COVID-19: Structural Predictions of Viral Success

Richard Stein[1] and Lauren M. Young[2]

[1]New York University Tandon School of Engineering
[2]University of Chicago

May 27, 2020

Since the beginning of the 21st century, three coronaviruses have crossed the species barrier and caused serious human disease: severe acute respiratory syndrome coronavirus (SARS-CoV) in November 2002 [1, 2], Middle-East respiratory syndrome coronavirus (MERS-CoV) in 2012 [3, 4], and SARS-CoV-2 in 2019 [5, 6]. SARS-CoV-2 [7], initially called 2019-nCoV, is the etiological agent of COVID-19, a highly contagious infectious illness that was first reported in December 2019 in Wuhan, China and subsequently spread globally [8]. As of May 24, 2020, COVID-19 has caused >5,370,000 infections and >343,000 deaths worldwide [9].

Unfortunately, nearly 20 years after the SARS outbreak, and despite many attempts for vaccines and therapeutic agents directed against SARS and MERS, no approved prophylactics or therapeutics exist. As a result, the management of COVID-19 largely relies on supportive care [10, 11] and on hopes surrounding compounds that appeared promising against previous coronaviruses [12, 13]. This lost opportunity, in itself, offers a valuable lesson for current and future outbreaks, and the need for new experimental rationales to accelerate discovery.

The cellular entry of coronaviruses is fairly conserved across members of the *Coronaviridae* family and is mediated by the transmembrane spike (S) glycoprotein [14], a homotrimer [15, 16] that is often heavily glycosylated [17] and protrudes from the viral surface. Each of the three monomers of the spike glycoprotein consists of two functional subunits, S1, involved in membrane attachment, and S2, required for membrane fusion [15, 18]. In many coronaviruses, the spike glycoprotein is cleaved at the S1/S2 interface by host cell proteases [19]. Within the S1 domain, the receptor binding domain (RBD) attaches to the cellular receptor, which in the case of both SARS-CoV and SARS-CoV-2 is the angiotensin-converting enzyme 2 (ACE2) [19-21]. Another cleavage site, S2', is located within S2 [17, 19]. The spike glycoproteins of SARS-CoV and SARS-CoV-2 share 76% identity at the amino acid level [22, 23], although biophysical assays indicate that SARS-CoV-2 binds their common receptor, ACE2, with a 10-20 fold higher affinity than SARS–CoV [14].

As we contemplate the dynamics of COVID-19 and the development of prophylactic and therapeutic interventions, one of the key considerations is the emergence and potential relevance of viral mutations. In the short time since the pandemic started, several missense mutations have been observed in various SARS-CoV-2 isolates [24]. One of these, the 23403A>G variant, substitutes the aspartic acid at position 614 of the viral spike glycoprotein with glycine (D614G), and is frequently documented in European countries but rarely observed in China [25].

In the current issue of the *IJCP* , Becerra-Flores and Cardozo interrogate the impact of this mutation on pathogenicity and offer a structural correlate for their findings [26]. Their analysis includes confirmed COVID-19 cases and deaths as reported by the European CDC during the first week of April 2020 and examines the viral spike genomic sequences deposited in the GISAID database over that period, correlating the prevalence of the D614G mutation with fatality rates in the same regions. The authors then use cryo-electron microscopy data and*in silico* mutagenesis of this key residue to predict conformational preferences

of the two variants of the spike protein.

The analysis indicates that viruses isolated from European patients predominantly expressed a glycine at position 614 of the spike glycoprotein, while a high percentage of the isolates collected from Far East patients favored aspartic acid at the same position. The proportion of viral isolates having a glycine at this position significantly correlated with higher average and median case fatality rates across geographic areas. Interestingly, their data also imply a rationale for divergence in the behavior of the disease between the East and West coasts of the United States, based upon the provenance of the viral 'founders' in these regions, from the European and Asian variants, respectively.

