Importance of SARS-CoV-2 Protein Functions

Ivan Ho Grade 11 Bayview Glen School

Abstract: Severe acute respiratory syndrome coronavirus (SARS-CoV-2), causes the coronavirus disease 2019 (Covid-19). This wide-spreading virus, SARS-CoV-2, is classified as the seventh coronavirus known to humans. It was first reported in Wuhan, China, in December 2019 (Zhou et al., 2020); it rapidly became a pandemic with devastating effects. Since the discovery of the virus in late 2019, the coronavirus crisis was officially announced "pandemic" globally in 2020 by the World Health Organization. Our world and society have been changed everlastingly. It has caused immense distress to our health systems, economies, and clinical care and is still a prominent issue in our society today. In this paper, we will be addressing the viral proteins presented inside SARS-CoV-2, specifically the 16 non-structural proteins and structural proteins.

The genome of the SARS-CoV-2 is 30kb and consists of three sections: (1) ORF1ab protein (16 non-structural proteins (NSP1-16)), (2) structural protein (including Spike, Envelope, Membrane, and Nucleocapsid protein). (3) accessory protein (ORF3a, ORF6, ORF7a, ORF8, ORF9b/c, and ORF10). (reference)

(1) The ORF1ab protein is the first viral protein created in the replication step inside the infected host cell and it consists of a bond of 16 non-structural proteins (NSP).

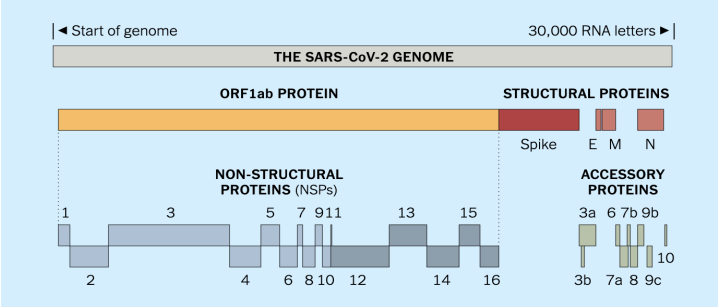

fig1(https://www.nytimes.com/interactive/2020/04/03/science/coronavirus-genome-bad-news-wrapped-in-protein.html)

The NSP1, the first non-structural proteins in the coronavirus genome, decreases the infected cell's protein production rate. This function inhibits the cell's ability to assemble its antiviral proteins, which are used to stop the virus. Since there is a decreased amount of antiviral proteins produced, the virus can execute its tasks without adversity. (NEWYORKTIMES)

The functions of NSP2, an endosmal-associated protein, are still obscure as we do not have sufficient information (https://www.nature.com/articles/s41392-022-00884-5). Although, scientists think that the proteins around it possibly hint that the NSP2 is involved in biological processes like endosomal support (NEWYORKTIMES). More biological processes include viral replication, host immune regulation, and mitochondrial biogenesis (https://www.nature.com/articles/s41392-022-00884-5).

NSP3 is substantially longer than the majority of other non-structural proteins, and has many parts each known as a domain. It is the largest primarily because it has many functions. Most importantly there are two main functions that it must carry out. The first is to cut and separate the parts of the NSP3 protein using specific scissor like molecules of the NSP3 called papain-like protease domains, as this further allows each of the NSP3 proteins to play their role properly. The other is to inhibit the host's natural capacity for protein synthesis by changing several host cell proteins and

This process alters the protein balance and lessens the host cell's resistance to the virus. NSP4 is a membrane-spanning protein that functions as a factory that creates fluid-filled bubbles from the copies of the viral building blocks(newyorktimes). NSP4 plays a role in modifying the endoplasmic reticulum (ER) membranes by anchoring the viral replication-transcription complex, which includes NSP3, NSP4, and NSP6. Together, this complex forms the double membrane vesicles that are involved in the replication and transcription of the RNA (frontersin). It plays a crucial part in both replication and the construction of new replicative structures that enable the creation of different viral clones.(ncbi link)

NSP5 the main protease functions similarly to NSP 3; it acts as a scissor separating the proteins linked together in this genome, allowing each protein to perform its duty. NSP 5 hydrolyzes viral polyprotein 1ab and recognizes over 11 cleavage sites, therefore playing an essential role in the viral replication and maturation of nonstructural proteins (https://www.nature.com/articles/s41392-022-00884-5).

NSP6 has three leading roles as identified by Simona Ricciardi et al.: First, it acts as a filter between the replication organelle and the endoplasmic reticulum by allowing lipid flow but restricting the access of ER luminal proteins to the DMV's (double membrane vesicles). NSP6 also plays a role in creating, positioning, and organizing DMB clusters, and NSP6 mediates contact with lipid droplets(FINDDDD LINK).

NSP7 and 8 assist the NSP12; NSP12 serves as the RNA-dependent RNA polymerase (RdRp) and NSP7/NSP8 as the cofactors with priimase activity that help mediate the RdRp activity(https://www.nature.com/articles/s41392-022-00884-5). These three together forms the replication machinery that helps replicate the RNA genome, which will then be stored inside the new copies of the virus (NEWYORKTIMES).

