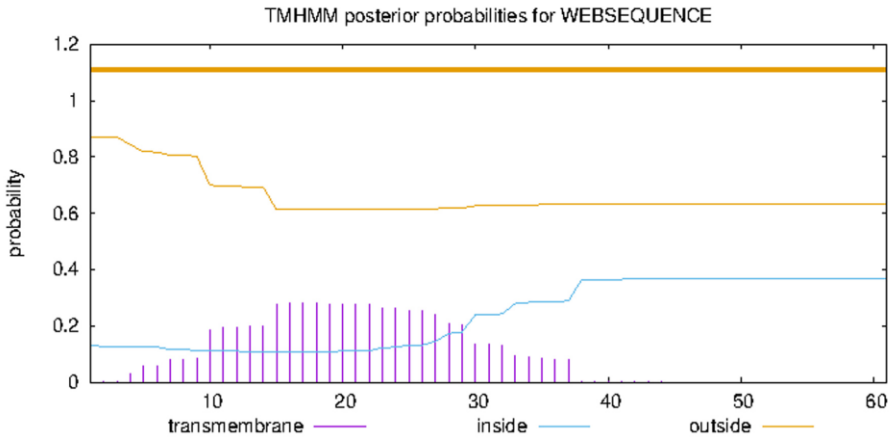
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

Figure 2. The topology of the accessory protein Orf6 of SARS-CoV-2. (A) The topology of Orf6 of SARS-CoV-2 determined using Protter. (B) TMHMM of Orf6. (C) The topology of Orf6 as determined by Phobius database (D) The domains (cytoplasmic, transmembrane, and luminal) of Orf6 of SARS-CoV-2. (E) The predicted Orf6 protein structure of SARS-CoV-2 (ribbon diagram) determined using the software I-TASSER.

A

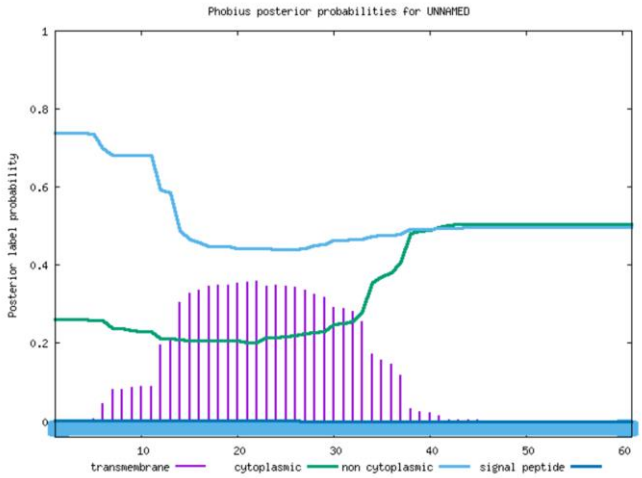


B



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

C

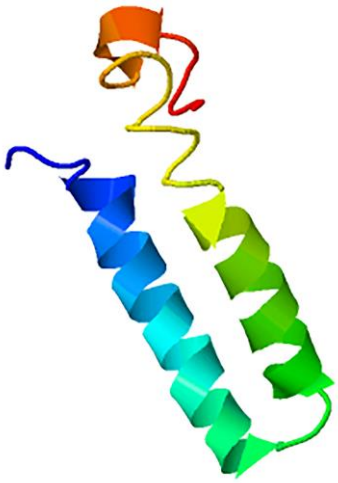


D

MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIIKNLSKS
LTENKYSQLDEEQPMEID

