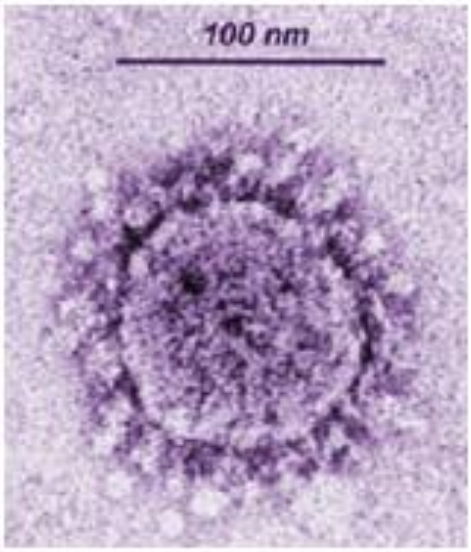
Analysis of Wuhan Coronavirus

Deja Vu



William R. Gallaher, Ph.D.
and
Andrew D. Gallaher

Comments and inquiries should be directed to:

William R. Gallaher

Mockingbird Nature Research Group

PO Box 568

Pearl River, LA 70452 USA

profbillg1901@gmail.com

Cover photo: Electron micrograph of Coronavirus, Centers for Disease Control and Prevention, US Federal Government, Public Domain

Preface

As this is intended as a serious scientific article, analyzing a deadly viral disease, a preface on qualifications is in order.

I, WRG, the first author, have been involved in experimental virology since the summer of 1967. My first paper (Bratt and Gallaher 1969) was communicated by Prof. John Enders while I was a grad student. I hold a Ph.D. in Microbiology and Molecular Genetics from Harvard University, since 1972, having done my graduate work at Harvard Medical School in Boston. For purposes of present identification, I have held a faculty appointment in the Department of Microbiology, Immunology and Parasitology of LSU Schools of Medicine and Dentistry, New Orleans, continually since August of 1973. I formally retired after 32 years of active service, but continue to work as Professor Emeritus and publish in peer-reviewed scientific journals. In 2008 I established Mockingbird Nature Research Group as a Louisiana Corporation, for collaboration and consultation outside the aegis of LSU, my former employer.

Especially when expressing opinions, as here, I do not represent LSU and state explicitly that my views are entirely my own. I take full responsibility.

For the present article, I have decided to go outside of the peer-review system and publish this directly. Not only do I avoid delays and dialogue with editors, but also the expense of professional publication, which can exceed $2000, in my case from personal funds. I can also feel free to express myself more personally. Most of what I publish on Amazon Kindle

is fiction, my retirement second act. This article is not fiction. It is serious science, as I am trained and experienced to conduct and report.

Acute viral respiratory disease is very personal to me. Influenza has nearly killed me more than twice, and in 1965 came close to doing so. My experience with the Asian flu as a 12 year old witness to the 1957 pandemic, as well as a patient later, was an important motivator in deciding to become a virologist. Viral pathogenesis has been my consistent passion for 53 years.

In 1967 I watched my first viral infection of a monolayer of cells growing on the bottom of a glass prescription bottle in a warm room, periodically peering through an inverted microscope. For 6 hours nothing happened, then cells began to change shape, then fuse together, and then, by 12 hours after infection, all hell had broken loose. The monolayer detached from the glass and floated off as debris. I went home determined to find out how that minute virus, with very limited genetic material, could do that. I was also determined to someday find a way to stop it.

Over the course of my career, I achieved both objectives. I went on to yet other families of viruses with the same goals.

My earlier work was with animal viruses, such as Newcastle disease virus of chickens and mouse hepatitis virus, as experimental surrogates for similar viruses causing human disease. The emergence of AIDS brought me more into human viruses. I was first to publish the identification of the fusion and entry peptide of HIV-1 (Gallaher 1987) and thereby identify HIV gp41 as the fusion and entry protein. I was first to develop a structural model of HIV gp41, built on a scaffold of the influenza surface hemagglutinin, and thereby discovered the superfamily of viral

fusion/entry proteins (Gallaher et al. 1989) that have subsequently been called “Class I Fusion/Entry Proteins” by those who later confirmed the “Gallaher model” by high resolution x-ray crystallography.

I later extended the superfamily to Ebola of the Filovirus family (Gallaher 1996), and to the Arenaviruses such as Lassa fever virus (Gallaher et al, 2001).

As will be cited in the article, I was first to develop a detailed molecular structure of the S2 fusion/entry glycoprotein of SARS virus, within 24 hours of publication of its genomic sequence. I then collaborated and consulted with my colleagues in the labs of Dr. Robert Garry and William Wimley in characterizing membrane destabilizing regions of the SARS S2 glycoprotein.

When the “pandemic Influenza H1N1/09” emerged, I happened to be the founding Deputy Editor of Virology Journal, and published on May 5 a commentary on the outbreak (Gallaher 2009). The present article is intended to be presented with the same purpose and tenor, albeit with greater molecular detail about the novel Wuhan Coronavirus.

In 2014, I identified the Ebola Delta Peptide as a membrane destabilizing agent and cytotoxin more potent than cholera toxin. (Gallaher and Garry 2015; He et al. 2017).

Since 2016, I have collaborated with my son, Andrew D. Gallaher, in discovery of additional viral cytotoxic motifs, to which allusion will be made in the accompanying article. Since I am now 75 years old, he also provides valuable assistance in sequence research and analysis, as well as in preparation of manuscripts. Since June of 2019, he has been appointed a

Staff Scientist at Mockingbird, and is engaged in several ongoing investigations in his free time, while still serving his country as an active duty Master Sergeant in the United States Marine Corps. Given the nature of the current epidemic, he also brings an understanding of national security to the table.

William R. Gallaher, Ph.D.

January 29, 2020

DEDICATION

To

Mayinga N'Seka

Carlo Urbani

S. Humarr Khan

And so many others

Who care for their patients

With deadly viral disease.

Knowing the risk

They go in anyway

They do so even now.

No greater love.

Analysis of Wuhan Coronavirus

Deja Vu

I. INTRODUCTION

Well, here we go again. Emerging viruses happen.

On December 30, 2019, the People's Republic of China (PRC) released information that an outbreak of significant acute respiratory disease was occurring in Wuhan, a city of 11 million souls, in the southeastern central province of Hubei. At the time, the etiologic agent was unknown. However, since the outbreak was associated with a seafood and meat market that sold a variety of live, wild animals, it was feared that SARS had again erupted in mainland China.(Galinski and Menachery 2020)

On January 5, 2020, SARS coronavirus was ruled out as the etiologic agent, along with influenza or MERS (Middle Eastern Respiratory Syndrome) or other known viral agents of respiratory disease. On January 9, the World Health Organization (WHO) announced that a novel Coronavirus appeared to be the etiologic agent. The genome sequence of the viral RNA was released to Genbank the next day. (This article uses the third iteration of that single release sequence, dated January 20). The release of such proprietary information, that is normally held until publication, was an unusual and highly laudable public service gesture by those at Fudan University, Shanghai, responsible for the genetic sequencing. It is enabling an informed approach to developing intervention strategies against the virus by literally countless laboratories around the world. There is an unusual amount of information sharing, in a scientific world where confidentiality is more the rule. Additional sequences have

also been posted, and thus far are 99.5% identical to one another, supporting a clonal, single, one-time source for the virus.

Retrospectively, the first cases were detected on December 8, 2019. This would place the date of initial transmission from its animal source around Thanksgiving through December 1.

The animal source of the virus is almost certainly dead, and the live animal markets, long a cultural fixture in the Far East, are closed. A second source point for the virus is, for a time being, highly unlikely.

There are no currently licensed drugs or vaccines against Coronaviruses. A number of candidate drugs against SARS have been investigated, as well as anti-SARS antibodies, but none have even been tested for safety except for some in small animals, and there are no significant stockpiles remotely adequate to the task that is likely to be at hand. The Wuhan virus, currently being abbreviated as **"nCoV2019"**, for novel Coronavirus 2019, is ***not*** SARS; at the molecular level it is only 80% similar to SARS overall. However, as will be discussed below, in certain protein regions it has a much higher similarity to SARS, high enough that some anti-SARS strategies, or drugs directed at other RNA viruses already in development, may be of some use in treatment or prevention of nCoV2019 infection and disease. Indeed, even now, some combinations of pre-existing drugs developed for other viruses are being tried in the field on a compassionate-use basis.

Current state of epidemic

The development of the outbreak, both within China and exported to other countries, is a daily, even hourly, evolving phenomenon, literally

changing with every sentence I type. Case report data is necessarily a view of the past, not the present, no matter how prompt and conscientious the reporting. Coronaviruses typically have an incubation period, from time of exposure to onset of clinical symptoms, of between 2 and 10 days, on average 5 days to development of significant illness. So, with current data we are essentially looking at what happened a week ago, which in rapidly developing epidemics might as well be an eternity ago. The specific incubation period for the Wuhan virus is only beginning to become known, but it has shown itself capable of readily passing from human to human.

What appears clear from existing data, from the first few thousand or so clinical cases, is that this is a virus of high morbidity (clinical illness) but low mortality (death). It is not unlike influenza in this regard thus far, except with perhaps somewhat higher mortality concentrated in compromised patients, i.e. elderly, cardiopulmonary compromise, or infants. It is not clear whether deaths are caused by the virus infection itself, or a result of pre-existing illness or opportunistic bacterial superinfection, as is common with flu. If this pattern persists, it is a good bit less virulent than SARS was in 2002-2003, when a 10% mortality was observed. Advanced respiratory supportive care, such as that commonly available in the United States health care system, would be anticipated to be effective in combatting the disease even in the absence of specific antivirals.

So, then, this is not SARS, or MERS, or Ebola. It does cause acute respiratory disease that currently requires hospitalization to control, characterize and quantify the disease and its spread. On the one hand, it

has already killed many; on the other hand, many have already fully recovered and been released from care.

Ironically, a less virulent virus is harder to contain. For most viruses, the most efficient period of spread is what is termed the “prodrome”, a day or two before development of frank symptoms, when the virus is already extensively replicating in the respiratory system of an individual, and the individual is shedding virus in respiratory droplet secretions. Two or three days of shedding virus may precede the patient presenting themselves to a clinical setting. By that time, the virus has already moved on to its next victim(s). The average number of secondary infections from an individual is known as the virus’s “R_0 value” (pronounced “R naught”). For pandemic influenza H1N1(2009) the R_0 value was 1.4 to 1.6 – each person infected on average about 1.5 other human beings. The R_0 can be inferred from an epidemic profile of increasing case incidence, but is best determined only retrospectively. The R_0 for nCoV2019 is unknown, and may not be clear for some time. However, an R_0 equal to or greater than that of pandemic flu would not be surprising.

There is no way to reliably predict the future course of the outbreak, as there are too many variables in play. Chief among these is the capability of nCoV2019 to mutate, as RNA viruses are well known to do (Goba et al. 2016), and better adapt to human infection and spread through the human population. This began as an animal virus trying to make its way through the human population. It languished for a while, but now it is truly becoming a human virus. The more it remains an animal virus, the course of the outbreak will be flatter and self-limiting in response to efforts to suppress opportunities to spread to new victims. By this time, it must be

admitted that it is spreading exponentially, as a very efficient human virus would.

At the beginning of the outbreak, few if any human beings globally had any prior exposure or immunity to the virus. Everyone is susceptible.

Despite rapidly increasing case reports, there is hope, however. The initial outbreak occurred over a month, indeed nearly two months, ago in the center of a major city in China, a high-density population of 11 million, less than a half mile from the central high speed rail station of a major transportation hub for China. Even if the actual number of cases is now 20,000 (already greater than all SARS cases), given the reporting lag, this is a small fraction of the population within which it emerged and with whom it had contact by high speed rail. I concur with statements made by a number of US health professionals, such as Dr. William Shaffner of Vanderbilt and Dr. Anthony Fauci, longtime Director of the US National Institute of Allergy and Infectious Diseases, urging a measured response and remaining calm.

Even as a virus mutates and adapts to a new host, one characteristic that does not generally change is its inherent virulence. As cases have increased in number, the relatively small percentage of critical patients, or of deaths, has not changed significantly relative to the total caseload, i.e. 20% severe illness in reported cases, 3% mortality.

Expressed more positively, a patient hospitalized with acute respiratory disease due to nCoV2019 has a 97% chance of recovery, probably higher if they are neither very old nor very young, nor afflicted with a preexisting cardiopulmonary illness.

In our recent experience is Ebola 2014 that, in contrast, exhibited a very high percentage of apparent illness, virtually 100%, and mortality of 50% (Goba et al. 2016). So nCoV2019 is nothing like Ebola in terms of severe illness or mortality.

Much of what I could say in support of a sane and rational approach, and against an atmosphere of hysteria, I already addressed in response to the pandemic influenza H1N1(2009) (CDC 2009; Aras et al. 2009) in May of 2009. Rather than repeat myself, I refer the reader to my comments at that time, publicly available for free (Gallaher 2009). For much of that 5000 word commentary, one can simply substitute nCoV2019 for pandemic influenza H1N1(2009) as the basis for approaching the current outbreak.