NSP9 is an essential protein with RNA binding activity. This protein infiltrates the tiny channels of the infected cell's nucleus or at the heart of the cell, which holds the host's DNA genome and attaches its RNA to it for replication. It can also influence the movement of molecules in and out of the nucleus, yet the purpose is still elusive (NEWYORKTIMES). NSP9 of SARS-CoV and SARS-Cov-2 are almost identical in sequence as they share 97% sequence identity, both belonging to the oligonucleotide binding family (OB-fold). The NSP9 plays an important role in viral replication by participating in the formation of the replication and transcription complex (RTC)(https://www.nature.com/articles/s41392-022-00884-5).

NSP10 conceals the presence of the viral RNA to prevent the host immune response andantiviral proteins from attacking them. NSP10 also plays a significant role in regulating replicate function. It also acts as a cofactor of NSP14 and NSP16 and helps to activate the exoribonuclease.

NSP11 is a minor non-structural protein (13 amino acids long) and is an intrinsically disordered protein with unknown functional status. (FIND LINKKK)

NSP12 is the RNA-dependent RNA polymerase, and it works in complex with NSP7 and NSP8 to replicate the RNA letters which make up the newly replicated genome.

NSP13 is a helicase enzyme that can unwind nucleic acid strands using ATP hydrolysis energy.(FIND LINKKK) The helicase is essential as the double helix must be separated for copying and genome replication.

NSP14 is the proofreader. As NSP12 duplicates the coronavirus genome, it may sometimes add incorrect letters into the new copy resulting in a dysfunctional coronavirus genome. NSP14 function is proofreading and eliminating these errors so that the correct letter can be added instead(NEWYORKTIMES).

NSP 15 is a specific endoribonuclease, and it protects the viral RNA by concealing it and evading detection from the host defence system.

Finally, the NSP 16, a ribose 2'-0-methylthransferase plays a role in immune evasion.

Structural Proteins

At the end of the genomic sequence, we have the final four structural proteins which contribute to the cell shape and even movement.

The spike (S) protein of virion is a glycoprotein, known for it's crownlike spikes on the surface of the virus, hence the virus name "corona". S protein has 2 subunits, S1 and S2, both involved in the virus's ability to enter host cells. The S1 generally controls the extension of the spikes during the spike proteins interaction with membrane bound ACE2 (Angiotensin conversion enzyme 2) receptor. The S2 mediates the fusion process between the viral and host cell membrane. S protein plays a major role in the host immune response and it has the potential to

substantially alter the viral pathogenesis if it undergoes mutation. (ncbi.nlm.nih.gov/pmc/articles/PMC8067447/)

The Membrane (M) protein is a glycoprotein and is the most abundant of the structural protein (25-30kDa) as it is the main component of the viral envelope. Its main function is to protect the virus nucleocapsid from damage, helps assemble virus particles and plays a role in pathogenesis. The M protein's structure is generally conserved among all the coronavirus types, except the β-CoVs and δ-CoVs which shows a different type of glycosylation. Glycosylation is important for determining the organs that the virus can infect and IFN signalling. (ncbi.nlm.nih.gov/pmc/articles/PMC8067447/)

The Envelope (E) protein is the smallest structural protein in Sars-CoV-2. Although it is abundantly produced within the infected host cells during the replication process, only a small portion of it is used to form the viral envelope.(link for above sentence) The main functio is it's important role in virus assembly and release, as the majority of the protein in the cell is located near the endoplasmic reticulum (ER), virus infectivity, and pathogenesis. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8067447/)

The Nucleocapsid (N) protein, a complex RNA-binding protein, is what makes up the structure of the virus's nucelocapsid and the viral protein coat that surrounds the genome protects the RNA from the harsh environment within the host cell(cant find source, I think its from Dr Alina slides). Its main functions include RNA packaging, virion assembly and virus transcription efficiency. Due to its highly immunogenic nature, N protein can be targeted for vaccine development. Additionally, the N protein controls the host-pathogen interactions, apoptosis and cell proliferation. (https://virologyj.biomedcentral.com/articles/10.1186/s12985-023-01968-6)

To summarize, this article gives an introduction to the genome of the SARS-CoV-2 and specifically goes into the main types of proteins; structural, nonstructural each of their roles that allow for the virus to be so successful at attacking immune systems.

Major Reference:

1.**A pneumonia outbreak associated with a new coronavirus of probable bat origin** Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL. https://www.nature.com/articles/s41586-020-2012-7

2. **A pneumonia outbreak associated with a new coronavirus of probable bat origin** Peng Zhou #1, Xing-Lou Yang #1, Xian-Guang Wang #2, Ben Hu 1, Lei Zhang 1, Wei Zhang 1, Hao-Rui Si 1 3, Yan Zhu 1, Bei Li 1, Chao-Lin Huang 2, Hui-Dong Chen 2, Jing Chen 1 3, Yun Luo 1 3, Hua Guo 1 3, Ren-Di Jiang 1 3, Mei-Qin Liu 1 3, Ying Chen 1 3, Xu-Rui Shen 1 3, Xi Wang 1 3, Xiao-Shuang Zheng 1 3, Kai Zhao 1 3, Quan-Jiao Chen 1, Fei Deng 1, Lin-Lin Liu 4, Bing Yan 1, Fa-Xian Zhan 4, Yan-Yi Wang 1, Geng-Fu Xiao 1, Zheng-Li Shi 5Nature. 2020 Mar;579(7798):270-273. PMID: 32015507

-