Yellow: Luminal domain
Blue: Transmembrane domain
Green: Cytoplasmic domain

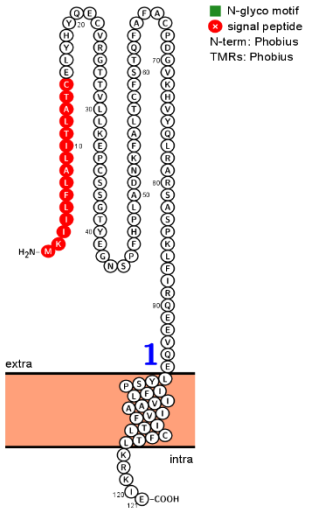
E



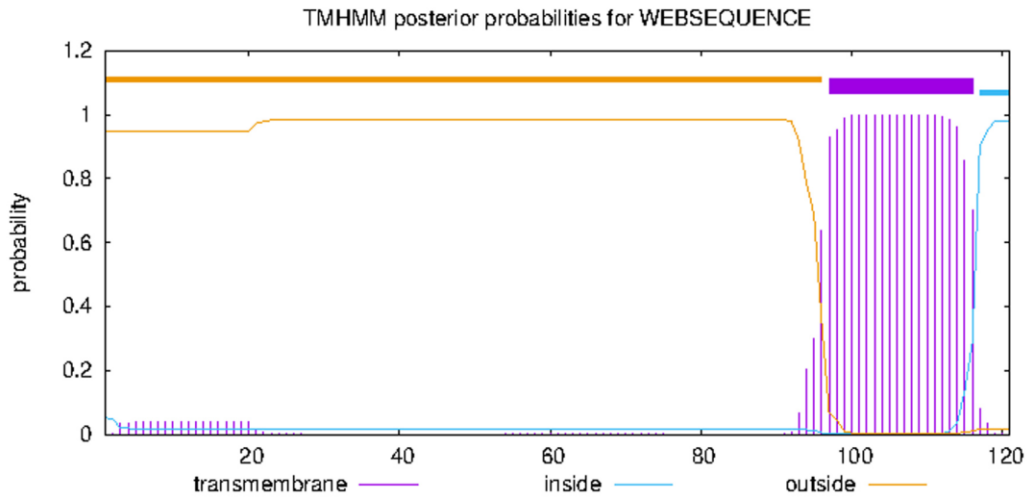
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

Figure 3. The topology of the accessory protein Orf7a of SARS-CoV-2. (A) The topology of Orf7a of SARS-CoV-2 determined using Protter. (B) TMHMM of Orf7a. (C) The topology of Orf7a as determined by Phobius database (D) The domains (cytoplasmic, transmembrane, and luminal) of Orf7a of SARS-CoV-2. (E) The predicted Orf7a protein structure of SARS-CoV-2 (ribbon diagram) determined using the software I-TASSER.