Common sense, as in covering a sneeze or cough, limiting exposure to crowds and close (less than 3 feet) contact to others, and, perhaps most importantly, frequent hand washing and use of hand sanitizer, will do more than boxcars of face masks and latex or nitrile gloves. Infection control can be as simple as never touching your nose with your fingers; many do so incessantly, potentially inoculating themselves with someone else's fresh respiratory droplets containing their freshly produced infectious virus.

When WRG sees a crowd of people wearing face masks in pictures and television, or in an overcrowded venue or an emergency room waiting room, he feels like screaming "Get away from all those people!" Too often the mask or glove induces us to take chances our common sense should tell us not to take. Unless you are a health professional, if you feel you need a face mask, your common sense is telling you that you should not be there!

Wearing a face mask in a dense crowd is rather like a man taking condoms into a bordello, and feeling safe. Nothing about his decision to visit a bordello is safe.

Avoid crowds whenever possible, and try to maintain a personal space on the edge, facing away from others.

A lot of people in close contact is what a virus regards as lunchtime. The most predictable result of Super Bowl, or Mardi Gras, or Sunday at a hugging and kissing church, with all those newly infected people mixing with all those new susceptibles from elsewhere, is spread of viral illness. It is not a matter of ***if***, but only how much. Viruses need to find a new host quickly, within a day or two, or become extinct. Most human viruses manage to do that incessantly, which is why they are still around. We make it easy for them. Quite simply, don't make it easy for them.

In the wake of SARS (Rota et al. 2003; Tsang et al. 2003; Ksiazek et al 2003; Poutanen et al 2003) and Ebola (Goba et al 2016), we have also learned a great deal about intercepting imported emerging viruses and screening arrivals from outside the US or across any international border. Every hospital and clinic, every health professional, has received training and drills in well-developed protocols for dealing with imported viral agents far more deadly than nCoV2019 now appears to be.

As Dr. Shaffner has reminded us, even if more nCoV2019 should reach our shores, influenza virus is already among us and a far greater danger to Americans (Kilbourne 2006; Taubenberger and Morens 2006). Flu kills ten or more thousands of Americans every year. Over the decade since 2009, the global death toll of pandemic influenza H1N1(2009) has

been well over 300,000 persons. Indeed, flu is not measured in cases, but in deaths due to influenza/pneumonia. Each winter, we should already be using our common sense and the measures listed above, as well as getting the flu vaccine, to limit our exposure to a dangerous viral agent that is already in our neighborhood.

On the other hand, the PRC has announced that all 70,000 theaters in China are to be closed, and a number of cities in China, with an aggregate population of over 35 million, have been placed on lockdown in an effort to suppress the outbreak. The Lunar New Year, that began January 25, is a huge deal in China; this year events are closed and travel severely curtailed. We can reasonably assume that this reflects private briefings given to the Chinese leadership which inspired such drastic measures. Containment may be difficult; indeed, the genie may never be returned to the bottle from which it emerged.

As of January 26, three cases, all originating in China, are in isolation in US hospitals. There are a few such cases in many countries, with many more suspected. More are doubtlessly coming. It has been documented that the virus may be spread before its victim shows any signs of illness. (https://en.wikipedia.org/wiki/Timeline_of_the_2019%E2%80%9320_Wuhan_coronavirus_outbreak)

Regardless of the future course of the Wuhan nCoV2019 outbreak, whether it explodes or fizzles in the face of draconian public health measures, it will be at least prudent and probably essential to our national health security to better understand the specific nature of the virus. We need to explore in detail its mode of infection and develop antiviral strategies to inhibit or prevent further spread and future outbreaks. If the

20th Century has taught us nothing else, it is that emergence of a virus happens repeatedly. Even if it goes away, it will be back. Somehow, some way, someone will go back and get it.

Culture is immutable. Those live animal markets will reopen one day or flourish on the black market. Emerging viruses happen. SARS is still out there. That Asian flu (H2) is still out there, even though the human population has not experienced it since 1967 (Kilbourne 2006). As human populations continue to increase, we impinge on environments and animal populations we have never experienced before.

The following is intended to apply our long-developed insights into Coronavirus infection, in specific molecular terms, to aid in the development of antiviral strategies to have on hand when nCoV2019 comes our way, sooner or later.

II. CORONAVIRUSES OF HUMAN RESPIRATORY DISEASE

Coronaviruses comprise a diverse family of viruses, in both animals and humans, that use RNA as their genetic material. They consist of a viral RNA-protein core that is surrounded by a membranous envelope. They are named for their appearance in electron micrographs, as shown on the cover of this article, spheroid particles festooned with extended surface projections, resembling the solar corona. The projections are surface proteins of the virus that facilitate attachment and entry into host cells, and are called “spikes” and “spike proteins (de Groot et al. 1987 Song et al. 2018). The spike protein complex of nCoV(2019), compared to that of

SARS, will be discussed in some detail later. A general outline is shown in Figure 1.

Figure 1.

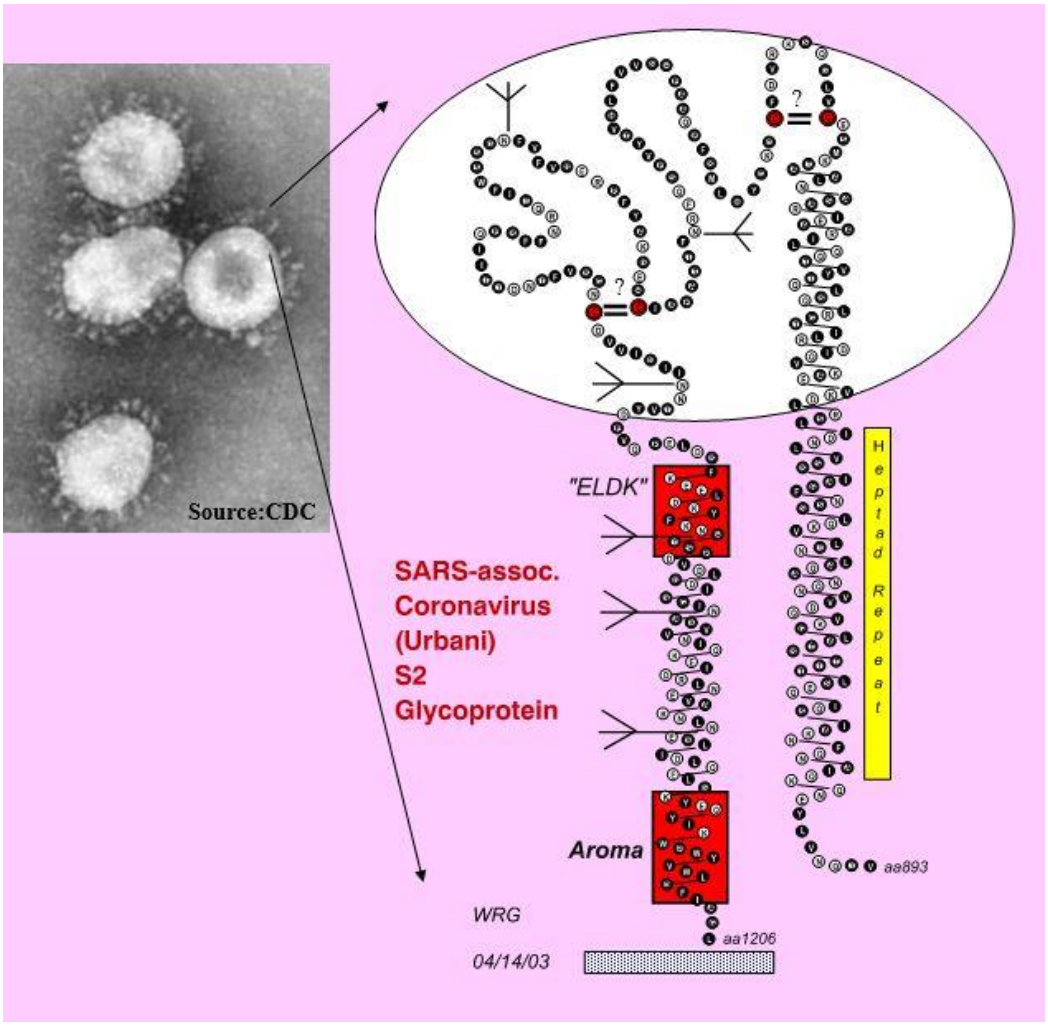


Figure 1: On the left is shown an electron micrograph of three enveloped Coronavirus particles, from the CDC, showing the prominent surface spikes. On the right is a blown up cartoon of one monomer of the spike protein complex, as described in the text. The modeling methodology is described in antecedent papers modeling corresponding proteins of HIV-1, Ebola and Arenavirus (Gallaher et al. 1989; Gallaher et al. 1995; Gallaher 1996; Gallaher et al 2001).

The spike consists of two proteins, a globular head group about 160 kilodaltons in size, S1, shown here simply as an oval, and a fibrous leg

region of about equal size, S2, illustrated for SARS virus in greater detail. This is the first molecular model of SARS S2, drawn as two antiparallel alpha helices of exceptional length, in what turned out to be its post-membrane fusion configuration. This complex is discussed in far greater detail below.

A single S1/S2 protein complex, as illustrated, constitutes only one of three monomers of S1/S2 that form a trimeric structure to form a single spike on the surface of the virus (Song et al. 2018). Each spike is therefore three very long polypeptides, each over 1200 amino acids long. Consisting of over 3600 amino acids in all, with an aggregate molecular weight of over 1 million daltons, there is little wonder the trimeric spikes are so prominent on the surface of the virus.

The viral RNA genome is a unique, single-stranded RNA molecule that is by far the largest known, about 30,000 nucleotide bases long. The replication and expression of this huge RNA is complex, and the virus encodes many non-structural proteins (nsp) to accomplish it. These are generated by endoproteolytic cleavage of large precursor proteins using a viral protease. The structural proteins of the virus are made separately. They include the spike protein complex (S), a membrane (M) protein and the core nucleocapsid protein (N) as principal components.

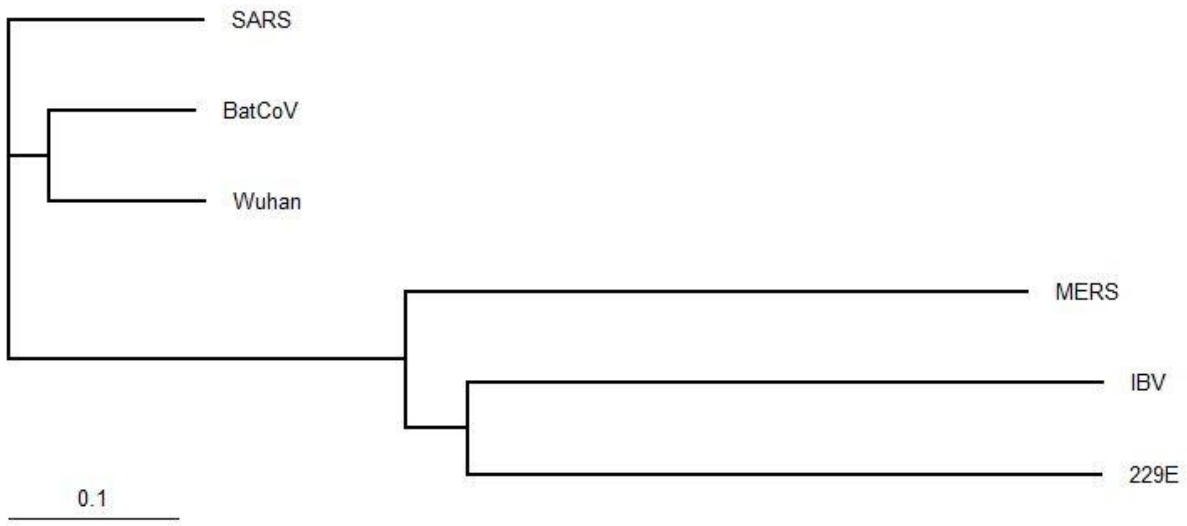
Coronaviruses appear to have diverged most significantly at the end of the most recent Ice Age, about 8000 years ago. The RNA and protein sequences can be quite different, while maintaining similar structure and function. With at least one cycle of infection per day, each virus today is the product of millions of replicative cycles, while capable of generating multiple mutations in its genome each cycle.

There are seven different Human Coronaviruses, each subdivisible into separate strains. The first two, 229E and OC43, were discovered in the 1960s by Tyrrell and others in surveys of volunteers for common cold viruses. Strains of each together contribute about 30% to the common cold throughout the world. They rarely cause serious infections.

