

How science demolishes the right-wing fiction of a Wuhan ‘lab leak’ as the source of coronavirus

Part two of a three-part series

Benjamin Mateus

21 June 2021

This is part two of a three-part series. Part one can be viewed here, and part three can be viewed here.

Glossary

In the interests of clarity, since much of the following discussion includes scientific terminology, further obscured by the shorthand of Twitter threads and other online exchanges, a glossary of key terms and abbreviations may be helpful to the reader.

- ACE2: Angiotensin-converting enzyme II, an enzyme on the surface of many human cells, targeted by the spike protein of SARS-CoV-2
- CTCCTCGGCGGG: Letters signifying a particular string of the four nucleotides that make up all basic genetic material, such as amino acids
- FCS: Furin cleavage site, a point on the spike protein of SARS-CoV-2 where the protein is easily cut by a protease called furin, helping it invade cells in the host human
- MERS-CoV: The virus that caused the Middle East Respiratory Syndrome pandemic in 2012
- Protease: An enzyme that helps proteins break down into smaller components; furin is one
- nCoV2019: An early acronym referring to the virus now called SARS-CoV-2
- RaTG-13: A bat virus found in caves in Yunnan province in China, genetically similar to SARS-CoV-2, but not a precursor, despite the claims of the conspiracy theorists
- RRAR: A sequence of amino acids found in some viruses
- S1/S2 boundary: The point on the spike protein of SARS-CoV-2 where the furin cleavage site is located in order to split the protein most efficiently
- SARS-CoV-1: The virus that caused the SARS pandemic in 2002, which was largely contained
- SARS-CoV-2: The virus that causes COVID-19
- WIV: Wuhan Institute of Virology, the lab where Dr. Shi Zhengli conducted research on bat coronaviruses

Dr. Kristian Andersen

Regarding Dr. Kristian Andersen and his explanations on the work he conducted on SARS-CoV-2 and the scientific conclusions he reached, the information below is reconstructed from Twitter threads he posted in discussions with other scientists and the public in general. The material he presented, which was never reported by the mainstream press, provides extensive replies to essential scientific questions and concerns being raised.

On June 6, Dr. Andersen’s Twitter account was inexplicably turned off. According to a Newsweek report on June 7, “Dr. Andersen removed his page on the social media platform after the release of email exchanges between himself and Dr. Fauci.”

This refers to the blizzard of right-wing distortions in which the initial discussions among scientists about the origins of the virus, which included consideration, as one hypothesis, of a laboratory leak, are portrayed as proof that the lab leak hypothesis was deliberately suppressed for political reasons, rather than discarded because it simply did not fit the facts.

Andersen has been steadfast on the natural origin of the SARS-CoV-2 virus since his work on the issue early in the pandemic, which led to the critical “Proximal Origins” report in *Nature Medicine* in March 2020.

In the context of his previous frank and open responses, one can surmise that this act of self-censorship was done under the influence of extreme pressures he must have faced from the media and political establishment. There is a striking contrast between the media promotion of Wade, an overt supporter of pseudo-scientific racism, and the gagging of Andersen, a foremost authority in the science of virology.

The threads cited below were copied and saved by this writer in anticipation of such an occurrence.

In one of these exchanges on Twitter, replying to Roger Pielke Jr., professor of environmental studies at the University of Colorado at Boulder, who was commenting on Andersen’s initial speculations about a lab origin, Dr. Andersen explained that when he had initially found the genome of SARS-CoV-2 inconsistent from the standpoint of evolutionary theory, this meant “we thought, on a preliminary look, that the virus could have been engineered and/or manipulated. It turns out the data suggest otherwise—which is the conclusion of our paper.”

One saved document deserves a fuller quotation. Dr. Andersen wrote, “Following up on my emails to Dr. Fauci from early 2020 about SARS-CoV-2 (nCoV), a couple of important questions came up: 1) What looked engineered to you? 2. What made you change your mind?”

His answer to question 1:

For our preliminary studies, there was very limited data with

only about ten genomes from Wuhan and the genome of RaTG13 not yet available, nor were the CoV genes from the pangolins [*It has been speculated that these animals may be the intermediate host of the SARS-CoV-2*]. We were very well aware of the CoV research ongoing at the WIV. There were features of the Sars-CoV-2 that, to us, appeared unique and at the time did not have an immediate obvious evolutionary precursor.

The features that stood out to us then were, 1) The furin cleavage site (unique in the sarbecovirus subgenus to which SARS-CoV-2 belongs, although it is present in several betacoronaviruses, the genus of SARS-CoV-2), 2) The SAR-CoV-2 Receptor Binding Domain (unique at the time and our modeling suggested it could be a strong if not perfect human ACE2), 3) A unique restriction enzyme BAMHI followed by a higher level of conservation towards the end of the spike protein, and 4) A few other residues found to be important from research from SARS-CoV-1.

