The Mechanism of SARS-CoV-2 Entry into Human Cells and the Discovery of Multiple Viral Strains

Gabriele Kiana Jamain

Biochemistry Lab
Mentors: Dr. Daniel Owens and Rina Carrillo

Coronaviruses, belonging to the Coronaviridae family, were first discovered in the 1960s having been characterized by the presence of distinct, protruding spike (S) proteins. This family of viruses is zoonotic, can be transferred from animals to humans, as demonstrated by the development of SARS-CoV-1 in 2002 from palm civets and MERS-CoV in 2012 from dromedary camels. The SAR-CoV-2 variety, the causative agent of the COVID-19 disease, emerged in Wuhan, China in December, 2019 from a currently unknown intermediate host. Like other coronaviruses, viral entry was facilitated by the binding of the S proteins to human ACE2 receptors. Although the symptoms, transmission, and morphology of SARS-CoV-2 virus were similar to SARS-CoV-1, the two types of coronaviruses had significant differences in their S proteins at the amino acid sequence level. The SARS-CoV-2 S protein amino acid sequence increases its affinity for ACE2 receptors, potentially explaining why it has been especially virulent. Additionally, two strains of SARS-CoV-2 have been discovered, indicating the possible need for two different types of CoV-2 vaccinations.

Introduction

The emergence and rapid spread of the novel coronavirus has shaken the entire world to its core. The novel species has proven to be especially resilient and infectious in comparison to prior coronavirus outbreaks, creating a global sense of anxiety. Coronaviruses, which belong to the *Coronaviridae* family, are single-stranded, positive-sense RNA viruses characterized by the presence of protruding spike (S) glycoproteins on their outer plasma membrane (Hulswit et al., 2016). This family of viruses are zoonotic, meaning they are transmissible from animals to humans or, most specifically, a virus that originates in animals but can also infect humans. Bats have been determined to be the natural host for the *Coronaviridae* family; however, an intermediate host is often needed for human infection (Hoffmann et al., 2020).

Although new information about the differences between previously identified strains and the current forms of coronaviruses are being uncovered daily, it has been concluded that

My name is Gabriele Jamain and I graduated from UH Mānoa in Spring 2020 with a Bachelor's Degree in Animal Science. I currently work as a veterinary assistant and am interested in pursuing veterinary medicine. Due to COVID-19, I had the unique experience of adapting to remote learning half-way through my final undergraduate semester. I wrote this article for my Biochemistry Lab course as an alternative to the standard lab reports I had been writing all term. Though the stress of the pandemic brought unexpected challenges, I took this assignment as an opportunity to relate my current coursework to this frightening global issue. I hope my finding will benefit those seeking information on SARS-CoV-2 during these uncertain times.

Horizons, Vol. 5, 2020, pp. 97–101
Copyright © 2020 by the University of Hawai'i at Mānoa

it is these distinctive spike (S) glycoproteins which enable viral entry into target cells (Walls et al., 2020). The S proteins allow the virus to dock onto and fuse with the host cell, followed by the release of its viral genetic material for replication (Hulswit et al., 2016). Furthermore, the COVID-19 disease, originating from the SARS-CoV-2 virus, has been shown to be especially virulent as its spike proteins have undergone mutation, creating "S" and "L" SARS-CoV-2 strains (Tang et al., 2020). Understanding the mechanism of how this S protein/receptor mechanism functions is critical in the development of treatments and vaccines for this deadly, novel strain. Therefore, in this review the mechanism of entry for SARS-CoV-1 and SARS-CoV-2's S proteins is analyzed and the establishment of two SARS-CoV-2 strain types is considered.

History of Coronavirus

Discovery of Coronavirus

Experimentation on human coronaviruses dates back to 1960s (Kahn & Mcintosh, 2005). Terrell and Bynoe were the first to begin analysis of coronaviruses, specifically a strain named B814, which was obtained from the respiratory tract of an adult experiencing a common cold (Kahn & Mcintosh, 2005). Concurrently, scientists Hamre and Procknow began cultivating another viral strain (named 229E) from cold-ridden volunteers in tissue cultures (Kahn & Mcintosh, 2005). Both B814 and 229E were found to be ether-sensitive, indicating the viruses possessed a lipid-containing coat to aid infection (Kahn & Mcintosh, 2005). Further recovery and culturing of multiple ether-sensitive strains were performed by another laboratory, who termed these viruses "OC" to highlight they were grown in organ cultures (Kahn & Mcintosh, 2005).