The other five Human Coronaviruses have only been discovered in the 21st century. SARS in 2002, NL63 in 2004, HKU1 in 2005, MERS in 2012, and nCoV2019 only last month. These tend to cause more lower respiratory infection, with SARS, MERS and nCoV2019 the most serious trend towards pneumonia and critical disease. Each of the last three are documented to have crossed over from animals to the human population when first discovered. SARS proved itself quite capable of human to human spread, and nCoV2019 appears to be similar in that regard. MERS, derived from Dromedary camels, has less potential for human to human spread.

The immediate source of SARS in 2002 was palm civet cats, wild animals in captivity. The immediate source of nCoV2019 is still unknown. However, both SARS and nCoV2019 are most similar to a group of bat Coronaviruses as the probable ultimate source in nature. Indeed, mCoV2019 is 88% similar to a Bat coronavirus, while only 80% similar to SARS. A rough family tree of the spike protein region of SARS, MERS, nCoV2019, BatCoV and 229E, using chicken infectious bronchitis Coronavirus (IBV) as an outgroup, is shown in Figure 2.

Figure 2.



It can readily be seen that SARS, BatCoV and the Wuhan virus cluster separately from the others, with the Wuhan nCoV2019 virus clustering most closely with the batCoV sequence. So we are dealing with a bat virus gone rogue, and not a virus derived in any way from previously existing Human Coronaviruses. Absent the special circumstances of the wild animal markets in China, it was very unlikely that humans would have come into contact with SARS or nCoV2019 at all, even in an extraordinarily populous place as mainland China.

SARS was quickly eliminated from the human population in 2003, due to an extraordinary public health effort and outright heroism. How we will fare with the newly arrived nCoV2019 is an open question, but the same measures that eliminated SARS and, more recently, Ebola in West Africa, from the human population are now underway in earnest. There is no shortage of heroism among medical staff coming forward to treat infected patients, in spite of the obvious danger to themselves. There are

press reports of illness among medical staff, but not yet any identified deaths among them.

This is personal to myself and my colleagues. On the description of the 2014 Ebola outbreak in West Africa, where I was privileged to be included to be one of many co-authors, the first author was living, but the next five authors, led by Dr. Khan, died in the course of trying to help Ebola patients (Goba et al. 2016).

No greater love.

III. THE CRITICAL CONCEPT OF VIRAL LOAD

Decades of study has demonstrated in diverse systems the importance of the concept of viral load. Viral load is defined as the concentration of viral genomes in a patient at a given point in time. In the case of HIV-1, the virus has almost never been eliminated from an infected individual. However, even in the case of the less effective early antiretroviral drugs, there was improvement in patient health by reducing their viral load. Patients lived longer and more normal lives, even if many ultimately succumbed.

Once the protease inhibitors and combined therapy were introduced as antivirals in the late 1990s, it has been possible to reduce HIV viral load to an undetectable level, without actually curing anyone of the virus. But even patients with detectable, but lowered, viral load may show marked improvement. Antiviral therapy has changed a uniformly fatal infection into a manageable one, provided the patient is compliant with that therapy.

During the 2014 Ebola outbreak in West Africa, it was found that older patients did less well than younger, and compromised patients less well than previously healthy patients. The unifying factor underlying these statistics was shown to be viral load. Patients with 1 million or more genomes per ml of serum did less well and showed high mortality; those who for some reason displayed a viral load under 1 million per ml of serum had better prospects and often recovered (Gobs et al. 2016).

Bottom line here is that, while reducing viral load to undetectable levels in a laudable goal for any prevention or treatment program that might be deployed against nCoV2019, it may well be just as good to accept reduction of viral load below a certain level correlated with serious disease. As Voltaire said, “Do not let the perfect be the enemy of the good.” Anything that helps, helps. Reduction of critical illness and mortality is the ultimate goal, even if elimination of any level of illness or total control of nCoV2019 eludes us.

IV. COMPONENTS OF CORONAVIRUS AND ANTIVIRALS

The following is about to get more technical, but an effort will be made to make it accessible to one with very little or no science background. To find out how a Coronavirus protein is like a “Transformer”, stay tuned!

A number of potential targets present themselves within the genome and protein products of nCoV2019 for either vaccines, protective antibodies, or antiviral inhibitors. These targets are modeled after similar approaches that have been used against other viral infections in the past, or approaches that have been developed in the event that SARS should return.

We will discuss each in turn. In the process we will cover proteins that are encoded by approximately 25% of the viral genome.

Before the advent of SARS, in late 2002, all of the active Coronavirologists in the world could have fit into a single large, university classroom. So, much of the antiviral approach is derived from other enveloped viruses such as HIV-1 and influenza virus. It so happens that the correlate in HIV (Kowalski et al. 1987)\., or in influenza (Wilson wt al. 1981; Gething et al. 1981), of the spike complex is a similar, albeit much smaller, version of the S1/S2 complex. In HIV and other retroviruses the globular head group is called SU, for surface, and the fibrous leg region, TM for transmembrane. In flu they are called HA1 and HA2, respectively. Given the great importance of the two latter viruses to human health, we know a great deal about how such a spike protein complex works and how its function can be inhibited or an immune response mounted against it (White 1992; Eckert and Kim 2001; Morrison 2003). Almost immediately after the SARS emergency began, virologists moved into the study of Coronaviruses, many with experience in these other relevant viral systems. Because of SARS, there is now no shortage of virologists or other medical scientists to turn their attention to nC0V2019, and they can be depended on to do so in droves.

1. Spike Glycoprotein

Overall similarity

The arrangement of a globular attachment protein with fibrous fusion/entry protein is an incredibly ancient molecular machine for specific transit across the cellular plasma membrane. We know this from endogenous retroviruses that infected animals long ago in geologic time and were incorporated into the animal genome. In many cases, the fusion mechanism was hijacked by the animal for its own purpose of fusing cells in the cellular layer of the placenta that separates the maternal from the fetal blood circulation. The human genome is littered with an enormous amount of what was originally retroviral genome, RNA made into DNA, and then embedded into the primate genome.

Two of these captured retroviral SU/TM complexes are known as Syncytin-1 and Syncytin-2, on human chromosomes 7 and 6, respectively (Mi et al. 2000; Renard et al. 2005). Their expression is controlled by human regulators, and only occurs during pregnancy, expressed in syncytiotrophoblasts of the placenta. They are immunosuppressive in that location, and are actually responsible for the failure of a mother to reject the tissue of her non-identical fetus. Based on the geologic timeline for development of primate species, we are fairly certain that these SU/TM complexes, homologous to freely circulating Retrovirus Group D viruses of today, entered the primate genome 40 to 50 million years ago. A similar SU/TM complex in carnivores entered the carnivore genome even earlier, up to 60 million years ago. Yet, the structure and even the protein sequence of the endogenous retroviruses is eerily similar to presently circulating viruses, including being exactly the same length and structure.

The retroviral SU/TM complex is half the size of that in Coronaviruses. Usually in evolution, smaller is later and more efficient. So

the S1/S2 complex may be far more ancient than retroviruses, perhaps back beyond the Cretaceous/Tertiary (K/T) boundary at the great extinction event that occurred 65 million years ago. The viral attachment/fusion machine may have originated in some Jurassic Virological Park, and conserved in form and function ever since (Shi et al. 2018). The point being that its principal functional parts are extremely well preserved over time in each virus that uses the complex for attachment and entry. What one learns about one frequently applies to all of the others, albeit with some protein sequence variation.

The same model I proposed for SARS has been found to be equally applicable, to some degree, to a wide variety of enveloped viruses that use what has been termed the Class I Fusion/Entry Glycoprotein complex and its accompanying receptor-binding globular attachment protein. (Hsu et al. 1981; Collins et al. 1984; Moscona et al. 1992; Bousse et al. 1994; Bousse et al. 1995; Morrison 2003; Eckert and Kim 2006).

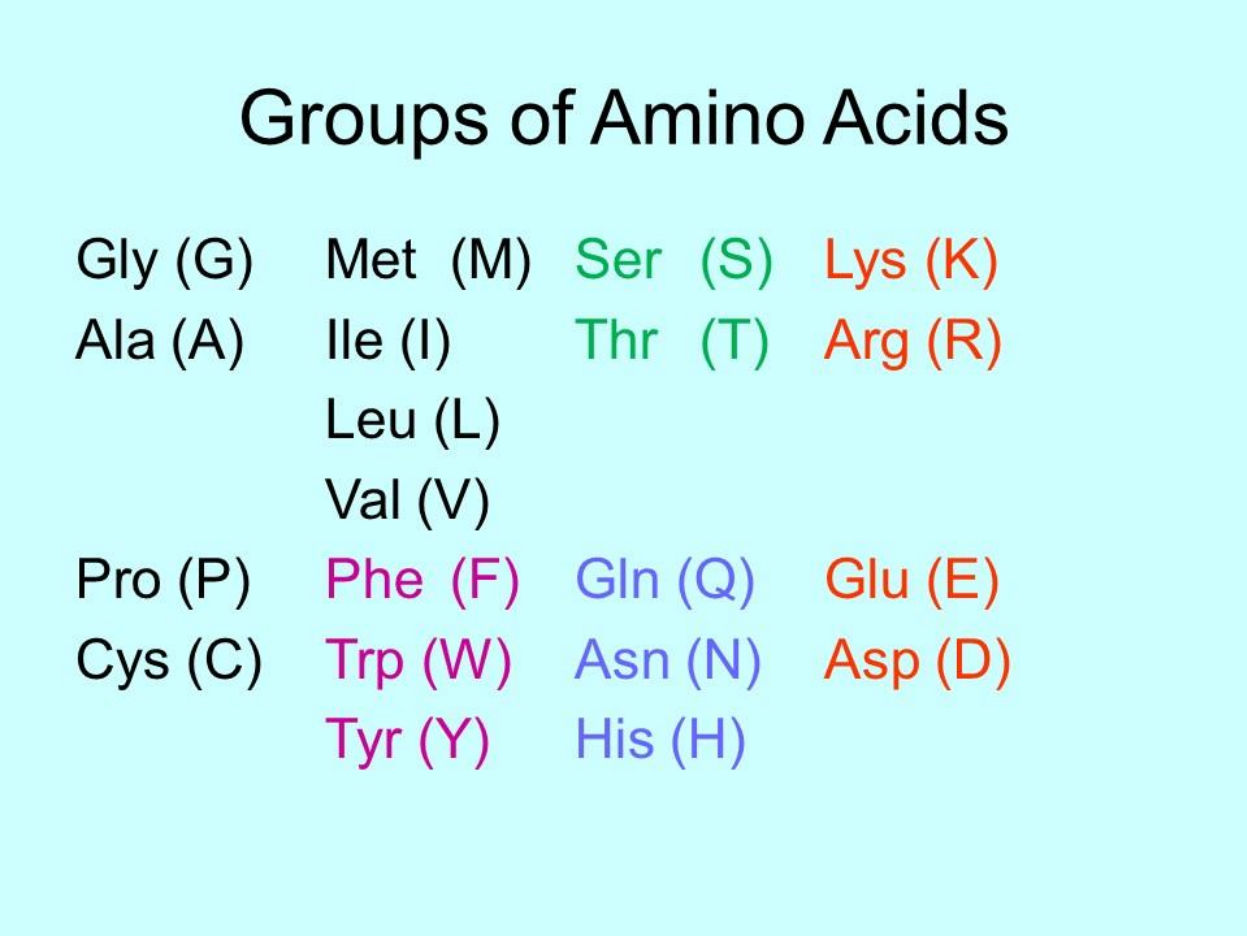
Figure 1 illustrates this in the protein sequence ELDK highlighted on the shorter helix. The reason for the highlighting is that the sequence ELDK is also found in a similar position conserved in HIV-1, where it is part of a site for antibody neutralization of diverse strains of HIV-1. The next amino acid in SARS is Y, while in HIV-1 it is W, both in the same group of aromatic amino acids. It is not unusual to be able to jump between dissimilar virus families and find comparable peptide regions of both, in both function and even sequence. They are, after all, cousins with the same job in viral infection. Examining comparable amino acid sequences in different viruses has been a key method is discerning the form and function of a novel virus such as SARS or nCoV2019.

Specifically with respect to comparing SARS, on which a great deal of knowledge has accumulated over the last 17 years, to the novel Wuhan Coronavirus, there is a good deal of similarity that allows us to go back and forth between one virus protein sequence and the other, using molecular landmarks.

2. The Amino Acids and Their Properties

Proteins are constructed of a series of amino acids in one continuous string synthesized together in the cellular protein synthetic machinery of polyribosomes. The front end first to be synthesized is called the N-terminus, because the nitrogen at one end of each amino acid remains exposed, while the other end is called the C-terminus, because the C at the other end each amino acid in the growing chain is exposed. Synthesis always goes N to C terminal, generally shown as left to right. In the example above from SARS, ELDKY would be a five-amino acid peptide, with the E N-terminal and the Y C-terminal. The letters are from the single letter code for the 20 different amino acids that commonly occur in human and animal proteins. The single-letter codes for the amino acids are shown in Figure 3, grouped into 8 separate clusters of amino acids with similar properties.