A

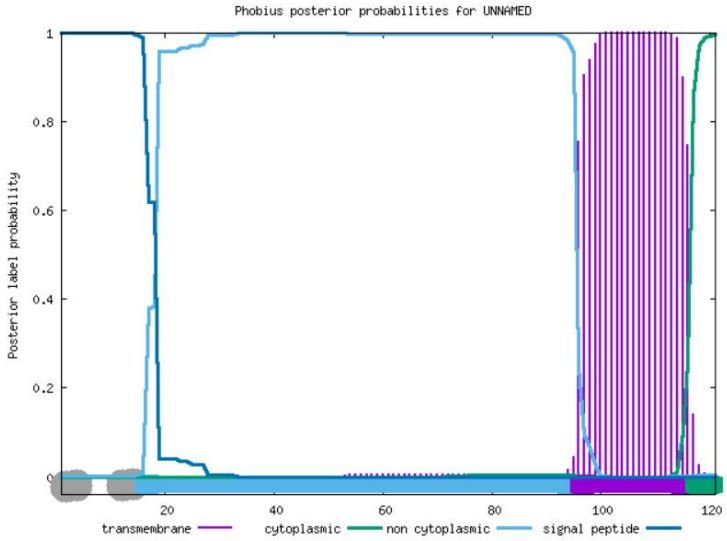


B



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

C

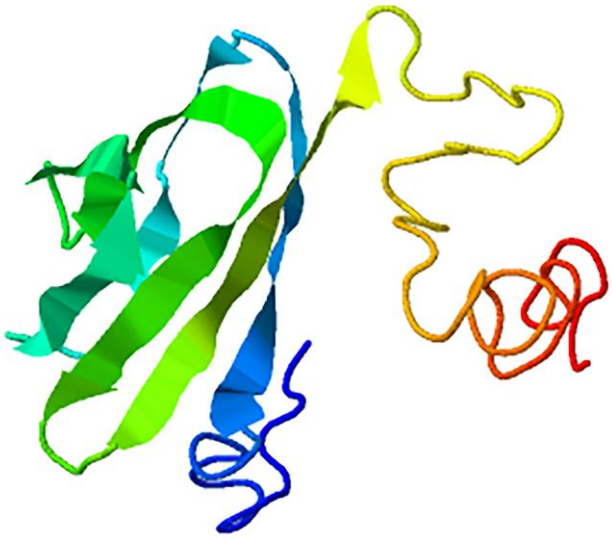


D

MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNSPFHPLADNKFALTCFSTQFAFACPDG
VKHVYQLRARSASPKLFIRQEEVQE LYSPIFLIVAAIVFITLCFTL KRKIE

Yellow: Luminal domain
Blue: Transmembrane domain
Green: Cytoplasmic domain

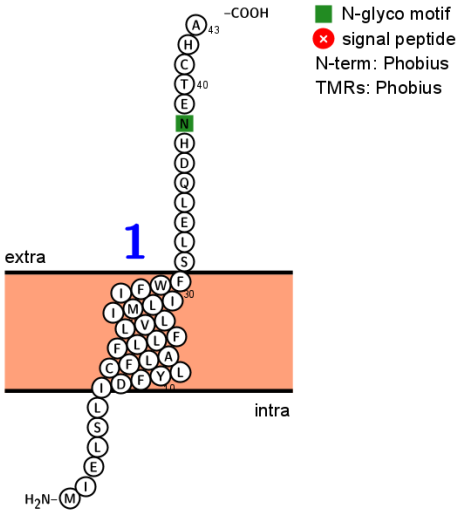
E



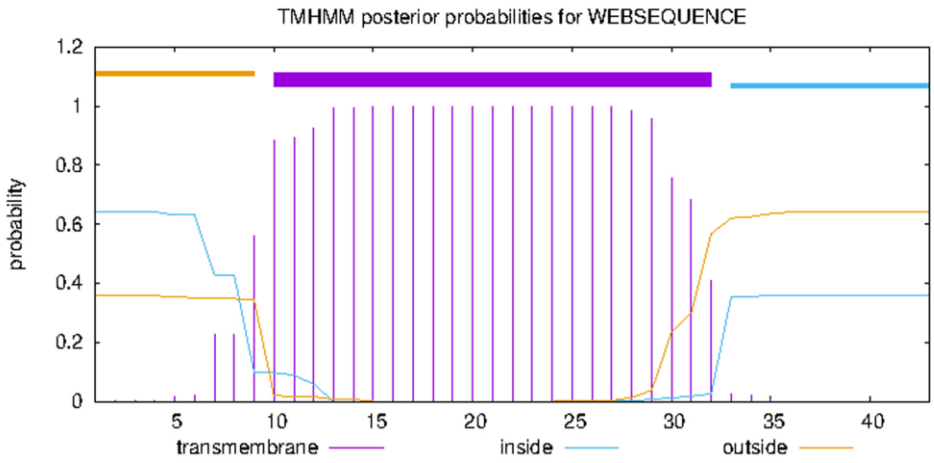
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

Figure 4. The topology of the accessory protein Orf7b of SARS-CoV-2. (A) The topology of Orf7b of SARS-CoV-2 determined using Protter. (B) TMHMM of Orf7b. (C) The topology of Orf7b as determined by Phobius database (D) The domains (cytoplasmic, transmembrane, and luminal) of Orf7b of SARS-CoV-2. (E) The predicted Orf7b protein structure of SARS-CoV-2 (ribbon diagram) determined using the software I-TASSER.