Kahn & Mcintosh (2005) stated that the electron microscopy of fluids from B814 infected organ cultures indicated the presence of particles similar to the infectious bronchitis virus in chickens. The observed particles were roughly 80–150 nm in size, pleomorphic, membrane-coated, and covered with club-shaped surface projections (Kahn & Mcintosh, 2005). Electron microscopy of 229E and the OC strains displayed similar morphology. Additionally, electron microscopy of various animal viruses including infectious bronchitis virus, mouse hepatitis virus, and transmissible gastroenteritis virus of swine, demonstrated the same morphological characteristics, hinting at the zoonotic capabilities (Kahn & Mcintosh, 2005). These viruses were named "Coronavirus", with *corona* symbolizing the crown-like surface projections (Kahn & Mcintosh, 2005).

These projections are now known to be the spike (S) glycoproteins, which have been found to interact with human ACE2 receptors (Walls et al., 2020). ACE2 receptors are commonly found in epithelial cells of the lung, intestine, kidney, heart, and blood vessels. Activation of these receptors induces vasoconstrictive, proinflammatory, and pro-oxidative effects (Bavishi et al., 2020). Consequently, binding of the S protein to ACE2 receptors causes undesirable increased blood pressure and inflammation in the lungs, and in turn leads to the common respiratory symptoms (Bavishi et al., 2020).

Currently, there are four common human coronaviruses and three novel strains which cause much more severe symptoms (Kooraki et al., 2020). Epidemiological and volunteer inoculation studies on human strains OC43 and 229E gave a basic knowledge on how this family of viruses' functions (Kahn & Mcintosh, 2005). The main symptom related to infection was upper respiratory infection with occasional cases of pneumonia in infants and young adults, which often induced asthma in children or chronic bronchitis in adults and the elderly (Kahn & Mcintosh, 2005). Additional research on animal coronaviruses was conducted, though not all studies focused on respiratory disorders, indicating the coronavirus family is capable of infection through a variety of disease mechanisms (Kahn & Mcintosh, 2005).

Aside from the current pandemic, there have been two other worrisome coronavirus outbreaks since its discovery in 1965: SARS-CoV-1 and MERS-CoV (Kooraki et al., 2020). The *Coronaviridae* family is classified into four sub-groupings (alpha-CoV, beta-CoV, gamma-CoV, and delta-CoV), however, only the alpha-CoV and beta-CoV variety are capable of infecting humans (Ou et al., 2020). Furthermore, only the beta-CoV strains are known to cause severe upper respiratory symptoms and pneumonia as seen in the SARS-CoV-1 and MERS-CoV outbreaks. The emergence of the aggressive SARS-CoV in 2003 and MER-CoV in 2010 sparked scientists' interest in understanding the zoonotic transmission of coronaviruses (Kahn & Mcintosh, 2005).

SARS-CoV-1

The first case emerged in Guangdong, China between 2002 and 2003 (Kooraki et al., 2020). This virus was named SARS-CoV (now SARS-CoV-1 with the emergence of the novel strain in 2019), as patients developed severe acute respiratory syndrome (SARS). Additional symptoms included fever, pneumonia, and lower respiratory symptoms like coughing and dyspnea (Guarner, 2020). These symptoms were much more debilitating than the generalized upper respiratory infection exhibited by OC43 and 229E strains (Kahn & Mcintosh, 2005). The virus spread to 29 countries globally, infecting 8,422 individuals, with a mortality rate of 11% (Guarner, 2020). SARS-CoV-1 is transmitted person to person via large droplets expelled during coughing or sneezing, as well as direct personal contact like touching contaminated surfaces (Guarner, 2020). It was found that the SARS virus is viable up for up to 24 hours on dry surfaces but is degraded by common disinfectants like bleach (Kooraki et al., 2020). Palm civets and raccoon dogs were determined to be the intermediate host for SARS-CoV-1 (Hulswit et al., 2016).