Some explanatory points about the above discussion.

First, though the furin cleavage site is unique for a virus like SARS-CoV-2, such a site is necessary for many viral diseases, including HIV, Ebola, and even influenza. It was discovered that MERS-CoV possessed such a site and may explain the highly lethal nature of an infection with this virus. A recent phylogenetic analysis, done by Suwen Zhao and Yiran Wu in the journal *Stem Cell Research*, found that furin cleavage sites at the spike portion of the genome have occurred independently several times in their evolution, supporting the natural origin conception. It has been surmised that these sites can make the virus more transmissible. At the time of their first glimpse, this unique finding raised their worries.

Second, Andersen noted that though the virus did not efficiently bind to the human ACE2 receptor, the binding was sufficiently strong that it pricked their interests as to possible bioengineering. Subsequently, Dr. Edward Holmes of the University of Sydney has found viruses similar to SARS-CoV-2 that can weakly bind human ACE2 receptors without the need for a furin cleavage site at all.

Nicholas Wade and several other pro-lab-leak proponents have raised the furin cleavage site to suggest that these are the hallmarks of calculated manipulations. Furin is a particular enzyme/protein in humans that cuts special sections in other proteins to activate them. The SARS-CoV-2 virus contains such a site on its spike protein. When this cut occurs, it allows the virus to change into its active form and bind the ACE2 receptor and enter the host's cells. The suggestion by the conspiracy theorists is that the furin cleavage site was "seamlessly" inserted into a precursor, the bat virus RaTG13, to create SARS-CoV-2.

Let us hear what Dr. Andersen stated in another Twitter thread on this subject:

The SARS-CoV-2 furin cleavage site is yet again in the news - this time because of a quote by Nobel laureate David Baltimore. The site is not a 'smoking gun,' nor does it 'make a powerful challenge to the idea of a natural origin. The furin cleavage site (FCS)/polybasic cleavage site is present in SARS-CoV-2 at the S1/S2 junction of the spike protein, where it mediates the cutting (by the host protease furin, among others) of the spike, which is required for infections of cells.

The FCS was created by an out-of-frame insertion of 'CTCCTCGGCGGG' creating the '(P)RRAR' amino acid sequence, which constitutes a suboptimal polybasic cleavage site that is important for expanding SARS-CoV-2 host range, its transmission and pathogenesis, etc. FCSs are abundant, including being highly prevalent in coronaviruses. While SARS-CoV-2 is

the first example of a SARSr virus with an FCS, other beta coronaviruses (the genus for SARS-CoV-2) have FCSs, including MERS and HKU1.

There is nothing mysterious about having a 'first example' of a virus with an FCS. Viruses sampled to date only give us a teeny-tiny fraction of all the viruses circulating in the wild. Fragments - such as the CTCCTCGGCGGG - come and go all the time.

Nucleotides, codons, and amino acids

Briefly, by way of explaining the above Tweet, all living cells or viruses using the machinery of living cells must translate the genetic material of their DNA or RNA into proteins. The four basic nucleotides represent the building blocks of the genetic material, or letters, for spelling out the protein that will be constructed. C stands for Cytosine, A for Adenine, G for Guanine, and T for Thymine. A fifth nucleotide is Uracil, used in RNA.

A triplet of nucleotides makes a "codon" that designates an amino acid. The triplet nucleotides can be sequenced in various forms. For example, the amino acid Alanine, designated by the symbol A, can be made by a combination of GCT, or GCC, or GCA, or GCG. In all, 20 amino acids are used as the building blocks of all proteins. These are essentially the components used by cells to create the proteins and enzymes they need to conduct their biological functions.

The sequence of nucleotides mentioned by Dr. Andersen leads to amino acid sequences designated by the letters RRAR (here R is Arginine and A is Alanine). Early in the COVID-19 global outbreak, questions were focused on this sequence, raising suspicions of a lab-generated virus.

Dr. William Gallaher from LSU Health New Orleans School of Medicine, professor emeritus in the microbiology department, provided the following explanation back in February 2020, which aligns with the answers provided by Dr. Andersen.

I see no evidence at all to support such a claim. In sharp contrast, I have studied the question in detail, using the RaTG13 and Wuhan sequence at the S1/S2 boundary. I find convincing proof of exactly opposite conclusion—that RaTG13 could NOT be a proximal source of the Wuhan virus.