Figure 3



All amino acids are built on the same backbone that forms the protein chain itself. They are differentiated on the basis of the very different side chains that impart specific molecular character to each one.

The first group is composed of Glycine (Gly, G) and Alanine (Ala, A) that both have very short side chains. In the case of glycine, none in fact, just a hydrogen. In the case of alanine, a single methyl group of only three atoms. Glycine is effectively a spacer, allowing free rotation for that spot in the protein sequence. Alanine provides very little bulk, and fits almost anywhere.

The second amino group to the right comprises the aliphatic series of amino acids, imparting hydrophobicity (greasiness) to their position in the

protein chain. Valine (Val, V) is the smallest, just three more atoms than Alanine, and Leucine (Leu, L) three more. Leucine is one of the most common amino acids in proteins, providing basic hydrophobic bulk wherever its place in the protein chain. Both V and L are symmetrical in shape. Isoleucine (Ile, I) is similar in size to Leucine, but asymmetrical. Methionine (Met, M) differs from the others in having a sulfur atom near its outer end, rather than a carbon. It is distinguished by always being the first amino acid in any protein chain, because gene expression always begins with RNA that codes for it.

At the lower left are two amino acids grouped for their uniqueness, while at the same time being hydrophobic. In Proline (Pro, P) the backbone atoms are cyclized into a ring structure. Instead of free rotation around each backbone bond, the two ends of Proline are locked into a 130 degree angle to one another. Proline creates kinks in the protein chain at critical locations. Cysteine (Cys, C) is unique in that it has a terminal sulfhydryl group (-SH). Two cysteines can become covalently bound to one another, formed an S-S or disulfide bond between different regions of the protein chain, locking them together in a fixed configuration to one another. Cysteines are often highly conserved landmarks in proteins very important for stabilizing secondary structure.

To the right of P and C are Phenylalanine (Phe, F), Tryptophan (Trp, W), and Tyrosine (Tyr, Y), the aromatic amino acids with either a planar benzene ring or an indole double-ring for highly hydrophobic side chains. In this regard, W might well stand for whopper. It is much larger than any other hydrophobic side chain. Wherever it is found constitutes a veritable

center of hydrophobicity. It has a high natural affinity for cholesterol found in cellular target membranes.

The four groups to the right of the above groups in Figure 3 are more hydrophilic, and tend to be found on the outside of proteins. Next at the top are Serine (Ser, S) and Threonine (Thr, T). These are hydroxylated amino acids (-OH). Apart from readily interacting with water, they may also be the site where polysaccharide adducts can be added to the protein chain in what is called an O-glycosidic linkage, effectively sugar coating to that region of protein.

Below S and T are Glutamine (Gln, Q), Asparagine (Asn, N) and Histidine (His, H). These are mostly neutral amino acids with secondary amines. Q is notable because it has a strong propensity to be part of an alpha helix; N, because it can serve as a site for N-linked polysaccharide adducts, another type of sugar coating. H is relatively rare, with an imidazole ring for a side chain that can impart a slight charge. H is frequently found where proteins interact with some sort of substrate, with H as the reactive group on the protein.

On the top right are the two basic amino acids, Lysine (Lys, K) and Arginine (Arg, R) for which the side chains end in a free amino groups, imparting a positive charge to that part of the protein chain. Since cell surfaces are negatively charged, K and R have a natural affinity. They are also sites for endoproteolytic cleavage of proteins by cellular proteases similar to trypsin and furin. As such, they have a key role in maturation of viral fusion/entry proteins when in certain critical locations. Arginine is very large, comparable to tryptophan in size, but on the hydrophilic side.

Finally, on the lower right are Glutamate (Glu, E), and Aspartate (Asp, D), the acidic amino acids that terminate in a carboxylic group (-COOH) and impart a negative charge to the protein chain. In terms of protein structure, they differ significantly in that E has a strong propensity (like its neutral homologue Q) to reside in alpha helices, whereas D, only shorter by a three atom methylene group, much less so.

When modeling for alpha helices (see Lupas 1996), such as those shown in Figure 1, my basic rule has always in fact been simple, “watch your Es and Qs”, especially when clustered with A, L, F, W and K.

G, P, S, and T are often found in turns, especially when clustered together.

The other amino acids are more malleable in terms of protein structure, but clusters of I, V, Y, and M are common in beta-pleated sheet regions.

To those who study protein sequence and structure, the amino acids are not just beads on a long string, in the case of S1/S2 over 1200 amino acids long. The sequences are interpretable in terms of character, structure and function of each particular stretch of protein, as we are about to examine in comparing the S1/S2 sequences of nCoV2019 and SARS.

3. **Model of nCoV2019 Spike Protein**

Protein modeling has come a long way. We are now able to take known structures from one protein, the S1/S2 of SARS, and create a model of a novel but similar protein, the S1/S2 of nCoV2019, that will not be very

far from what x-ray crystallographers are likely to determine months or years hence.

Figure 4 is a Swiss-Model computer-generated structure for the nCoV2019 S1/S2 spike protein complex, presented here courtesy of my longtime collaborator, Dr. Robert F. Garry of Tulane School of Medicine in New Orleans.

Figure 4.

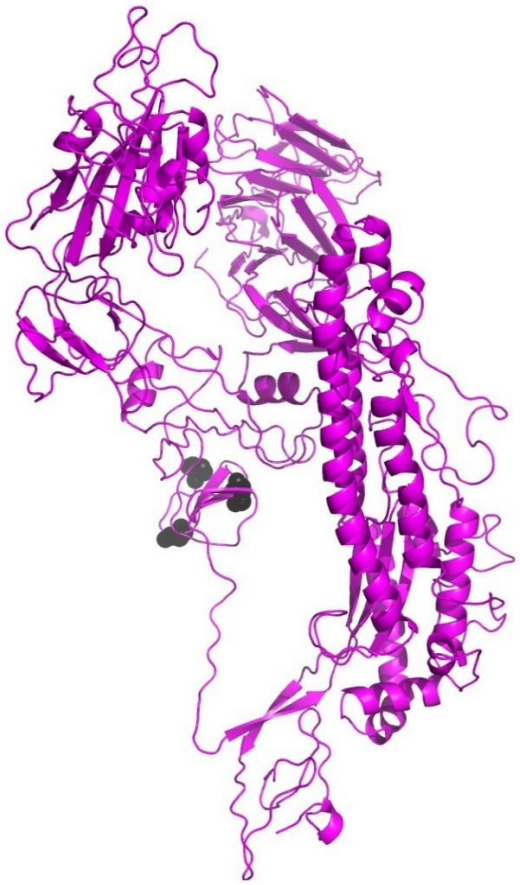


Figure 4. Ribbon model of the monomer S1/S2 Glycoprotein Complex of Wuhan Coronavirus (nCoV2019) via the Swiss Protein suite of modeling programs, based on the archival known structure of the SARS S1/S2 determined by x-ray crystallography. (courtesy of Dr. Robert F. Garry)

In Figure 4, the globular S1 N-terminal protein is on the upper left, while the fibrous S2 C-terminal protein is slightly below it and on the right. S1 consists mostly of beta-pleated sheets of amino acids, displayed as arrows to indicate the N to C direction, and connecting random coils. The site for attachment to receptor binding lies on the top of S1.

Cryo electron microscopy (Song et al. 2018) has shown that each S1 monomer of SARS is wedge shaped, subtending an angle of 120 degrees on the surface of the trimer, with contacts to the other S1 monomers on each side. The three monomers of S1 together form a cap over the S2 complex below them.

S2 is shown here in its prefusion conformation. The longer helix from Figure 1 is fragmented here into two helices with a connecting bridge. The shorter helix is not yet of final length, but consists of its shorter alpha helical core.

A cluster of black balls are depicted to the left side of S2, near the junction with S1. They indicate a group of S residues calculated to be likely sites of O-glycosylation, that would tend to sugar-coat and protect the region around what we shall see is the fusion peptide motif of the S2 protein.

This modelling is possible because, overall, the S1/S2 protein of SARS is 75% identical to that of nCoV2019. This breaks down to 67% identical, and total 71% highly similar, for S1; 90% identical, and 96% highly similar for S2. This imparts to the model high confidence for the S1 structure, and extremely high confidence to the S2 structure.

Indeed, when one looks at the x-ray structure of SARS S2, virtually everything one sees is identical in the highly probable S2 structure, even though the latter is yet to be determined. Seeing one is seeing the other. As we shall see, this is an enormous advantage in analyzing nCoV2019 for the structural and functional landmarks of the protein and potential targets for inhibition of the viral fusion/entry glycoprotein complex.

While a labeled bead, two dimensional, model is still useful for more easily visualizing the position of a given amino acid peptide sequence in the structure, for S2 of nCov2019 a new model would be superfluous, given the virtual identity of the S2 for both viruses. With updating for new information gleaned over the last 17 years. the old model is the new model. For the post-fusion configuration, Figure 1 is still a good approximation of nCov2019 S2.

4. S1 Overall Similarity

The S1 protein is a globular surface protein that binds SARS to its receptor. (Li et al. 2003; Mathewson et al. 2008). An alignment of the protein sequences of S1 glycoprotein for the Wuhan nCoV2019 and SARS is shown in Figure 5.

Figure 5.

S1 Glycoprotein (Attachment)

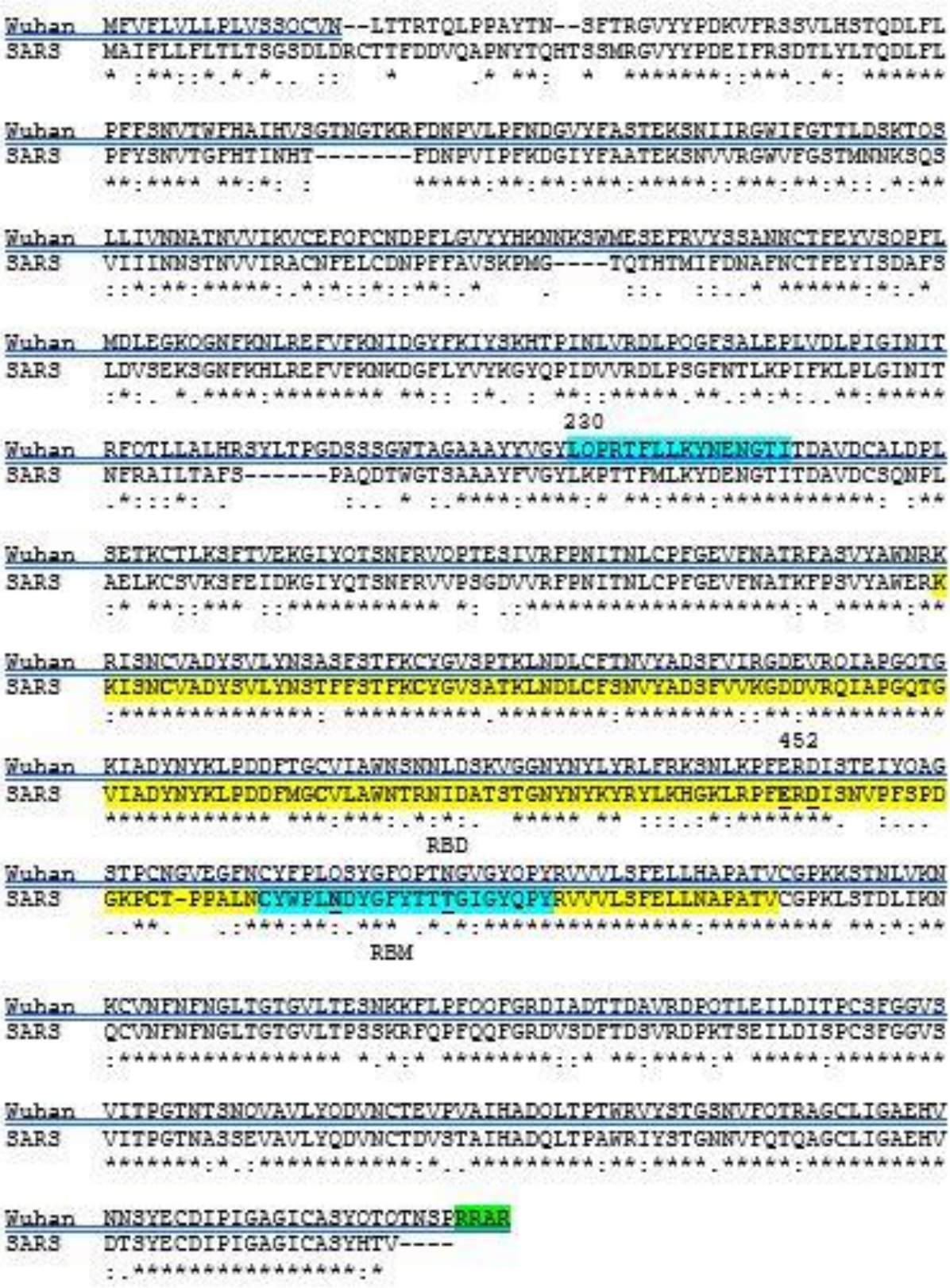


Figure 5: Protein sequence alignment of Wuhan and SARS S1 proteins. Proteins sequences were obtained from Genbank entries MN908947 for Wuhan nCov2019, and the SARS reference standard NC004178. Sequences were aligned using CLUSTAL W. Asterisks indicate identical amino acids at that position, two dots high similarity, and one dot modest similarity. No symbol indicates a non-conservative amino acid substitution, and a dash in the sequence indicates an inferred gap to best align the two sequences.