MERS-CoV

In 2012, the second CoV outbreak occurred in the Middle East. This virus was travel-specific, meaning all of those who were infected had come from the Arabian Peninsula (Guarner, 2020). Although the virus affected a smaller population of people, the death rate was much higher. Dromedary camels were found to be the intermediate host between bats and humans (Walls et al., 2020). MERS-CoV was much more severe and had a higher death rate, it was much less infectious than the original SARS-CoV-1 virus. The R0, the predicted average number of people who will contract the disease from an infected individual, of MERS was only 1, even though the fatality rate was 35% (Guarner, 2020). It has been thought that MERs is transmitted by human to animal, rather than SARS which was human to human (Guarner, 2020).

Coronavirus Structure and Mode of Action

Zoonotic Host

Phylogenic analysis has shown that bats are the natural host and reservoir of the coronavirus, implying the human coronavirus strain is evolved from the bat coronavirus genome (Tang et al., 2020). Both SARS-CoV-1 and MERS-CoV were transmitted from bats to palm civets or dromedary camels, respectively, before finally reaching humans (Walls et al., 2020). This suggests that SARS-CoV-2 has its own specific intermediate host before infecting humans. Pangolins are currently considered as the most likely suspect as they have their own novel coronavirus variant with 85.5%-92.4% similarity to the 2019-nCoV strain genome (Tang et al., 2020). However, further research is required in order to establish that pangolins are the definitive intermediate host of SARS-CoV-2 (Tang et al., 2020).

Viral Structure

A coronavirus particle is comprised of at least four different structural proteins: Envelope proteins (E), membrane proteins (M), nucleocapsid proteins (N), and spike proteins (S) (Hulswit et al., 2016). Additionally, beta-CoV's contain membrane-anchored hemagglutinin-esterase proteins (HE), a type of Class I fusion protein responsible for facilitating influenza virus infection (Figure 1) (Hulswit et al., 2016). R.J.G. Hulswit et al. (2016) stated that S proteins are also Class I fusion proteins, but they are the key determinant of susceptibility to CoV infection. This is a significant point that suggests why beta-CoV's are potentially more infectious and severe in humans. Although the S proteins are the main fusion source, the HE proteins may play a role in the resilience and intensity of this virus throughout time.

The S protein is a trimeric protein composed of two functional subunits: the S1 subunit, which binds to a host cell receptor, and the S2 subunit, responsible for the fusion of viral