He goes on to explain:

One has to consider that the PRRA is an unusual sequence to introduce to generate a furin site—others even among coronaviruses like MHV A59 are so much better. Also, that the underlying code CCTCGGCGGGCA introduces an unnecessarily G and C rich region where none otherwise exists. Not likely scenarios for something a gene jockey would do. Then one looks at the actual RNA alignment. The 'insert' is actually not in frame, but CTCCTCGGCGGG, or -2 out of frame [*see Dr. Andersen's comments above.*] Again, who does that?

This implies that the furin cleavage site present in the SARS-CoV-2, despite Wade's claim, is sufficiently clumsy and inefficient in its construct that a "gene jockey" wouldn't employ such a construct to make a furin cleavage site from scratch.

He also noted at the time that the RaTG13 and the SARS-CoV-2 were sufficiently divergent that RaTG13 couldn't be the "proximal source of the nCoV2019."

Dr. Gallaher concludes:

Given that furin cleavage signals are present in other coronaviruses at exactly that point in the S1/S2 boundary region, it only LOOKS unusual, especially against the backdrop of SARS. The preponderance of evidence, coupled with Occam's razor (that the simplest explanation is preferred) dictates that the PRRA sequence has been conserved in nCoV2019 from a long-ago ancestor virus. It is not of suspicious origin. The closest bat virus sequence is really not close at all.

The scientific method vs. conspiracy theories

Moving on Dr. Andersen's response to question 2:

All the features in SARS2 that to us suggested possible engineering were identified in related CoVs in the first half of 2020, which largely invalidated our previous hypothesis of engineering and instead bolstered the argument for a natural origin. In the days immediately following my email to Dr. Fauci, additional data was released (Or we became aware of it) including the full genome of RaTG13.

Following up on our preliminary analysis, we did much more extensive investigations both on RaTG13 and other CoV genomes to compare genomic diversity more broadly across CoVs. We looked at all the literature from the WIV, investigated common virus backbones and molecular and cellular cloning techniques used at WIV and UNC, investigated sequence data sets produced from WIV and ECOHEALTH, and performed KMER-based (phylogenetic studies) and recombination analyses on SARS-CoV-2. We also had a lot of considerations about likelihoods related to virus emergence, virus discovery-capture, virus manipulation, lab escape, etc. Many of the analyses were completed in a matter of days and allowed us to relatively quickly reject our preliminary hypothesis that SARS-CoV-2 might have been engineered.

This is a textbook example of the scientific method where a preliminary hypothesis is rejected in favor of a competing hypothesis as more data became available and analyses were complete. Yet, more extensive analyses and significant additional data led to scientifically supported conclusions in our 'Proximal Origin' paper which was a peer-reviewed study published in *Nature-Medicine* in March 2020. Due to limitations on length and number of citations for the article format, not all analysis performed could be described and all relevant articles referenced.

The response by Andersen in his Tweets is in a considerable measure his attempt to fight for scientific truth and its methods in addressing the repeated attempts by politically motivated scientists using their access to the bourgeois press to put forward their unreviewed and biased conclusions, which have dangerous geopolitical consequences.

In a case in point, the *Wall Street Journal* published an opinion piece by Stephen Quay and Richard Muller on June 6, who wrote, "The Chinese Communist Party has been reluctant to release relevant information.

Reports based on US intelligence have suggested the lab collaborated on projects with the Chinese military." Actually, US intelligence has put forth only hypothesis based on low-level evidence, meaning these remain in the realm of speculation.

The pair go on to write:

In gain-of-function research, [*which Dr. Shi has emphatically denied conducting*] a microbiologist can increase the lethality of a coronavirus enormously by splicing a special sequence into its genome at a prime location. Doing this leaves no trace of manipulation. But it alters the virus spike protein, rendering it easier for the virus to inject genetic material into the victim cell. Since 1992 there have been at least 11 separate experiments adding a special sequence to the same location. The end result has always been supercharged viruses.

In the case of the gain-of-function supercharge, other sequences could have been spliced into this same site. Instead of a CGG-CGG (known as "double CGG") that tells the protein factory to make two arginine amino acids in a row, you'll obtain equal lethality by splicing any one of 35 of the other two-word combinations for double arginine. If the insertion takes place naturally, say through recombination, then one of those 35 other sequences is far more likely to appear; CGG is rarely used in the class of coronaviruses that can recombine with CoV-2.

Once more, the issue of the furin cleavage site (FCS) reappears.

In response to Quay and Muller, Dr. Andersen replied, "The FCS itself is not an optimal site (for cleavage) and has never previously been used in CoV experiments to the best of my knowledge - unlike more optimal sites, which have been inserted into SARSr CoVs for basic research."