It can be seen that, while the overall identity of the two proteins in 67%, and high similarity 71%, this is not at all uniform over the length of S1. The closer one gets to the C terminal end of S1, the higher the identity between Wuhan and SARS S1. The closer to the N terminal end, the greater the breakdown in identity between the two. However, as indicated in Figure 4, the breakdown in identity does not undo the propensity of the N terminal region of S1 to adopt a similar secondary structure enriched in beta pleated sheets.

The alignment infers two significant gaps in the SARS sequence relative to that of Wuhan, of 7 and 6 amino acids, respectively. Close examination of the sequences in Wuhan, GTNGTKR and SYLTPG, show them to have a high overall turn propensity, indicating they are probably extensions in Wuhan to turns between beta sheets in both Wuhan and SARS S1.

The S1 proteins end unevenly, which is probably a function of different patterns of endoproteolytic cleavage between the two virus S1 proteins, as will be discussed below.

Binding Domain

Highlighted on the sequence of SARS is the known Receptor Binding Domain (RBD) and, within that region, the known Receptor Binding Motif

(RBM)where SARS S1 actually contacts its receptor, angiotensin converting enzyme 2 (ACE2) on the surface of susceptible cells (He et al. 2004). Four amino acids within the RBD and RBM, known to affect receptor binding, are underlined, E452, D454, N479 and T487. In the Wuhan sequence, only two of the four are conserved. Also, the sequence within and N terminal to the RBM are not well conserved. Such sequence variation could create problems in ACE2 binding to the RBM of Wuhan, despite an overall similarity in structure.

However, other laboratories have reported that they have confirmed that ACE2 is also the receptor for Wuhan nCoV2019, despite the sequence differences shown here. They have opined, however, that the affinity of Wuhan S1 to the receptor may be reduced relative to that of SARS. If so, then this may be a factor in the apparently lower virulence of Wuhan nCoV2019 relative to the high virulence of SARS in humans who use ACE2 to bind the virus.

The difference in sequence within and around the RBD and RBM also has possible significance with regard to highly neutralizing monoclonal antibodies that bind to the region of S1 of SARS at or around the binding site. Only actual experimentation can resolve that question, but these sorts of sequence differences far more than usually lead to interference with the close apposition typical of high affinity binding of neutralizing antibodies. In other words, anti-SARS antibodies might or might not work on nCoV2019.

Possible Immunosuppressive Peptide in S1 (ISP)

Close re-examination of the Wuhan nCoV2019 sequence reveals a feature not seen anywhere in the SARS S1/S2 sequence but common to a number of other Class I Viral Fusion/Entry proteins, namely, a potential immunosuppressive domain (Cianciolo et al. 1985; Morozov et al. 2012). We mentioned this feature earlier, in our discussion of Syncytin-1 and Syncytin-2 expressed during pregnancy from the human genome. An alignment of this region of S1 with several known immunosuppressive domains, is shown in Figure 6.

Figure 6.

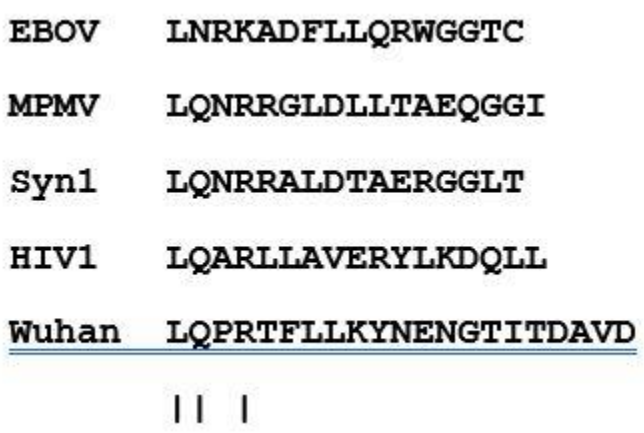


Figure 6: Alignment of the Wuhan S1 sequence with similarity to known immunosuppressive domains of Class I Fusion/Entry Proteins of different virus families. EBOV76, Ebola Mayinga 1976; MPMV, Mason-Pfizer monkey virus; Syn 1, Syncytin-1 from the HERV-W endogenous retroviral sequence on chromosome 7 of human genome; HIV-1, human immunodeficiency virus, BH10; Wuhan, nCoV2019. Vertical lines indicate the known key residues in inducing immunosuppression.

This is the first description of a possible immunosuppressive domain in Coronaviruses or nCov2019. The three key residues common to the known immunosuppressive domains are also in common with the sequence from S1. In addition, other sequence motifs seen in the known

immunosuppressive domains are found in Wuhan, even if not precisely aligned, i.e. FLL, GT, and RY vs KY.

While Coronaviruses are not known for general immunosuppression of the style shown by HIV-1, this does not rule out immunosuppression at the site of active infection in the lung, which would prolong and potentially worsen infection at that site. Work with HIV-1 peptides shows that it is relatively straightforward to test for the induction of apoptosis by immunosuppressive peptides in vitro, that correlate well with effects in vivo. It would be well to not include an immunosuppressive peptide in any vaccine candidate for Wuhan nCoV2019. In this respect, the work with HIV-1 is instructive on which types of amino acid changes to introduce in order to abrogate any immunosuppressive effect.

S1/S2 Cleavage site

As shown above, the alignment of S1 from Wuhan nCoV2019 and SARS is uneven at the C-terminal end. In SARS it is well known that a typical furin cleavage sequence K/RxxK/R is not found at the typical S1/S2 boundary for other Coronaviruses. Instead, SARS uses host cathepsin to cleave S1 from S2 a few amino acids into the classic S2 sequence (Belouzard, Chu and Whittaker 2009),. In SARS there is then a secondary minimal furin susceptible site, RNTR, further into the classic S2 sequence, just prior to the fusion peptide motif in S2.

In Wuhan nCoV2019, there is a strong furin susceptible site at the typical S1/S2 junction, RRAR, which would be more than sufficient as a

cleavage site. There is no second site that aligns with RNTR, suggesting that the one potential cleavage site suffices in the case of nCoV2019.

A different pattern of S1/S2 cleavage is a biologically significant difference between the two viruses.

Inhibitors and Therapeutic Agents

While ACE2 is the receptor for both SARS and nCoV2019, the active site of the enzyme is not part of the binding site of S1 to the ACE2 molecule. So using ACE2 inhibitors available to control blood pressure would have no effect on binding of Coronavirus S1.

Monoclonal antibodies provide high specificity in reacting with discreet sites on viral proteins, and have been developed for many viruses (Elshabrawy et al. 2012). Several panels of monoclonal antibodies with excellent neutralizing activity have been developed against the RBM region of SARS S1. However, the changes in sequence shown in that region of nCoV2019 may make their use problematical. The only way to know is to try, and it is presumed that this is ongoing.

Cocktails of monoclonal antibodies have never been used in humans who are subsequently exposed to SARS. Safety trials have not been conducted, in the wake of SARS disappearing from the human population over 15 years ago. One thing is probable, that a cocktail of antibodies differing in specificity would be needed. Treatment with just one monoclonal has been shown to rapidly result in selecting viral mutations that evade an antibody with a single specificity.

Further development of a SARS vaccine, of any of several constructions, has also been put largely on hold. Several vaccines tested in mice did result in the development of neutralizing antibodies. However, on challenge of the mice with a SARS construct, the mice suffered histopathological changes suggesting that they had been sensitized to components of the virus such that they developed severe allergies to SARS proteins. This experience demonstrates that production of a vaccine is more art than science. The science can be perfect, yet fail to yield a useable vaccine. Over the last 70 years, such failures have often happened. Most vaccines carry with them the possibility of side effects that must be considered in ultimately deploying them to the human population.

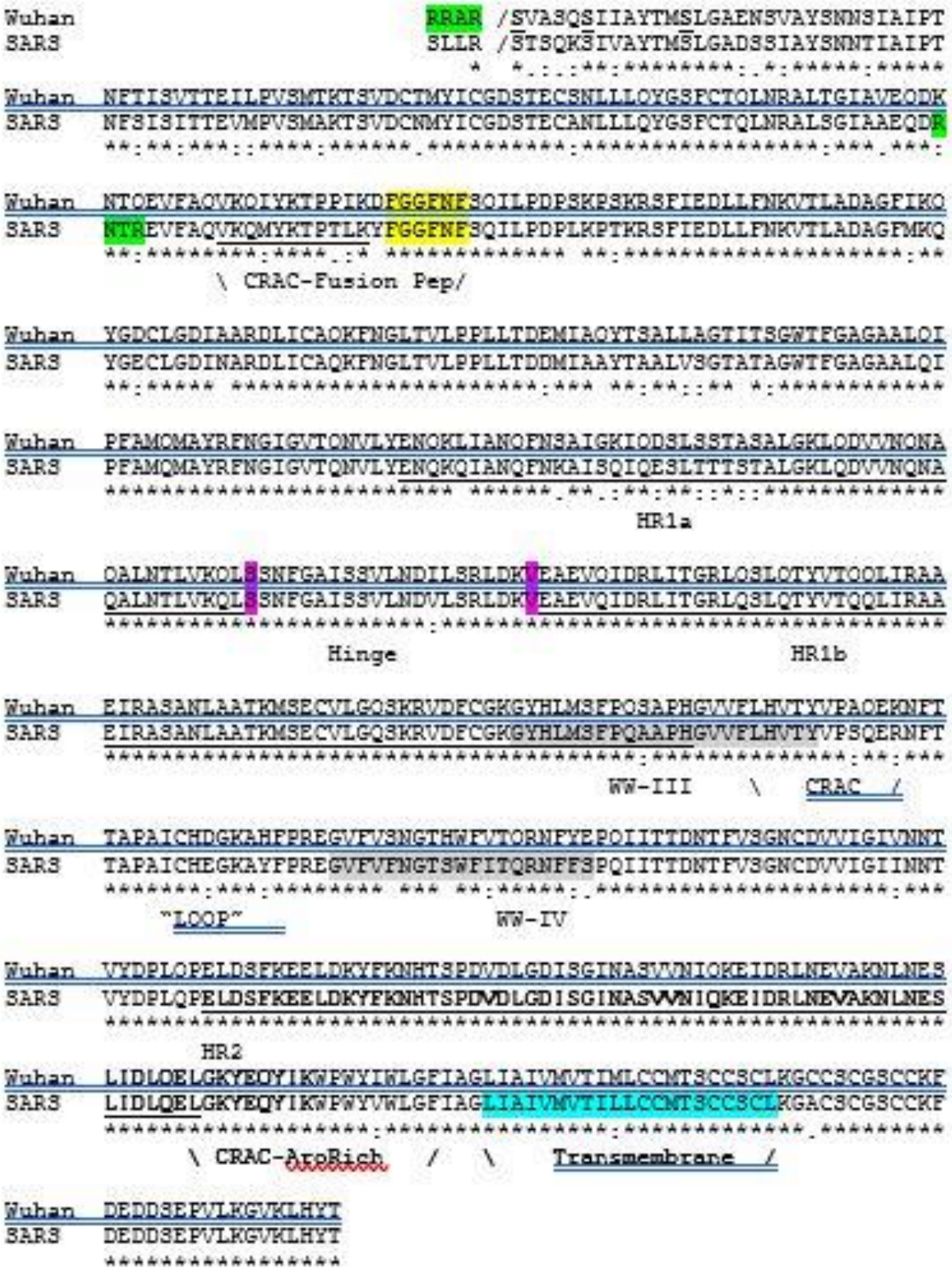
There is no licensed SARS vaccine approved for human use at this time. Nevertheless, the goal has been announced to develop an anti-nCoV2019 vaccine by May of 2020, based largely on developing neutralizing antibodies to S1. The scientific community and the public should understand that, while a laudable goal. this is far from a sure thing.

5. S2 Overall Similarity

As stated earlier, the sequence of S2 for SARS and nCoV2019 are practically identical. However, there are still observations to discuss that have not been previously identified. The annotated alignment of the two viral proteins is shown in Fig 7.

Figure 7.

S2 Glycoprotein (Fusion/Entry/Transmembrane)



Structure

Both the S2 of Wuhan nCoV2019 and that of SARS have all of the essential elements of a Class I Viral Fusion/Entry Glycoprotein, along with lots of extra protein sequence in between the identifiable elements.

It must be emphasized that S2 is a highly conformationally variable protein, while still maintaining a substantial helical, and thereby fibrous overall structure. Radical rearrangements occur after cleavage between S1 and S2, on binding to receptor, and after changes to the environment of the protein involved in the interactions that lead to membrane fusion. Indeed, S2 has at least three different structures, rather than one.