Surprisingly, the authors' molecular modeling indicates that the presence of a glycine at position 614 *diminishes* binding to the cellular receptor when replacing the aspartic acid at that residue, mainly by reducing the spike protein's occupancy of the "up" or liganded state, when it is most amenable to receptor interaction. While seemingly counterintuitive, this finding opens at least two fascinating scenarios. As the authors hypothesize, a spike glycoprotein that harbors glycine at this position might be better protected from immune recognition, elicit the production of harmful antibodies, flood the host with ineffective antibodies, or some combination of all three. A delay in immune recognition may impact viral transmission by delaying symptomatic presentation or allowing unfettered infection without effective immune response. An aberrant response, suited to the viral conformation at large but not the infective conformation, could equally allow for an increased—but poorly targeted—inflammatory cascade. The possibility of a harmful immune response is particularly thought provoking, as antibody-dependent enhancement, the phenomenon by which antibodies facilitate viral entry into host cells that do not necessarily have viral receptors [27, 28], has been reported for many viruses, including coronaviruses [27, 29], dengue virus [30, 31], feline infectious peritonitis virus [32] , Ebola virus [33], and HIV [34]. Another possibility, not mutually exclusive, is that the D614G mutation creates or exposes a novel cleavage site in the spike glycoprotein.

Delving into these molecular mechanisms with confirmatory *in vitro* studies will hopefully reap the benefits of decades of scientific strides while simultaneously highlighting deficiencies in key areas that can guide our approach to the current pandemic. One of the immediate questions involves the impact of this and other mutations on vaccine efficiency and the potential need to develop multiple candidate vaccines that cover a range of epitopes and their variants. In all likelihood, there is a lengthy and tortuous road ahead, but characterizing significant variants will allow us to better understand many elusive aspects of this virus' success – the latent/incubation period, immune evasion and hyper-response, variable receptor binding, replication dynamics, and organ-specific pathogenesis—and discover host vulnerabilities that mutations such as D614G seem to exploit.

The D614G mutation appears to become more common as the pandemic unfolds [35]. That this phenomenon is simply the result of a founder effect is possible but unlikely, and rather may be explained by this variant's selective advantage allowing more efficient spread. Whether this advantage is conferred by infectivity, immune evasion, or pathogenicity—or some combination of these—is yet to be understood. Interestingly, this mutation is now known to travel simultaneously with other mutations, including one that affects the RNA-dependent RNA polymerase, with implications for proofreading, replication efficiency (and thus viral titer), and the emergence of drug-resistant viral phenotypes [36].

Addressing these molecular questions relies heavily on widespread efforts to assemble accurate and comprehensive data on population infection rates and mortality, and frequent sampling of the genotypes of circulating isolates on a global basis. So far, this feat has been challenging and continued deficiencies will translate into missed singular opportunities to link molecular findings with population-level consequences, ultimately leaving us less prepared to address both this and future pandemics.

The valuable and timely experimental strategy used by Becerra-Flores and Cardozo serves as an important analytic model that should be employed routinely to understand the 'molecular strategy' of this virus in the context of the evolving pandemic. This approach will also prove to be an indispensable instrument if also employed routinely at the onset of future outbreaks, which are all but guaranteed in the coming years, given

the only recently appreciated ease of global spread of viruses in the modern world. In summary, this set of tools allows us to perform active surveillance, monitor the emergence of deleterious mutations prior to their widespread distribution, and use informed *in silico* and structural data to make informed decisions guiding molecular research and epidemic preparedness.

References

1. Zhong NS, Zheng BJ, Li YM et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet* 2003; **362** : 1353-8.

2. Xu RH, He JF, Evans MR et al. Epidemiologic clues to SARS origin in China. *Emerg Infect Dis* 2004; **10** : 1030-7.

3. Al-Omari A, Rabaan AA, Salih S et al. MERS coronavirus outbreak: Implications for emerging viral infections. *Diagn Microbiol Infect Dis* 2019; **93** : 265-85.

4. Mackay IM, Arden KE. MERS coronavirus: diagnostics, epidemiology and transmission. *Virol J* 2015; **12** : 222.

5. Chatterjee P, Nagi N, Agarwal A et al. The 2019 novel coronavirus disease (COVID-19) pandemic: A review of the current evidence.*Indian J Med Res* 2020.

6. Chen X, Yu B. First two months of the 2019 Coronavirus Disease (COVID-19) epidemic in China: real-time surveillance and evaluation with a second derivative model. *Glob Health Res Policy* 2020;**5** : 7.

7. Zheng J. SARS-CoV-2: an Emerging Coronavirus that Causes a Global Threat. *Int J Biol Sci* 2020; **16** : 1678-85.

8. Li Q, Guan X, Wu P et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med*2020; **382** : 1199-207.

9. Center JHCR. Available at: https://coronavirus.jhu.edu/map.html. Last accessed May 17, 2020 2020.

10. Dyall J, Gross R, Kindrachuk J et al. Middle East Respiratory Syndrome and Severe Acute Respiratory Syndrome: Current Therapeutic Options and Potential Targets for Novel Therapies. *Drugs* 2017;**77** : 1935-66.

