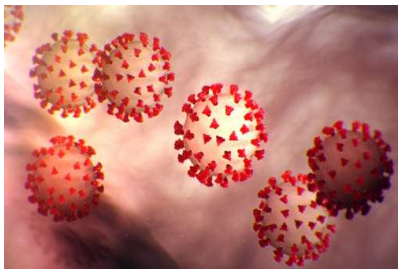


## COVID-19 Vaccine Will Close in on the Spikes

Posted February 20, 2020 by Ricki Lewis, PhD

As epidemiologists try to stay ahead of the spread of new coronavirus COVID-19, vaccine developers, like Sanofi and Johnson & Johnson, are focusing on the "spike" proteins that festoon viral surfaces. Following clues in genomes is critical to disrupting the tango of infectivity as viruses meet and merge with our cells.

Vaccine developers look specifically to the molecular landscapes where viruses impinge upon our respiratory and immune system cells. Targeting COVID-19 is especially challenging, because efforts to develop a vaccine against its relative, the SARS coronavirus (SARS-CoV), elicit only partial responses. But those steps are now serving as jumping off points for pharma.

The relationship between viruses and humans can seem like a science fiction plot. The viruses that make us sick may be little more than snippets of genetic material borrowed, long ago, from human genomes. Packaged with their own proteins, viruses return to our bodies, taking over to make more of themselves.

A zoo of animal hosts

Coronaviruses present a "severe global health threat," write researchers from Wuhan University and Sun Yat-sen University in the *Journal of Medical Virology*. The viruses aren't new, nor do they infect only people. They cause:

- diarrhea in pigs, dogs, and cows
- fever and vasculitis in cats
- fever and anorexia in horses
- severe lung injury in mice
- lung disease and death from liver failure in whales
- respiratory tract infection in birds (bulbuls, sparrows, and chickens)

Bats' immune response enables them to house coronaviruses without becoming sick, making them a dangerous reservoir of infection.

Some species spread coronaviruses without becoming sick, like the camels that carry MERS, and bats, which carry many viruses.

## Human Coronaviruses Before COVID-19

Before December 2019, six coronaviruses were known to infect humans. The first two, HCoV-229E and HCoV-OC43, were discovered in the 1960s. They cause about 30% of colds, with rare case reports of pneumonia in patients who had other viral infections or were immunocompromised.

In 2002 came SARS-CoV and in 2004 HCoV-NL63, which causes pneumonia and bronchitis, rarely. SARS (severe acute respiratory syndrome) caused more than 8,000 infections and 774 deaths, but most of us have antibodies to NL63, indicating past exposure that didn't make us very sick.

In 2005 came HKU1. It causes pneumonia in young children and 1.5% of cases of adult respiratory distress syndrome.

MERS-CoV (Middle Eastern Respiratory Syndrome) emerged in 2012 in the Arabian peninsula, and is rare but can be fatal. SARS and MERS show zoonotic (to other animals) as well as human-to-human transmission.

Two coronaviruses without a predilection for human bodies may also be important, epidemiologically speaking. HKU2, which killed 24,000 piglets in southern China from diarrhea in 2017, is the first "spillover" from a bat coronavirus to livestock. And Beluga whale CoV/SW1, although only distantly related to the human pathogens, could reveal how bat viruses get into sea creatures.

The coronaviruses are of four genera (equivalent to the "*Homo*" in "*Homo sapiens*"): alpha, beta, gamma, and delta. COVID-19, SARS, MERS, HCoV-OC43, and HKU1 are beta. Two other human coronaviruses, HCoV-229E and HCoV-NL63, are alpha.

Many viruses that plague humans have RNA as
their genetic material. It's copied into DNA in our cells.

## Anatomy of COVID-19

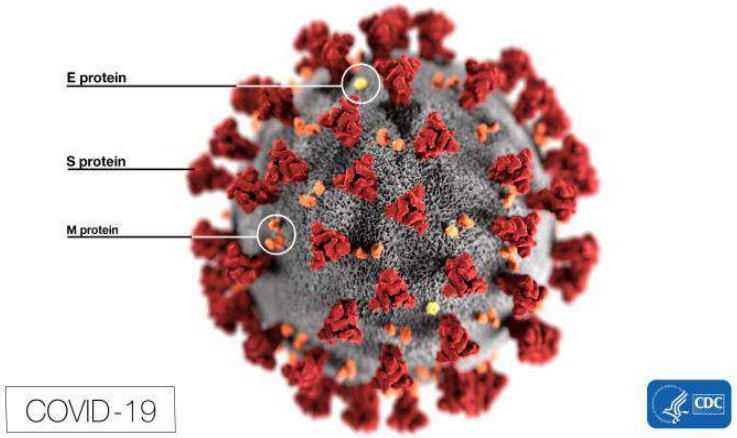
A virus isn't a cell, isn't even considered alive. It's a nucleic acid (DNA or RNA) wrapped in a coat of proteins, some attached to sugars (glycoproteins).