Is the FCS a "smoking gun"?

Dr. Andersen also takes up the comments of Dr. David Baltimore, a Nobel laureate and biologist who was widely cited by "lab leak" advocates when he called these genetic findings "the smoking gun" showing laboratory manipulation. He supplemented this assertion stating, "these features make a powerful challenge to the idea of a natural origin for SARS-CoV-2."

But more recently, according to the *Guardian*, Dr. Baltimore has walked back his statement attempting to strike a more balanced position to protect his reputation. In an email exchange with the *Los Angeles Times*, he wrote, "I should have softened the phrase 'smoking gun' because I don't believe that it proves the origin of the furin cleavage site, but it does sound that way. I believe that the question of whether the sequence was put in naturally or by molecular manipulation is very hard to determine, but I wouldn't rule out either origin." With the journal *Nature*, he refined his position, "There are other possibilities, and they need to [have] careful consideration, which is all I meant to be saying."

Placing these retractions in context, the *Guardian* wrote, "Given his considerable reputation, Baltimore's dramatic 'smoking gun' quote in early May had driven a lot of the resurgence of interest in the Wuhan lab leak theory in tandem with renewed reporting of unverified intelligence claims that three staff at the Wuhan Institute of Virology were hospitalized in November 2019 with symptoms consistent with COVID-19 or seasonal flu."

Continuing with Dr. Andersen's comments, which rebut Dr.

Baltimore's initial comment and elaborate, albeit in highly technical shorthand, why the hubbub over the furin cleavage sites is entirely misguided:

[However], the exact same FCS found in SARS-CoV-2 can be found in different viruses, including Feline coronavirus (FCoV), which is an alphacoronavirus. FCS isn't optimal and while it's "sufficient" for SARS-CoV-2's "success" as a pandemic virus, it's not an ideal site as defined by the canonical R?X?K/R?R FCS seen in many proteins (viral and otherwise). Importantly, however, in recent month we have started seeing the CoV's [FCS] mutating towards residues creating more optimal furin sites - P681H and, especially, P681R, which can be found in B.1.1.7 and B.1.617.x, suggesting the virus may evolve towards more efficient usage of the site.

So, Baltimore's first point—that the FCS found in SARS-CoV-2 is somehow unusual—is simply incorrect. FCSs are found in a multitude of different coronaviruses, indels [*a molecular biology term for an insertion or deletion of these nucleotides in the genome*] come and go frequently, and the exact (P)RRAR can be found in other coronaviruses.

Now, the codons. Here, Baltimore (and Quay/ Muller) is talking about the two codons coding for the first two arginines (R) following the P - CGG. The CGG codon is rare in viruses because it's an example of an unmethylated "CpG" site that can be bound by TLR9, leading to immune cell activation. Despite being rare, however, CGG codons *are* found in all coronaviruses, albeit at low frequency. Specifically, of all arginine codons, CGG is used at these frequencies in these viruses: SARS: 5% SARS2: 3% SARSr: 2% ccCoVs: 4% HKU9: 7% FCoV: 2% Nothing unusual here.

One final point about the CGG codons in the FCS - if they were somehow 'unnatural', we'd see SARS-CoV-2 evolve away from 'CGG' during the ongoing pandemic. We have more than a million genomes to analyze, so what do we find if we look at synonymous mutations at the 'CGG_CGG' site? Remarkably stable. Specifically, CGG is 99.87% conserved in the first codon and 99.84% conserved in the second. This is *very* strong evidence that SARS-CoV-2 'prefers' CGG in these positions.

So, Baltimore's second point is also false, invalidating his hypothesis that the 'FCS ... with its arginine codons ... was the smoking gun for the origin of the virus.' Baltimore does not provide any evidence to support his hypothesis and the data support a natural origin.

Does this disprove a lab leak? No. However, it disproves there being a 'smoking gun' in the FCS and lends further evidence to natural emergence - but it also does not *prove* that scenario. *To this day, we have yet to see any scientific evidence supporting a lab leak.* [Emphasis added]

In response to the WWS's question regarding assertions made by scientists on the SARS-CoV-2's genetic stability in humans, Dr. Andersen explained, "When it [SARS-CoV-2] spilled over, it is incorrect to say it was 'well adapted to humans.' We know this because 1) The emergence of variants of concern and human adaptation that is ongoing, 2) the virus can jump between species with no evolution—e.g., mink, and 3) Pangolin CoVs bind even stronger to human ACE2 receptors."

To be continued



To contact the WSWs and the
Socialist Equality Party visit:

wsws.org/contact