A

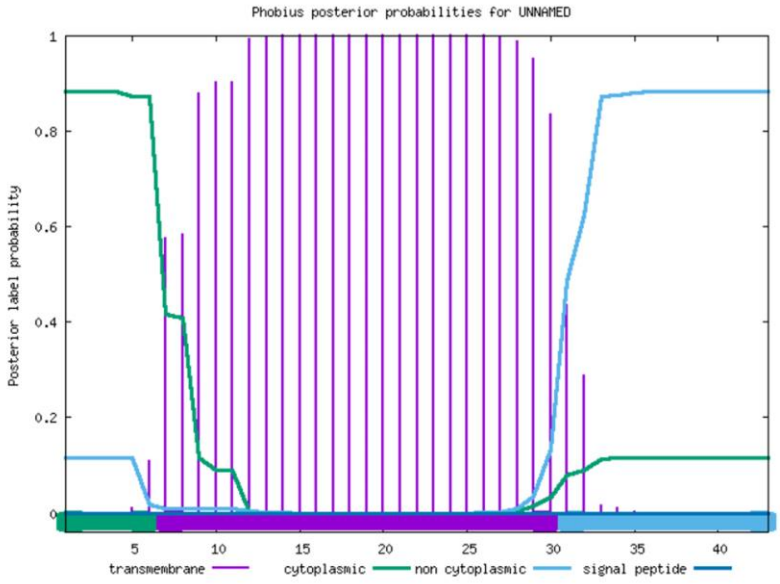


B



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

C

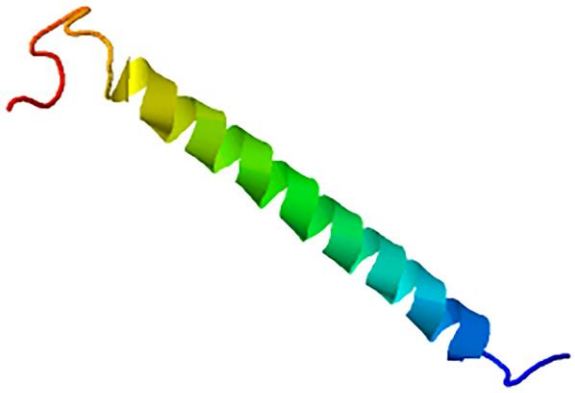


D

MIELSLIDFYLCFLAFLLFLVLIMLIIFWFSLELQDHNETCHA

Yellow: Luminal domain
Blue: Transmembrane domain
Green: Cytoplasmic domain

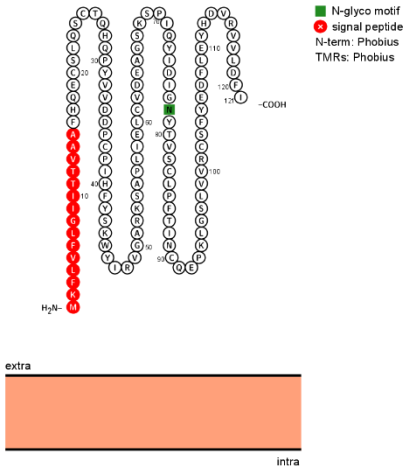
E



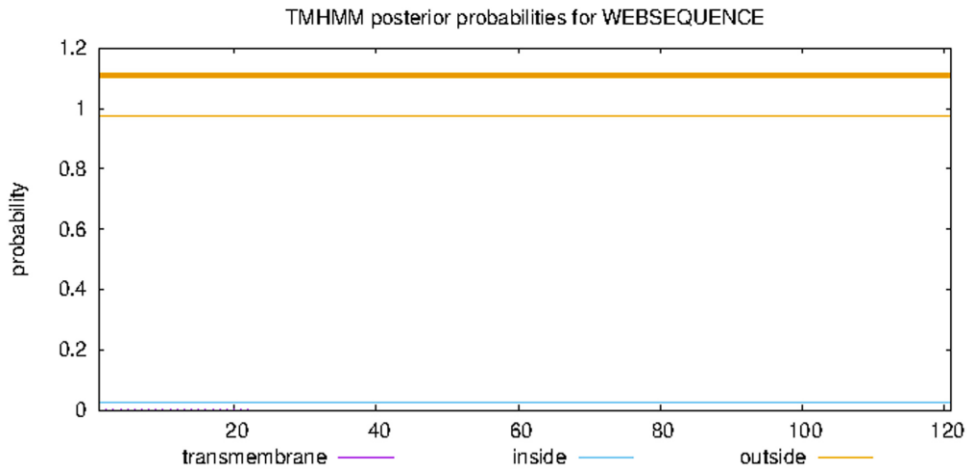
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