Many familiar viral pathogens – those that cause cold, flu, hemorrhagic fevers like Ebola, rabies, dengue, and yellow fever – are RNA viruses, notorious for mutating rapidly and unable to correct errors.

The "body" of COVID-19 is basically a genome enveloped in glycoproteins, with a smear of fat and bearing the crown of spikes that inspired the name "coronavirus."

The genome is a single strand of RNA that is termed "positive-sense." That means that the infected cell treats the viral genome as if were it's own messenger RNA (mRNA), translating it into proteins. A "negative-sense" RNA virus requires more manipulation; a host enzyme must make a positive-sense copy.

A coronavirus genome typically is 26,000-32,000 bases long. That's hefty for a virus, but tiny compared to a human gene. Our *BRCA1* gene, for example, is 125,951 bases long. Coronavirus RNAs are embellished with "caps" and "tails" like those of human mRNAs.




Once ensconced in a human cell, a half dozen or more viral mRNAs are peeled off. The first, representing about two-thirds of the viral genome, encodes 16 protein "tools" that viruses require to replicate. Making this toolkit is a little like downloading an installer for new software.

The tools ("non-structural proteins") are enzymes needed to produce the other encoded proteins, and transcription factors to continually renew the RNA instructions.

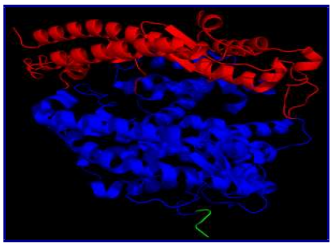
The other third of the viral genome encodes four "structural" proteins that are the nuts and blots that build the virus:

- Spike, or S protein, is made early in infection. One part of it, S1, grabs a receptor molecule sticking out of a host cell and another part fuses to the cell membrane. Three copies of the S protein form each spike.
- Membrane (M) glycoprotein lies beneath the spikes, where it shapes mature viral particles and binds the inner layers.
- Lipid (fat) is borrowed from host cell membranes during past infections.
- Envelope (E) glycoproteins control the assembly, release, and infectivity of mature viruses.
- Nucleocapsid (N) proteins knit a characteristic shell of identical subunits, like the panes of a greenhouse, that binds and packages the RNA genome. It also serves as a cloaking device, hiding viruses from our immune system's interferons and RNA interference.

All coronaviruses share the "tools," but differ in a few additional structural proteins tailored to the host species.

COVID-19's Spikes Bind at ACE2 Receptors

Viruses have co-evolved with us, using proteins that jut from our cell surfaces. HIV and West Nile virus enter through CCR5 receptors, which dot white blood cells. Influenza viruses bind sialic acid residues. Coxsackievirus and adenovirus target part of an antibody. And herpes simplex uses 3 different doorways.



ACE2 raises blood pressure in us, but is a receptor for COVID-19.

COVID-19 latches onto angiotensin-converting enzyme 2, aka ACE2.

To us, ACE2 is an enzyme that has an effect on blood pressure.

To COVID-19, ACE2 is a receptor, an entranceway, in the airways and alveoli (air sacs) as well as in blood vessel linings. ACE2 is also a receptor for SARS-CoV and NL63-CoV. (MERS-CoV uses a different receptor.)

The key to developing vaccines and treatments is the three-dimensional shapes of the parts of the virus that contact our cells.

SARS and NL63-CoV bind to a helical part of ACE2 that snakes up from cell membranes, forming distinctive tunnels and bridges that comprise a "hot spot" for viruses. The attraction of a virus to a cell receptor hot spot is a little like a tired commuter emerging from a subway station and seeing a Starbucks sign, moving towards the coffee shop as if dragged by a tractor beam. The viral hot spot that beckons both SARS and COVID-19 is a shared drug and vaccine target – and so all the work on developing a SARS vaccine is now in the spotlight.

Researchers knew from SARS that the S1 parts of the viral spikes hug the ACE2 receptor at a region of five amino acids (protein building blocks). Even though four of the amino acids differ in COVID-19, they are similar in size and charge to their counterparts in SARS.

If S1 attaches SARS to the ACE2 receptor like a boat docking, would COVID-9 tie up at exactly the same points?

Teaming a traditional crystal structure approach with computational methods, Pei Hao, of the Chinese Academy of Sciences and colleagues modeled the interface, showing that COVID-19 indeed binds ACE2 just like SARS does, with slightly less force. In a Letter to the Editor of *Science China Life Sciences*, they conclude that the new virus "poses a significant public health risk for human transmission via the S-protein-ACE2 binding pathway."

Even more recently, researchers from the University of Texas at Austin and the National Institute of Allergy and Infectious Disease used electron microscopy to zero in on a "pre-fusion" spike protein triplet nearing the receptor. They imaged one of the three spikies rotating upward to latch onto the receptor, revealing precisely where a vaccine must fit. The work is published in *Science*.