It is rather like a “Transformer” toy, that begins by looking like a boxcar, but unfolds to form a robot that reaches up to grab a target membrane, and then pulls it down toward the membrane at its own feet. It winds up completely bent at the waist, fingers touching toes, messing with the two membranes to fuse them together. The pore formed in this way allows the insides of the virus, including the RNA genome, to get inside of the cell and begin the process of viral replication (Carr and Kim 1993). Inhibit the “Transformer” and one blocks viral infection in its tracks.

We know these structures from studies of crystals of viral protein. In real life, the protein is not constrained by being in a crystal. It is rather a highly deformable protein, that may undergo reversible changes as well as the irreversible changes that occur during the fusion process as it snaps open and then closed while bound to membranes and closely interacting with cholesterol within those membranes.

After the classic furin cleavage site in Wuhan nCoV2019, there are 93 amino acids before the first recognizable sequence motif from Class I fusion proteins. With SARS most of this is simply deleted by a second cleavage at the sequence RNTR that is lacking in Wuhan. With Wuhan some may be lost to other proteases, or be used in inter-subunit interactions during the conformational changes the Wuhan S2 undergoes.

Within this region is the triad cluster of likely O-glycosylated Serines medtioned above, two conserved sequence motifs for N-glycosylation, and four Cysteines that probably participate in cross-linking of the protein chain.

This is followed by a recognizable fusion peptide (FP) region that contains the canonical FGGF sequence identical in SARS and Wuhan. (The "CRAC" label we will deal with below). There is then another relatively amorphous region of nearly 100 amino acids that includes two conserved cysteines for possible disulfide bridging.

The first heptad repeat (HR1a) region then begins and continues for 49 nucleotides or 14 alpha helical turns. This is followed by a conserved hinge region of 21 amino acids bordered by a conserved S at the N-terminal end and a conserved V at the other. This will eventually be incorporated into the extended HR1 helix in the post-fusion confirmation. This is followed by HR1b, the second component of HR1, a peptide region of another 72 amino acids, essentially another 10.5 alpha helical turns.

In all, the ultimate HR1 will extend 142 amino acids and 21 alpha helical turns, extending over 100 Angstroms from the viral membrane surface to the cell membrane surface.

HR1b is followed by a “loop” region, which in Coronaviruses is greatly extended, unlike other Class I fusion proteins with the exception of the ancestral Spumaretroviruses. The Loop region is 83 amino acids long, with two conserved cysteines and three conserved motifs for N-linked glycosylation.

This is followed by the HR2 alpha helix, that begins with the ELDKY peptide homologous to an HR2 peptide of HIV-1, and extends for 60 amino acids, or 8.5 helical turns. The ELDKY motif is potentially important because in HIV-1 it defines a broadly neutralizing epitope for monoclonal Cluster II antibodies such as 2F5 developed by the Robinson lab decades ago. This strongly suggests that homologues of the HIV-1 peptide region could induce broadly neutralizing antibodies to a conserved protein sequence.

Finally there is an aromatic-rich region we first identified in 1989 as common to the Class I superfamily, and then the transmembrane segment that traverses and anchors S1/S2 to the viral envelope. In both viruses, as in other coronaviruses, there is then a short peptide region, rich in cysteines, that is internal to the viral membrane.

Fusion Peptide

The specific peptide motif that interacts with host membranes is known as the fusion peptide (Richardson et al. 1980; Nieva and Aitziber 2003). There are three candidate fusion peptide sequences in S2, one of them favored by others (Yuan et al. 2017),but the most hydrophobic, with the highest interfacial hydrophobicity, is the FGGF motif, conserved in the S2 alignment of Figure 7 (Sainz et al. 2005). In SARS, this peptide winds up

being very close to the new amino terminus of the protein generated by furin cleavage at the RNTR motif, comparable to its position in HIV-1 relative to the amino terminus of gp41. In Wuhan nCoV2019, this does not appear to be the case. Rather, the fusion peptide is internal, far from the putative amino terminus resulting from furin cleavage at the RRAR motif. This also is a feature of other Class I viral fusion proteins, often occurring in a disulfide stabilized loop.

HR1a and HR1b

The HR1 region of S2 in Coronaviruses is extraordinarily long in all of the viruses thus far examined. While fragmentation of the HR1 domain into more than one helical segment is found in the native state of Class I Fusion/Entry Proteins from other virus families such as the Paramxoviruses (Morrison 2003), the segments there are much shorter. In both SARS and nCoV2019, the prefusion conformation contains three full sized alpha helices. The N terminal HR1a is oriented toward the interior of S2 trimer, while the HR1b forms a central helix in each of the three S2 proteins of the trimer. The visible stalk of spikes on SARS virus particles is therefore a complex of nine helices of substantial length in the prefusion form, with the three HR1a helices forming the core and internal inter-trimer interactions.

HR2

HR2 lies on the exterior of what becomes a six-helix bundle in the course of inducing membrane fusion (Wild et al. 1994:Chen 1994). Unlike the radical conformational changes typical of HR1, it remains more a constant feature of the structure during morphological evolution of the

protein complex. In the exterior position it is typically more exposed to antiviral antibodies, along with the loop region. Most monoclonal antibodies that react with S2 do so along the HR2 helix. The 2F5 monoclonal that reacts with ELDKWAS on HIV-1 gp41 is an example now decades old (Montefiori et al. 1988)). Because this region of the Class I Fusion/Entry proteins is often the most conserved region for a given virus, antibodies directed to the HR2 helix tend to be broadly reactive to multiple strains of a virus. In the case of SARS and nCoV2019, it would not be surprise if monoclonal antibodies generated to this region of SARS S2 would be equally reactive to nCoV2019.

Attempts to use peptides to inhibit viruses go back to 1968, when researchers at Parke-Davis made a series of random tripepides and found that Phe-Phe-Gly inhibited measles virus (Miller et al 1968;Nocolaides et al. 1968). This random finding was confirmed by Norrby (1971). The Choppin lab later realized that FFG corresponded to the N-terminus of the measles F glycoprotein and that the inhibition was specific (Richardson et al. 1980; Richardson and Choppin 1983). This led to the formulation of the fusion peptide hypothesis. However, the highly hydrophobic peptide was difficult to manage and not clinically useful, even while providing a proof of concept.

Inhibition of glycosylation has also been demonstrated to have an antiviral effect (Gallaher et al 1973), but inhibition of an essential cellular process makes such inhibition impracticle.

Both HR1 and HR2 have also been targets of peptide analogues that have inhibitory activity against the conformational changes involved in fusion. The basic principle of such experiments is “two’s company, three’s a

crowd", that I first articulated in 1989. Throw an extra helical peptide in the mix, and HR1 or HR2 pair with the wrong peptide. The experiment was performed at Tulane, Duke and the New York Blood Institute using peptide analogues of various lengths. All were successful to different degrees in inhibiting HIV-1 infection. The best of them, targeting HR1 and HR2, were direct peptide analogues of the antiparallel helices from the Gallaher model, and developed at Duke (Qureshi et al. 1990; Wild et al.1992; Jiang et al. 1993) The better of those were targeted to HR2, the better exposed. An extended peptide inhibitor of HIV-1, closely corresponding to the HR2 of HIV-1 designated the "charged helix" in the Gallaher model, was developed into the fusion inhibiting antiviral known as Fuzeon, now produced by Merck. Since it targets an essential and highly conserved helix of HIV-1, it has been used for salvage therapy of HIV patient resistant to other HIV drugs for over two decades.

The example of Fuzeon opened the floodgates to develop similar products derived from the HR1 and HR2 helices of JIV and various other viruses. (Lambert et al 1996; Eckert and Kim 2001; Medinas et al 2002; Pinon et al. 2003 Giannechini et al. 2003)

A Fuzeon-like inhibitor for SARS or nCoV2019, or a peptide analogue of either HR1a or HR1b, is a potentially valuable approach to one component of an antiviral therapy (Sainz et al. 2006: Xia et al. 2018) Three candidate peptides have been developed in SARS. One closely corresponds to the conserved HR2 domain, and is 68 amino acids long, i.e. ELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLN ESLIDLQEL. As extensive as it is, it is only effective at about 20 uM concentration in vitro. While a specific inhibition, given the size of the

peptide nearly twice as long as Fuzeon, that does not make it an attractive candidate for clinical development.

The Garry and Wimley labs also assessed peptides derived from the SARS sequence, specifically two named WW-III and WW-IV that are analogues of the loop region between HR1b and HR2 (Sainz et al. 2006). The peptides GYHLMSFPQAAPHGVVFLHVTY and GVFVFNGTSWFITQRNFFS also inhibited in the low uM range, but with higher specific activity than the longer peptide, precisely because they are much shorter. The peptide regions in nCoV2019 corresponding to these latter two inhibitors are virtually identical, so the corresponding nCoV2019 peptides may hold some promise as inhibitors of nCoV2019.

Jiang, one of the individuals originally in the hunt for such inhibitors of HIV-1 nearly 20 years ago, has developed peptide analogues for the human coronavirus 229E (Xia et al. 2018). However, these differ in sequence from the more divergent SARS and nCoV2019. They likewise are effective in the low uM range.

Fuzeon has been clinically successful against HIV-1 because it is effective in the nM range, a 100 or 1000 fold higher specific activity than the above described SARS peptides.

For treatment of HIV-1, peptide analogues have to be injected under the skin, an impediment to wide usage. However, for a respiratory infection such as nCoV2019, they may well be more effective when administered by nebulization and inhalation.

In the case of any peptide inhibitor, careful consideration must be given to whether the peptides can sensitize a subject as an allergen. Allergic

reactions to vaccines or drugs can be major detriments to their widespread use. While drug resistance has not been a major issue with Fuzeon, there are cases where only a single amino acid mutation will confer resistance (Rimsky et al. 1998). Also importantly, viruses may use different pathways for entry, and peptide inhibitors of HR1 or HR2 may not block and alternate endosomal pathway (Ujike et al. 2008)

Further development of fusion inhibitors continues along these lines. (Gait et al 1995; Root et al. 2001;Sia et al. 2002

CRAC motifs

Figure 7 is annotated in three places with the acronym “CRAC”. CRAC stands for Cholesterol Recognition Amino acid Consensus sequence. It is a peptide motif discovered by Li and Papadopoulos (1998) in cellular proteins. It seems like a quite broad definition of a consensus sequence, but it is remarkably rare in most proteins. The motif begins with a V or L, then 1-5 other amino acids, then a Y, another 1-5 amino acids, then K or R. In short, V/L(X1-5)Y(X1-5)K/R. The aromatic amino acid Tyrosine (Y) is at the center and is thought to be the residue that actually interacts with the equally planar 6-membered ring of cholesterol in target membranes.

Richard Epand and co-workers (Epand 2003; Vishwanathan et al. 2008) found a CRAC motif a short distance from the ELDKWAS sequence of HIV-1, specifically LWYIK, and examined the effect of changes in the peptide, alone or in HIV-1 gp41, in the interaction with cholesterol and induction of virus-induced cell fusion. They found a strong correlation, and abrogation of fusion if the L, Y or K were substituted with even a

conservative change in amino acid. The CRAC motif appeared to have a strong influence on the function of the entire Class I fusion complex.

In 2009, Corver et al noted the presence of a CRAC motif in a location of SARS H2 close to the aromatic region, but investigation of the motif was not included in that analysis, which was focused on the aromatic region itself (Corver et al 2009). This article is the first description of the discovery of CRAC motifs in several additional critical locations within the S2 glycoprotein of SARS and Wuhan nCoV2019. In each case the CRAC motif is conserved in both viruses.

The first CRAC motif, VKQMYKTPTLK, appears immediately before the fusion peptide of the virus. This is a strategic location for cholesterol recognition, just prior to a functional motif containing multiple aromatic amino acids.

The second location is in the loop region of S2, namely LHVTYVPSQER, close to the end of HR1b. A third location is within HR2, namely LDKYFK, in the interesting position of being an extension of the ELDKWY peptide mentioned earlier. As in two sequence analogues of conserved HIV-1 peptides in the HR2 region in one place.

The final location of a CRAC sequence, the one alluded to previously (Corver et al 2009), is just prior to the aromatic rich region, labeled "Aroma" on Figure 1, another critical location for cholesterol recognition in the vicinity of a highly unusual cluster of Tryptophan with a high propensity to stack with the planar surface of cholesterol.

There are no experimental data to indicate that any of these four CRAC motifs confer critical functions on S2 of either SARS or nCoV2019.

However, three of the four are in notably strategic locations where cholesterol recognition would be a valuable property for the action of the adjacent functional peptide regions in membrane fusion.