11. Sarma P, Prajapat M, Avti P et al. Therapeutic options for the treatment of 2019-novel coronavirus: An evidence-based approach.*Indian J Pharmacol* 2020; **52** : 1-5.

12. Devaux CA, Rolain JM, Colson P, Raoult D. New insights on the antiviral effects of chloroquine against coronavirus: what to expect for COVID-19? *Int J Antimicrob Agents* 2020: 105938.

13. Dyall J, Coleman CM, Hart BJ et al. Repurposing of clinically developed drugs for treatment of Middle East respiratory syndrome coronavirus infection. *Antimicrob Agents Chemother* 2014;**58** : 4885-93.

14. Wrapp D, Wang N, Corbett KS et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 2020;**367** : 1260-3.

15. Ou X, Liu Y, Lei X et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV.*Nat Commun* 2020; **11** : 1620.

16. Du L, Yang Y, Zhou Y et al. MERS-CoV spike protein: a key target for antivirals. *Expert Opin Ther Targets* 2017; **21** : 131-43.

17. Millet JK, Whittaker GR. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Res* 2015;**202** : 120-34.

18. Walls AC, Park YJ, Tortorici MA et al. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 2020;**181** : 281-92.e6.

19. Coutard B, Valle C, de Lamballerie X et al. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res* 2020;**176** : 104742.

20. Verdecchia P, Cavallini C, Spanevello A, Angeli F. The pivotal link between ACE2 deficiency and SARS-CoV-2 infection. *Eur J Intern Med* 2020.

21. Zhang H, Penninger JM, Li Y et al. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. *Intensive Care Med* 2020; **46** : 586-90.

22. Hoffmann M, Kleine-Weber H, Schroeder S et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020; **181** : 271-80.e8.

23. Chan JF, Kok KH, Zhu Z et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect* 2020;**9** : 221-36.

24. Stefanelli P, Faggioni G, Lo Presti A et al. Whole genome and phylogenetic analysis of two SARS-CoV-2 strains isolated in Italy in January and February 2020: additional clues on multiple introductions and further circulation in Europe. *Euro Surveill* 2020;**25** .

25. Koyama T, Weeraratne D, Snowdon JL, Parida L. Emergence of Drift Variants That May Affect COVID-19 Vaccine Development and Antibody Treatment. *Pathogens* 2020; **9** .

26. Becerra-Flores M, Cardozo T. SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. *Int J Clin Pract* 2020.

27. Wan Y, Shang J, Sun S et al. Molecular Mechanism for Antibody-Dependent Enhancement of Coronavirus Entry. *J Virol*2020; **94** .

28. Tirado SM, Yoon KJ. Antibody-dependent enhancement of virus infection and disease. *Viral Immunol* 2003; **16** : 69-86.

29. Wang SF, Tseng SP, Yen CH et al. Antibody-dependent SARS coronavirus infection is mediated by antibodies against spike proteins.*Biochem Biophys Res Commun* 2014; **451** : 208-14.

30. Katzelnick LC, Gresh L, Halloran ME et al. Antibody-dependent enhancement of severe dengue disease in humans. *Science* 2017;**358** : 929-32.

31. Wilder-Smith A, Ooi EE, Horstick O, Wills B. Dengue. *Lancet*2019; **393** : 350-63.

32. Takano T, Yamada S, Doki T, Hohdatsu T. Pathogenesis of oral type I feline infectious peritonitis virus (FIPV) infection: Antibody-dependent enhancement infection of cats with type I FIPV via the oral route.*J Vet Med Sci* 2019; **81** : 911-5.

33. Kuzmina NA, Younan P, Gilchuk P et al. Antibody-Dependent Enhancement of Ebola Virus Infection by Human Antibodies Isolated from Survivors. *Cell Rep* 2018; **24** : 1802-15.e5.

34. Beck Z, Prohászka Z, Füst G. Traitors of the immune system-enhancing antibodies in HIV infection: their possible implication in HIV vaccine development. *Vaccine* 2008; **26** : 3078-85.

35. Hadfield J, Megill C, Bell SM et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 2018; **34** : 4121-3.

36. Pachetti M, Marini B, Benedetti F et al. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant.*J Transl Med* 2020; **18** : 179.