Figure 5. The topology of the accessory protein Orf8 of SARS-CoV-2. (A) The topology of Orf8 of SARS-CoV-2 determined using Protter. (B) TMHMM of Orf8. (C) The topology of Orf8 as determined by Phobius database (D) The domains (cytoplasmic, transmembrane, and luminal) of Orf8 of SARS-CoV-2. (E) The predicted Orf8 protein structure of SARS-CoV-2 (ribbon diagram) determined using the software I-TASSER.

A

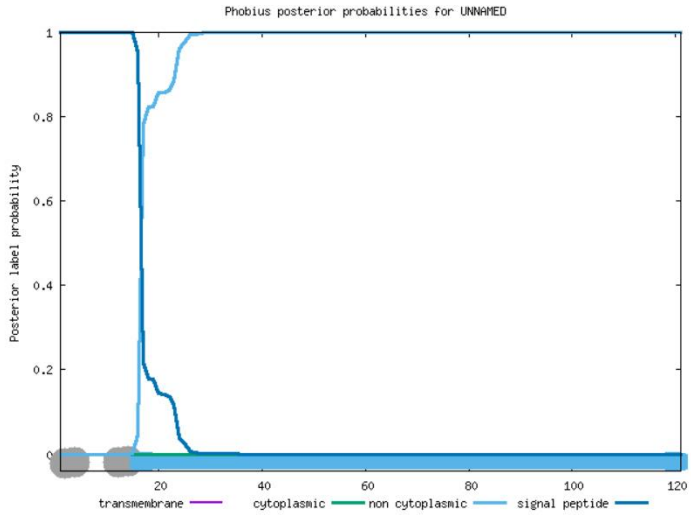


B



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

C

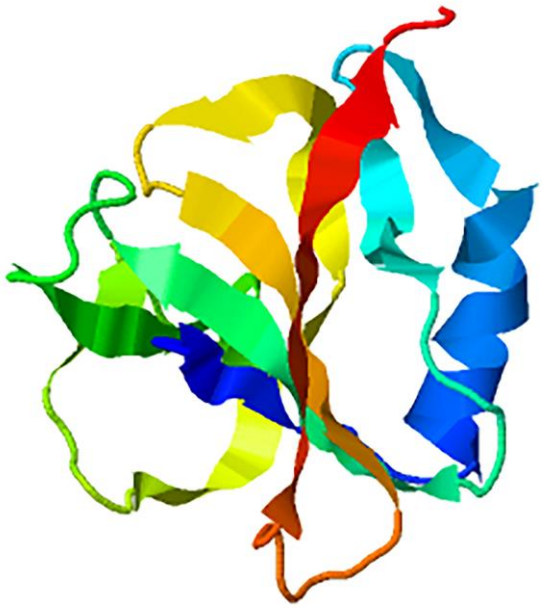


D

MKFLVFLGIITTVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEA
GSKSPIQYIDIGNYTVSCLPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRVVLDFI