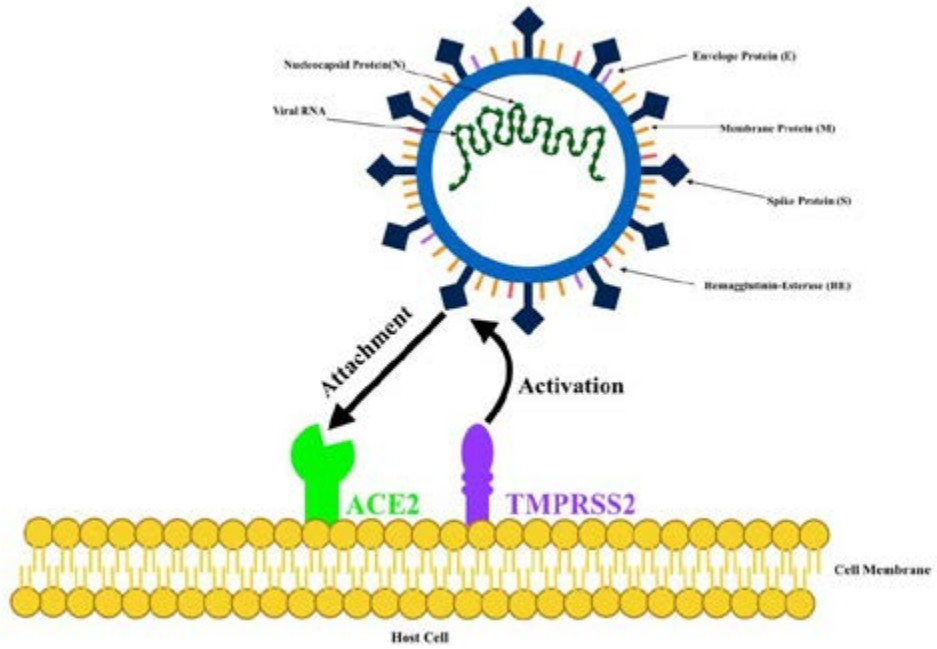

Figure 1 This figure has been adapted from an illustration by Hoffmann (2020). Although there are multiple types membrane proteins located on the viral membrane, attachment of the virus to human cells is facilitated by the S proteins. SARS-CoV-2's S proteins bind to human ACE2 receptors and uses the protease TMPRSS2 for activation.

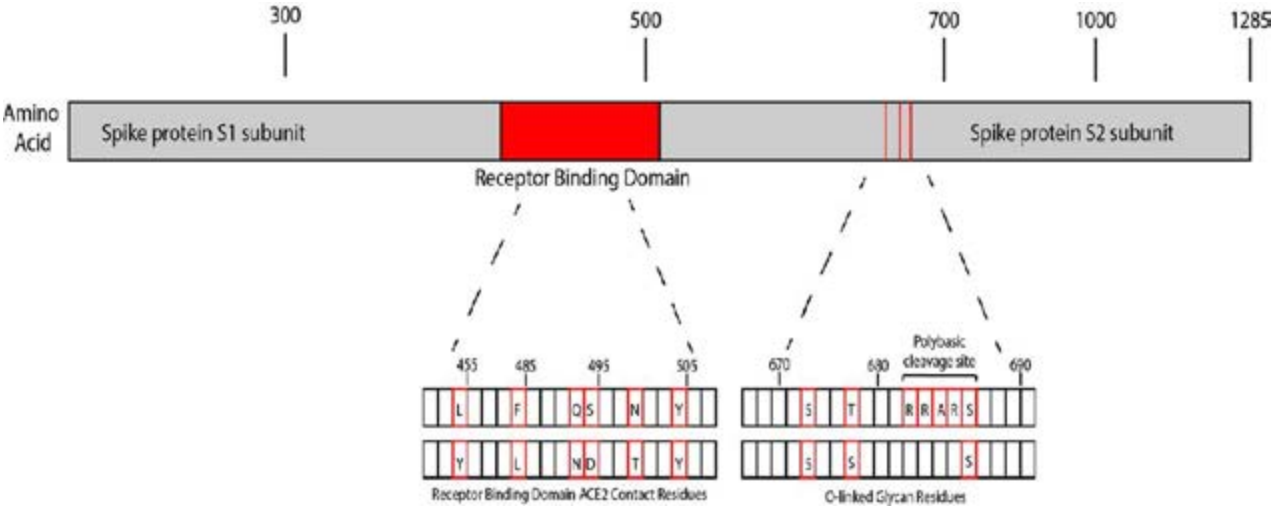

Figure 2 Features of the spike protein in human SARS-CoV-2 and SARS-CoV-1. This figure has been adapted from Andersen et al., 2020. The spike (S) protein genome was amplified to illustrate the different amino acid residues between SARS-CoV-1 and SARS-CoV-2. In the S1 subunit (RBD section) the six key ACE2 receptor binding residues for both coronavirus's are visualized. Only the final residue at Y505 is shared between the two species. The second section on the right visualized the polybasic cleavage site of the SARS-CoV-2 protein. Although both strains have O-linked glycerin residues, only CoV-2 strain has the cleavage site.