Aromatic rich Region

Close to membrane insertion of S2 lies the extraordinarily hydrophobic aromatic rich region, i.e. YEQYIKWPWYVWLGF, that I first described as “a certain aroma at the feet of TM” in 1989 for HIV-1 and other Retroviruses. Such a feature is a common, even standard, feature of Class I Fusion/Entry Proteins from many virus families. In each virus where it has been investigated, this peptide region, especially the tryptophans (W) has been found to be essential for the induction of membrane fusion. Within Coronaviruses the sequence WPWYVW is extraordinarily highly conserved even among divergent viruses (Corver et al.2009).

Tryptophan is normally a very rare amino acid. Many large proteins contain none at all. So this peptide region is highly unusual. One detriment of W is that it has but a single three letter nucleotide codon UGG that encodes it. If either guanidine mutates to an adenosine, the codon is mutated to a stop codon that terminates protein translation dead. G to A mutations are common in RNA viruses during transcription, so it would be anticipated that multiple UGG codons in a row would result in frequently generating RNA that cannot code for complete S1/S2 protein. So, to have multiple Ws in a row here the virus may pay a price. This underscores the great importance of this aromatic peptide region for SARS, nCoV2019 and other Coronaviruses.

They are perhaps best understood as members of a general class of membrane destabilizing peptides (Suarwz et al. 2000; Sainz et al.2005)

General Class of Cytotoxic Basic/Aromatic Peptides

One consequence of the high fusogenic activity of the aromatic rich region is that it carries a high degree of potential toxicity that is not appreciated. We recently investigated the Ebola Delta peptide made from an alternate gene to it surface glycoprotein (Gallaher and Garry 2015; He et al. 2017). We found that it combined basic amino acids with aromatic amino acids in such a way as to create a potent toxin that in the virus serves as a membrane-permeabilization region known as a viroporin. A 22 amino acid fragment of Delta peptide, generated by interaction with normal human serum, was indeed more toxic than cholera toxin.

More recently, Gallaher and Gallaher found that Picornaviruses such as D68, which is linked to acute febrile myelitis, encode a similar 27 amino acid peptide that combines basic amino acids with aromatic amino acids, and also has toxic properties.

Examination of the HIV-1 aromatic region, and the helical region of influenza PBf2 known to be toxic, again show that they also combine basic amino acids with aromatic amino acids, as also does the entry peptide for papillomaviruses.

The peptide region around the aromatic region of SARS and nCoV2019 follow this same pattern of combining basic amino acids, K or R with aromatic amino acids, especially tryptophan (W).

We suggest here that a pattern is emerging among very disparate viral agents of using this known membrane-destabilizing motif – K or R with multiple aromatic residues, especially W, to disrupt cell membranes yielding cell fusion, cell permeabilization or cell destruction.

Drawings of several of these peptides from divergent sources are illustrated together in Figure 8.

Figure 8.

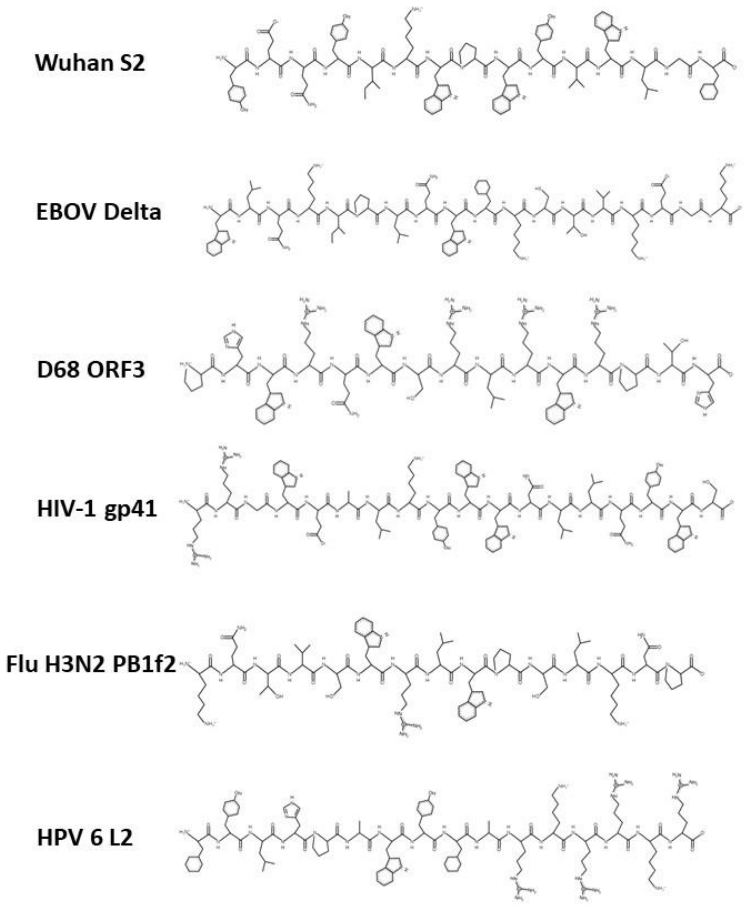


Figure 8: Family of Basic/Aromatic Membrane Destabilizing Peptides. Six peptides known to destabilize membranes or are involved in membrane fusion are shown that share the property of

combining basic amino acids K and/or R with aromatic amino acids W, F or Y. Shown are peptides from Wuhan S2, Ebola Delta, Enterovirus D68 ORF3, HIV-1 gp41, Flu H3N2 PB1f2, and Human Papillomavirus type 6 L2, as described in the text. Projections are from the PepDraw uility of the Wimley lab, Tulane School of Medicine.

While not directly alignable, the common character of these peptides in visually striking. We suggest that they constitute a superfamily of membrane destabilizing peptides across many families of both naked and enveloped viruses, wherever membrane perturbation is an essential mechanism to viral replication or pathogenesis.

Wobble Mutagenesis as Index of Relatedness

Since the HR2 of nCoV2019 and SARS are virtually identical, one may be tempted to regard that region as having a different evolutionary history than less identical regions of S1 and S2. While the identity in amino acids is indeed 98.5%, the underlying RNA code for HR2 is only 81.7% identical. The difference is what is known as “Wobble mutagenesis”.

The genetic code is redundant for many of the amino acids. Alanine may be encoded by the nucleotides GCA, GCG, GCC or GCU, with only the first two nucleotides constant. The third nucleotide is known as the “wobble base”, because it can be any of the four possible nucleotides.

Even when selection preserves the non-wobble first two nucleotides from mutation, mutation may occur freely in the third without changing the sequence of the amino acids in the protein. “Wobble mutagenesis” accumulates over a long period of time. In the 100 years Influenza H1N1 has circulated in the human beings since the 1918 pandemic, HA2 has remained constant in amino acid sequence. However, 7% of the codons in the viral RNA coding for that constant sequence have undergone wobble mutagenesis.

Figure 9 shows an alignment of the RNA encoding HR2 within the conserved region between nCoV2019 and SARS.

Figure 9.

Over this region encoding 71 amino acids, 70, or 98.5% are identical. However, there are 13 mutations in the RNA between the two, for an RNA identity of only 82.7%. The 2/1 pattern of nucleotide identity is an obvious sign of wobble mutagenesis, since only every third nucleotide is mutated. Even when the beginning and end of the gene is unknown, the reading frame of three letters can be directly inferred from the sequence.

Comparing the extent of wobble mutagenesis gives an estimate of the time since the Wuhan and SARS RNA were identical in this region. Using flu as a yardstick, the estimate here would be 260 years ago. So, even though the HR2 of nCoV2019 and SARS are nearly identical, the RNA tells us that they actually diverged from a common ancestor approximately 260 years ago.

Inhibitors and Therapeutic Agents

Overall, the spike proteins make attractive candidates for development of antiviral strategies (Du et al. 2009). These include monoclonal antibodies, and peptide inhibitors that are analogues of regions HR1 and HR2 critical for virus entry.

Peptide analogues of the aromatic rich region are so likely to be hydrophobic themselves that developing an inhibitor of these region is very unlikely. However, recognition of the potential toxicity of this region is

ample reason to exclude it from reagents intended for clinical use wherever possible. They are indispensable for production of virus for vaccine use, but the resulting viral protein should be purged of this area as much as possible before administration to avoid toxic side effects.

This has not been done for either the influenza or Ebola vaccines. Fortunately, this region of influenza HA is predicted to have low toxicity due to a low aromatic content. Some of the reactogenicity of the Ebola vaccine could be attributable to retaining these peptides in the vaccine responsible for membrane fusion.

Vaccine production is by its very nature an empirical process, much more art than science, with trade-offs between efficacy and side effects. However, the presence of essentially reactogenic peptides in potential vaccines and antivirals should be given greater attention.

6. Cysteine Protease 3CLpro

Overall Similarity

The main protease encoded by Coronaviruses is a cysteine protease, 3CLpro, which cleaves at 11 recognition sites in the viral polyprotein as it is being translated, producing the non-structural proteins (nsp) involved in viral replication (Muramatsu et al. 2016). As such, its function is absolutely essential to replicate virus and it has been a prime target to develop relatively small molecule inhibitors that can serve as antiviral drugs against SARS in the event of its return.

The drive to develop such inhibitors is inspired by the tremendous success of protease inhibitors in treatment of HIV-1 infection.

Regions of the polyprotein vary in their conservation between SARS and nCoV2019. However, in the region of 3CLpro, the Wuhan nCoV2019 and SARS proteins are 96% identical. An alignment of the proteins is shown in Figure 10.

Figure 10.

The critical amino acids for protease function are highlighted in the alignment. It can be seen that all of these determinants of protease activity are identical in both viruses. Of the 11 amino acid differences between the two viral proteins, 9 are conservative substitutions. This makes it highly probable that the large library of protease inhibitors developed against SARS 3CLpro will also be active against the 3CLpro of nCoV2019.

Inhibitors and Therapeutic Agents

There is an enormous variety of candidate protease inhibitors that have been developed against the SARS 3CLpro protein (Yang et al. 2005; Berry et al. 2015). High throughput systems have been used to screen for potential candidates, with designer stereochemistry guided by accurate x-ray crystallography of the viral 3CLpro dimer. These are followed by assessments of toxicity and bioavailability. These have been arrested in development because of the continued disappearance of SARS virus from human beings. As a result, none of these potential inhibitors are deployable as antiviral agents in the clinical setting. However, it is expected that a number of pharmaceutical companies will jump into this gap as soon as possible.

The aim is to plug two pockets in the enzyme to prevent its natural substrate from being able to bind. The stereochemistry of the pockets is well defined, and designer drugs would ideally plug the pocket stably without being toxic. Drugs that were designed to covalently bind to amino acids in the pockets proved to be to reactogenic and toxic to human processes. So the search has been increasingly to find a drug that binds by affinity, and stably sticks in the enzyme to block it without also being toxic. That magic bullet has thus far appeared to elude those trying to develop it.

It has been reported that existing protease inhibitors licensed against other viruses are being deployed on a compassionate use basis in China, allegedly “with some success”. Based on high throughput screening, after computational matching to a model of the NCoV2019 3CLpro structure, the old protease inhibitor Lopinovir has apparently been selected as the closest

match and is now being deployed in China on a trial basis, according to the pharmaceutical firm Innophore (https://innophore.com/2019-ncov/).

Scientists in China have been heavily involved in the development of 3CLpro inhibitors. If any drug is felt to be close to ready, with high specific activity against the nCoV2019 3CLpro enzyme, I suspect that this will be the lead pharmaceutical deployed once supplies are adequate. Chinese scientists and the National Health Council (NHC) will drive the process of if, when and which inhibitor will be yanked from early development and deployed. A process of bringing a new drug online that normally takes years will take months, and will be entirely different from the process that is primarily market-driven in the US.

7. Papain-like Protease

Overall Similarity

There is a second Coronavirus-encoded protease that cleaves the viral polyprotein at an additional 3 sites. It is similar in its specificity to the digestive enzyme papain, so called because it was originally derived from papaya. While most of the focus on protease inhibitors has been directed at 3CLpro, the papain-like protease, named nsp13, presents an additional advantage as a drug target (Baez-Santos et al. 2015). The protease also interacts with components of the human innate immune response, suppressing that response in several ways that commonly involve the ubiquitin pathway for tagging proteins for destruction. In this way, the protease short circuits the production of interferon, an important component of the natural human inhibition of RNA viruses.

The logic is that by inhibiting nsp13, the human host will be better able to inhibit the virus by its own natural means, especially by interferon converting cells to an antiviral status.

An alignment of the nsp13 proteases of Wuhan nCoV2019 and SARS is shown in Figure 11.