Yellow: Luminal domain
Blue: Transmembrane domain
Green: Cytoplasmic domain

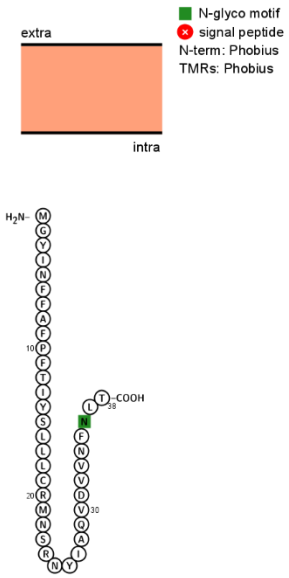
E



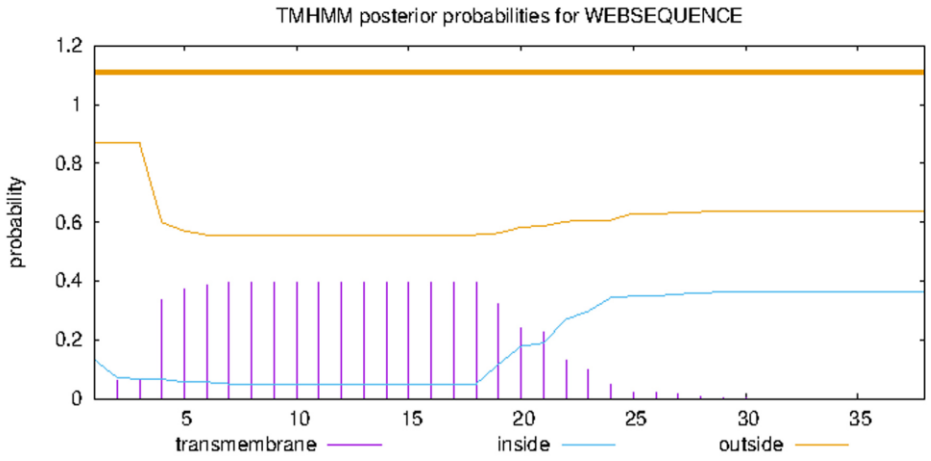
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

Figure 6. The topology of the accessory protein Orf10 of SARS-CoV-2. (A) The topology of Orf10 of SARS-CoV-2 determined using Protter. (B) TMHMM of Orf10. (C) The topology of Orf10 as determined by Phobius database (D) The domains (cytoplasmic, transmembrane, and luminal) of Orf10 of SARS-CoV-2. (E) The predicted Orf10 protein structure of SARS-CoV-2 (ribbon diagram) determined using the software I-TASSER.

A

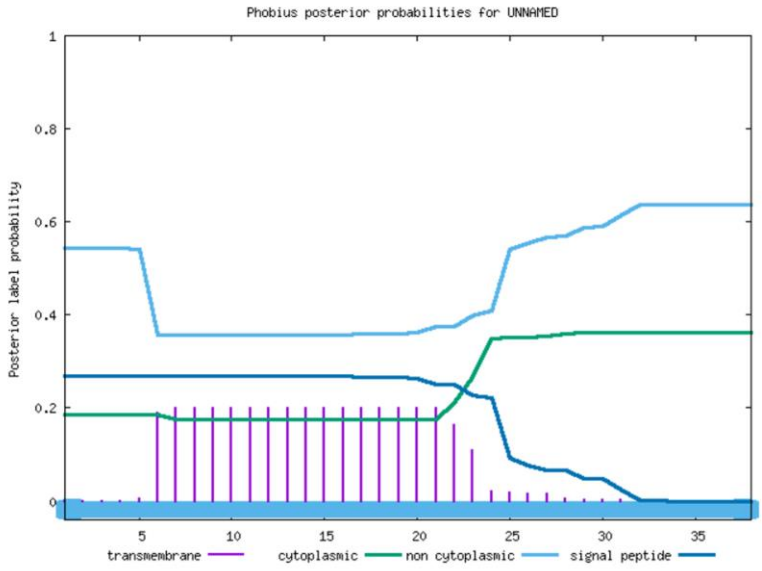


B



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 October 2022 doi:10.20944/preprints202210.0158.v1

C



D

MGYINFFAFPFTIYSLLLCRMNSRNYIAQVDVVNFNLT

Yellow: Luminal domain
Blue: Transmembrane domain
Green: Cytoplasmic domain

E