and cellular membranes (Figure 2) (Walls et al., 2020). Viral attachment is initiated by the binding of the S1 subunit to a host cell receptor. Entry of the viral RNA is achieved by "S protein priming" via a cellular protease whereby the S protein is cleaved between the S1 and S2 subunits (Hoffmann et al., 2020). The subunits remain non-covalently bonded to each other in a prefusion conformation after cleavage (Hoffmann et al., 2020). Walls et al. (2020) explained that the S1 subunit, containing the receptor-binding domains, stabilizes the prefusion state of the S2 subunit while anchored to the host membrane. The S protein undergoes further cleavage by host proteases at the S2' site of the fusion peptide, activating the protein and facilitating the incorporation of the viral membrane into the host's (Walls et al., 2020). This fusion event is irreversible due to highly specific conformational changes to the cell membrane (Walls et al., 2020).

S Protein Selectivity Between beta-CoV Species

The S1 subunit is comprised of multiple domains which serve to recognize and facilitate viral entry. As a result, attachment and entry vary for different viral species. One of these domains is the receptor binding domain (RBD), which is the specific region that interacts with the host cell receptors (Hoffmann et al., 2020). The RBD of MERS-CoV binds with dipeptidyl peptidase-4 (DPP4) as its entry receptor while the SARS-CoV class of coronaviruses bind to human angiotensin-converting enzyme 2 (ACE2) (Hoffmann et al., 2020).

Although both CoV-1 and CoV-2 can successfully bond to the ACE2 receptor, a portion of their RBD differ at the amino acid level. There are six crucial amino acids necessary for SARS-CoV-like viruses to bind to ACE2 receptors (Andersen et al., 2020). Andersen et al (2020) discovered that five of these six residues differ between SARS-CoV-1 and SARS-CoV-2 (Figure 2). These mutations in SARS-CoV-2 are hypothesized to lead a higher affinity for the ACE2 receptor, potentially explaining why SARS-CoV-2 has been especially infectious (Andersen et al., 2020). Slightly downstream from the RBD is a polybasic cleavage site composed of four specific amino acid residues. Although it is unclear what the exact function of the polybasic cleavage site it, it has been predicted that it facilitates viral entry (Andersen et al., 2020). This cleavage site falls right before the boarder of the S1 and S2 subunits, indicating that this could be the potential location of protease priming (Andersen et al., 2020). CoV S proteins may be cleaved by one or several host proteases, including furin, trypsin, cathepsins, and transmembrane protease serine protease-2 (TMPRSS-2) (Hulswit et al., 2016). Both SARS-CoV-1 and SARS-CoV-2 are susceptible to priming by TMPRSS-2; however, only SARS-CoV-2 contains a polybasic cleavage site (Andersen et al., 2020). It has been suggested that the presence of the polybasic cleavage site may influence protease priming by a currently unknown mechanism and increases the affinity for viral entry (Walls et al., 2020).

The Future for SARS-CoV-2

Mutation of SARS-CoV-2 S Proteins

In a study conducted by Tang et al (2020), 103 SARS-CoV-2 genomes were subjected to genetic analyses. The results indicated that the virus had evolved into two major strains: the "L"

and the "S" types. Majority of the mutations were singletons, indicating that the differences were caused by single nucleotide polymorphisms (SNPs). Two specific marker SNP sites were located at nucleotides 8,782, which caused no change in the amino acid residue, and 28,144, which changed an amino acid residue from serine to leucine (Tang et al., 2020). Of the 103 SARS-CoV-2 genomes analyzed by Tang et al. (2020), 72 strains possessed the leucine residue (now known as the "L" type) while 29 strains exhibited the serine residue ("S" type). It was concluded that L is the major type (70%) and S is the minor type (30%) (Tang et al., 2020).