A Fish, A Bat, and A Human Walk Into a Seafood Market …

Comparing genome sequences is a classic way to sort out evolutionary relationships, and the comings and goings of viruses are evolution in action. Evolutionary trees, whether going back millions of years to dinosaurs or just years to viruses, depict descent from shared ancestors. The meme of chimps leading directly to humans is and has always been incorrect.

The pinecone soldierfish *Myripristis murdjan* may pass COVID-19 from bats to us.

COVID-19 and the SARS-CoV have a common ancestor, a bat coronavirus. But COVID-19 is actually closer to the bat virus, sharing 96% of its genome sequence, compared to about 86% with SARS-CoV. And muddying the waters further, COVID-19's spike gene shares a 39-base insertion with a type of soldierfish that swims in the South China Sea.

Somehow, the virus that evolved into COVID-19 may have started in a bat in 2013 and gotten into fish that ended up in the Wuhan Huanan Seafood Wholesale Market at the epicenter of the pandemic.

The confusion arises from the promiscuity of RNA viruses. Over the ages, they swap genetic material, mutate, and lose and gain pieces of themselves. The result is a constant spawning of patchwork genomes that under some circumstances harm the host species. And our bodies help it all along, sneezing, oozing, bleeding, or crapping out zillions of viruses before either expiring or recovering.

Did COVID-19 Come From Us?

How does a virus find itself at a door to a host cell? The "escaped gene" hypothesis leads the classic list of three explanations for a virus, such COVID-19's, emergence:

- Viruses were ancient intermediates between collections of self-replicating chemicals and the first cells. (The virus-first or primordial viral world hypothesis).
- Viruses were once cells invaded by parasites that robbed them of the ability to manufacture their own proteins. As viruses, they infect cells to make the proteins they need to reproduce. (The cellular regression hypothesis)
- Viruses evolved, many times in many organisms, as mobile genetic elements – aka "jumping genes" – that produced protein coverings and took bits of fatty membrane from cells. (The escaped gene hypothesis).

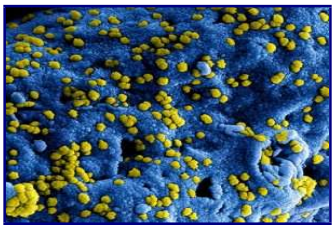
All three may have happened.

So what's COVID-19's story? Is a hint in what normally binds the receptor?

Perhaps sometime in the past, a virus formed, or came to include, human DNA or RNA instructions for making an integrin, which is a protein that binds to ACE2. Integrins glue our cells to surrounding connective tissue. The viral spike masquerades as the integrin, grabbing our cells.

In other words, a viral epidemic may arise as an accident, of sorts, of biochemistry and evolution.

Vaccine!



SARS coronaviruses (yellow)

Spike protein took center stage as a vaccine candidate early in the SARS outbreak, because it elicits an antibody response in mice. Various vaccine strategies – live weakened SARS, hitching spike genes to existing vaccines, circles of DNA housing spike genes and triplets of spike proteins in nanoparticles – haven't worked well enough. But genomic technology has exploded since SARS, leading to insights with astonishing rapidity.

In an exhaustive study preprinted in bioRxiv, Arunachalam Ramaiah of the University of California, Irvine and Vaithilingaraja Arumugaswami of UCLA catalogued where key parts of the four structural proteins of COVID-19 bind to the proteins that mark immune system cell surfaces. The work implicated the spike protein, but from the perspective of the immune response rather than the receptor.

Now pharma and biotech companies are charging ahead to create vaccines, partnering with BARDA, the Biomedical Advanced Research and Development Authority. Under Health and Human Services, BARDA was established in 2011 "to aid in securing our nation from chemical, biological, radiological,

and nuclear threats, as well as from pandemic influenza and emerging infectious diseases," by speeding the trajectory of diagnostic, vaccine, and treatment development.

But vaccine and drug development take time. Until then, treatment remains supportive. Meanwhile, epidemiologists are filling in the denominators of the case:fatality stats to determine exactly how deadly the infection is, and assess if asymptomatic individuals are unwittingly spreading it.

I'll end with a sobering thought from a pathogen's point of view. Even if we vanquish COVID-19, a continuing if not escalating perfect storm of events foreshadows other viral epidemics: climate change, human travel, and our encroachment into the turfs of other animals.


Ricki Lewis, PhD

Ricki Lewis is a science writer with a PhD in genetics. The author of several textbooks and thousands of articles in scientific, medical, and consumer publications, Ricki's first narrative nonfiction book, "The Forever Fix: Gene Therapy and the Boy Who Saved It," was published by St. Martin's Press in March 2012. In addition to writing, Ricki provides genetic counseling for parents-to-be at CareNet Medical Group in Schenectady, NY and teaches "Genethics" an online course for master's degree students at the Alden March Bioethics Institute of Albany Medical Center.