Furthermore, the "S" type was considered an ancestral form, as the nucleotides of the S type at locations 8,782 and 28,144 mirrored those of other coronaviruses (Tang et al., 2020). This suggested that the L form of the SARS-CoV-2 virus is evolutionarily favored, and it may be more aggressive (Tang et al., 2020). To test this hypothesis, Tang et al analyzed viruses based on location and the date they were isolated. They discovered that among the 27 viruses isolated from Wuhan, 26 (96.3%) were L type, and only 1 (3.7%) was S type (Tang et al., 2020). Additionally, among the other 73 viruses isolated outside Wuhan, 45 (61.6%) were L type, and 28 (38.4%) were S type (Tang et al., 2020). This information is slightly contradictory under the assumption that the L type is more aggressive than the S type. It is hypothesized that the spread of the more severe L time was mitigated by China's strict viral prevention and control methods (Tang et al., 2020). In other countries, failure to follow preventative measure may cause a resurgence of the more aggressive L type and consequently lead to increased mortality (Tang et al., 2020).

Conclusion

Although SARS-CoV-2 has some similarities to SARS-CoV-1, it is clear there are multiple differences that make it a much more intimidating virus. Firstly, its spike proteins have a higher affinity for human ACE2 receptors located in lung/gastrointestinal endothelial cells, which facilitates viral attachment and fusion with the host cell's membrane (Walls et al., 2020). Additionally, the RBD of CoV-2 has a polybasic cleavage site which could potentially induce protease priming, further promoting viral fusion (Andersen et al., 2020). Since its discovery in December 2019, two SARS-CoV-2 types have been identified by Tang et al.: an ancestral "S' type and a more infective "L" type. Although the types differ by only one amino acid residue, this single change may lead to the need for development of multiple CoV-2 vaccines (Tang et al., 2020). The S protein of CoV serves as the antigen for human immune cells, which create a highly specific antibody against it (Tang et al., 2020). Considering the two types have different characteristics, two antibody types could be necessary. (Tang, et al., 2020) It is possible that one type of antibody may be enough to block the ACE2 receptor from viral attachment; however, further research is needed in order to confirm this hypothesis.

Works Cited

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nature Medicine*, *26*, 450–452. doi: 10.1038/s41591-020-0820-9

Bavishi, C., Maddox, T. M., & Messerli, F. H. (2020). Coronavirus Disease 2019 (COVID-19) Infection and Renin Angiotensin System Blockers. *JAMA Cardiology*. doi: 10.1001/jamacardio.2020.1282

Guarner, J. (2020). Three Emerging Coronaviruses in Two Decades. *American Journal of Clinical Pathology*, *153*(4), 420–421. doi: 10.1093/ajcp/aqaa029

Hoffmann, M. (2020, March 5). *Blocking SARS-CoV-2 cell entry*. Retrieved from https://www.eurekalert.org/multimedia/pub/226254.php?from=457812

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor Cell. doi: 10.1016/j.cell.2020.02.052

Hulswit, R., Haan, C. D., & Bosch, B.-J. (2016). Coronavirus Spike Protein and Tropism Changes. *Coronaviruses Advances in Virus Research*, 29–57. doi: 10.1016/bs.aivir.2016.08.004

Kahn, J. S., & Mcintosh, K. S. (2005). History and Recent Advances in Coronavirus Discovery. *The Pediatric Infectious Disease Journal*, 24 (Supplement). doi: 10.1097/01.inf.0000188166.17324.60

Kooraki, S., Hosseiny, M., Myers, L., & Gholamrezanezhad, A. (2020). Coronavirus (COVID-19) Outbreak: What the Department of Radiology Should Know. *Journal of the American College of Radiology*, 17(4), 447–451. doi: 10.1016/j.jacr.2020.02.008

Ou, X., Liu, Y., Lei, X., Li, P., Mi, D., Ren, L., Guo, L., Guo, R., Chen, T., Hu, J., Xiang, Z., Mu, Z., Chen, X., Chen, J., Hu, K., Jin, Q., Wang, J., & Qian, Z. (2020). Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature Communications*, 11(1). doi: 10.1038/s41467-020-15562-9

Tang, X., Wu, C., Li, X., Song, Y., Yao, X., Wu, X., Duan, Y., Zhang, H., Wang, Y., Qian, Z., Cui, J., & Lu, J. (2020). On the origin and continuing evolution of SARS-CoV-2. *National Science Review*. doi: 10.1093/nsr/nwaa036

Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., Mcguire, A. T., & Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell. doi: 10.1016/j.cell.2020.02.058