Figure 11.

```
SARS    EVKTIKVFTTVDNTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPS
Wuhan   EVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPHNSHEGKTFYVLPN
        **:********** *****:*************************** .******:***.

                                                    107  112
SARS    DDTLRSEAFEYYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLAL
Wuhan   DDTLRVEAFEYYHTTDPSFLGRYMSALNHTKKWKYPQVNGLTSIKWADNNCYLATALLTL
        ***** ******** * *****************:***.**************::.**:*

SARS    QQLEVKFNAPALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLES
Wuhan   QQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELGDVRETMSYLFQHANLDS
        **:*:***.****:********:***********.**************::*:*****:*

SARS    AKRVLNVVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESS
Wuhan   CKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSYEQFKKGVQIPCTCGKQATKYLVQQESP
        .********* ***: ***.*************:::*.**.***.**::**:*******.

                                      273           287
SARS    FVMMSAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPV
Wuhan   FVMMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSKETLYCIDGALLTKSSEYKGPI
        *********:*:*::*** **.***********.***:***** **** *** ******:

SARS    TDVFYKETSYTTTIK
Wuhan   TDVFYKENSYTTTIK
        *******.*******
```

Overall, nsp13 from the two viruses is only 83% identical. The four amino acids that come together to form the catalytic active site are identical in both, but neighboring amino acids to the site show significant differences that may affect the stereochemistry of the area, and thwart binding by those inhibitors that have been developed for SARS.

Inhibitors and Therapeutic Agents

Another advantage of targeting nsp13 is that the development of inhibitors seems further along and that the inhibitors are of simpler construction. They are based on the multi-ring organic compound naphthalene, with the intent of interacting with the tryptophan in the active site and blocking it from participating in the enzyme activity. A particular candidate, named “inhibitor 3e” or simply “3e” has been shown to have good activity, low toxicity, and good metabolic stability, the hallmarks of a lead drug candidate.

Since “3e” is relatively small and fits in the enzymatic site, the surrounding differences between the nsp13 of SARS and Wuhan nCoV2019 may not be an issue in its high affinity binding. As a small molecule, it would be expected to have high specific activity per milligram administered.

A drawback is that it targets less the actual viral replication and more the salvaging of the normal immediate human interferon response. However, it may be able to be brought online more quickly.

8. Helicase

Viral RNA must unwind from a helical state in order to effectively replicate. With an enormous RNA genome, this is especially an issue with Coronaviruses, and so the viral genome encodes a helicase to facilitate this essential step in replication (Adedeji et al.2012). Inhibition of helicase would directly impair the ability of SARS or nCoV2019 to replicate its RNA genome or generate messenger RNAs for the synthesis of viral proteins. Inhibition would occur at an early step in viral replication, potentially rendering it a more effective mode of suppressing viral infection.

Targeting helicase is also inspired by the great success of the fluoroquinolone antibiotics against bacteria, like ciprofloxacin, by targeting the unwinding of bacterial DNA. It provides a proof of concept, even though the flurroquinolone antibiotics would themselves be useless in combatting the RNA replication by Coronaviruses.

A number of flavonoid derivatives, some of them natural products from plants, have shown success in inhibiting the helicase of SARS.

Identity

An additional advantage in targeting the viral helicase is that the amino acid sequence, all 601, are absolutely identical in SARS and nCoV2019. There is no alignment worth showing here. Work on one nsp13 is immediately translatable to the other. Therefore, all of the work done in identifying inhibitors of SARS helicase has essentially been work done on Wuhan nCoV2019 well before its discovery.

Inhibitors and Therapeutic Agents

Flavonoids such as myricetin and scutellarein are normal components of fruits and vegetables consumed by humans on a daily basis. Such nutrients are available as dietary supplements, and therefore are expected to have very low toxicity to human beings. Specific inhibitors of SARS helicase have been developed with high selectivity without affecting other enzymatic activities such as ATPase.

One such anti-SARS drug was developed in 2012, named SSYA10-001, in collaboration with Susan Weiss' lab at the University of Pennsylvania. It had the desirable properties of a lead drug in further development of this approach. However, as with most anti-SARS drug development, it appears to have stalled in the absence of any cases of SARS over the last 15 years.

However, this class of drugs show promise of being brought along rather quickly to a deployable formulation against mCoV2019. For one thing, as mentioned above, the complete identity between the SARS helicase and the nCoV2019 helicase means that there is zero uncertainty that what works on one would work on the other.

9. Peptide Elongation Inhibitors

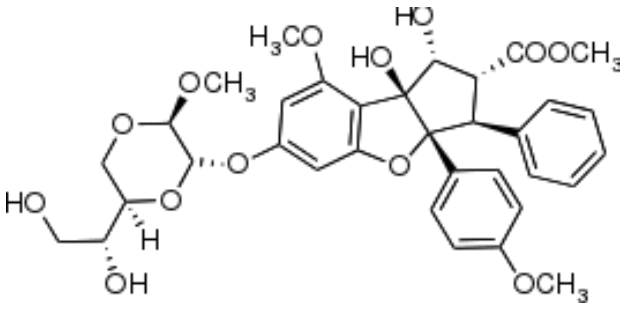
In the earliest days of the fight against AIDS, pharmaceutical companies screened everything they had on their shelves against the virus, to determine if anything already far along in development, or already licensed, had any effectiveness against HIV-1.

The major winner is that screening was a drug previously developed as an anti-cancer nucleoside analogue but had been shelved because of its relative ineffectiveness.

That drug, available off the shelf, so to speak, was azidothymidine, known as AZT. For those living in the 1980s and 1990s, AZT became a household word. Burroughs-Wellcome made an unimaginable fortune selling AZT before other antiretroviral drugs could be brought to bear. It was far from a perfect solution to the problem, but it was a solution. It had issues of toxicity and rather easily encountered mutation of HIV-1 to drug resistance, but it bought time for many.

There is a parallel situation today, in the form of another group of anti-cancer drugs licensed for human use in battling certain types of cancer. These are the drugs that inhibit elongation of growing polypeptide chains, using human elongation factor 4, abbreviated eIF4A. (Muller et al. 2018). One of these inhibitors is Silvestrol, a polycyclic natural product of a tropical member of the mahogany family of plants, Aglaia foveolate, widely available in southeast Asia in large quantity without endangerment of the species by harvesting. It has been found to inhibit a wide variety of RNA viruses without significant toxicity in vitro, including some human Coronaviruses. The structure of Silverstrol is shown in Figure 12.

Figure 12.



Of great advantage is that the elaborate compound need not be synthesized but purified from the leaves and twigs of a profusely growing plant. Another advantage is that it is targeted not against a viral protein but a cellular one that is essential for making viral protein. Thus, the virus cannot be expected to mutate to drug resistance. Likewise, the human host will not mutate to resistance in tissue of the respiratory tract.

The potential usefulness of such an inherently toxic drug to human cellular physiology is that virus replication, like growth of cancerous tissue, requires a much higher level of protein synthesis than resting cells of host tissue. The differential toxicity to virus at low doses of inhibitor lies in its reliance on very high levels of protein synthesis to result in significant viral load. Unlike cancer therapy, it may be possible to administer the drug for a much shorter period of time during the acute phase of viral infection

A disadvantage is that the patient is also ramping up protein synthesis in the immune response to the virus, which may be adversely affected, as well as any inherent toxicity of the drug due to its mechanism of action.

Specific antivirals that have high differential inhibition of viral proteins alone are obviously preferable to resorting to an anti-cancer drug of some likely toxicity to already sick patients. However, if all else fails, it may be a short-range alternative therapy, as AZT was against HIV-1 decades ago.

V. PROSPECTS

Continued Outbreak

As of this writing, the outbreak had entered an exponential phase of expansion. Given the incubation period of the illness, and even a minimal lag in reporting, it will not be until February 1 that any effect due to the lockdown and closures ordered by the Chinese government will be apparent. We will not project numbers that do not yet exist, except to say that the number of cases will be much higher than the 4515 reported as of January 28. In fact, startingly higher.

After the onset of the West African Ebola outbreak in 2014, teams happened to be in place capable of isolating and sequencing the RNA from sequential cases. Very rapid mutation of Ebola was observed, as it adapted to the new human environment. This is likely to occur in the case of nCov2019 as well, especially given the massive viral replication currently underway.

Eighteen days after the original Wuhan nCoV2019 was posted on Genbank, only five others have been posted, greater than 99% identical to the first. This is likely to change with additional postings.

All known cases are being hospitalized and put in isolation from the public, with medical staff wearing full protective gear developed for SARS and Ebola. There is, however, significant risk that medical facilities and staff may become overwhelmed, and soon.

Approximately 20% of the apparent cases of nCoV2019 are being reported as “severe”, a high number. Mortality thus far is hovering around

3%, when weighed against current caseload, but is likely to be higher given the time lag between report of illness and when the patient expires. An ultimate mortality rate relative to the caseload of apparent infection would be projected to be in the range of 5 to 7%, about half or a bit more than half of that ultimately found in the case of SARS. This may worsen if medical facilities are overwhelmed and existing medical staff can no longer provide intensive care for the severely ill.

This does not count the number of inapparent cases that are going unreported, that are unknowable at this time.

Every infection has a unique epidemic profile that is colloquially known as its “epidemiological iceberg”. An iceberg lies 7/8 below the surface of the water, so is much larger than what is seen. The same is true with most viral infections, with the exception of a few. Even in the case of flu, that kills more than 10,000 Americans each year, most patients rely on over the counter cold remedies, even when there are good antivirals available for use early in infection.

For nCoV2019, it is likely that a much higher percentage of infections will be above the surface, but there will always be that which is unseen, yet capable of spreading the virus.

The prospect for an early suppression of this outbreak is not good at this point. Also, since lockdowns cannot be continued long term, containment may be problematical beyond another week. It will be imperative that some successful form of antiviral therapy to be deployed to reduce severe illness and death due to nCoV2019. Fortunately, as described above, many such drugs have been in the works for years, so the world is

not starting from scratch. There may even be an effective drug deployed quite soon from among current lead candidates, or a pre-existing drug with sufficient potency to reduce severe illness or death.

There will certainly be no lack of trying.

Off the Shelf Drugs

The protease inhibitor Lopinavir has been deployed. Other existing antiretroviral drugs may show some antiviral activity as well. It may be presumed that all have already been assessed in screening assays.

Some human monoclonal antibodies against SARS may prove to be useful against nCoV2019, provided a "cocktail" of such antibodies can be assembled and production ramped up quickly. Presumably this is being done.

Flavonoids already used as dietary supplements could be tested in a nebulized form, to see if they were tolerated well and had any effect on the helicase activity of nCoV2019.

Silvestrol is also a licensed human drug that may have some antiviral effect without excessive toxicity to the lung.

In terms of "off the shelf" those are the only such drugs known to exist that have potential in inhibiting nCoV2019 without further development.

Several other anti-SARS drugs are far along in development, in vitro or in small animals, and may be quickly brought on the line after further testing in small trials. Inhibitors of 3CLpro, nsp13, spike S2 protein and helicase fall into that category. Normally, the process of bringing a drug to

market from the current stage of these experimental drugs would take years. We may expect considerable shortcuts to be taken, especially overseas where the many levels of drug approval can be more easily waived in an emergency than in the United States.

The stakes here are incredibly high. Drug manufacturers, no matter how large, will not be able to risk very expensive processes of bringing a drug to market without assurance that someone will be out there to buy their product. We should expect them to demand a hefty government contract, from China or from the US, before incurring great expense to inhibit a virus that, like SARS, may be eliminated by the time a saleable stockpile of the drug is ready for release.

In 1976, before agreeing to manufacture and deploy a vaccine against the swine flu that had appeared in Fort Dix, New Jersey, drug manufacturers demanded, and got, not only government contracts for the vaccine when it was produced, but also complete indemnification against lawsuits by the US Congress. In the rush, mistakes were made, several cases of Guillain-Barre syndrome were linked to the vaccine, and the US government was forced to make significant financial settlements with the victims.

In 2020, the risk of lawsuits and massive judgments are even more prominently in the minds of drug executives than in 1976, given the highly litigious atmosphere prevailing in the US today. Aversion to risk is a far greater force than opportunity for sales, and that is likewise a higher motivator than merely saving lives.

A successful inhibitor against nCoV2019 may well be a multibillion dollar pharmaceutical, if the outbreak continues to expand or even continue at a much lower level. It may also have to be written off at a loss, if the virus disappears as SARS did. One can be sure that pharmaceutical executives and boards are very conscious of these realities.

Combination Therapies

Combination therapies, especially with components drugs that do not have high activity, would be expected to be the norm. Even with relatively short term administration over days, RNA viruses have proven to be quite capable of developing drug resistance. There are examples of viruses developing resistance to a single monoclonal antibody with a single mutation in a single cycle of infection, given the frequency of mutation and heavy pressure of positive selection for resistance.

An appropriate mix would be two or more monoclonal antibodies, a protease inhibitor and a third inhibitor from the other potential classes mentioned above. Each should inhibit the virus approximately equally, to prevent development of drug resistance to any of the others.

Monotherapy against RNA viruses has worked in the short term. An excellent example is Tamiflu against influenza virus. However, combination therapy would be much preferable, especially with the level of mortality due to nCoV2019 exhibited thus far.

Vaccine

The goal of a nCoV2019 vaccine by May, however laudable a goal, will not have an impact on the current outbreak. It will either be massive or contained by then. The aim seems to b toward an engineered component vaccine, rather than attempting to use whole, killed virus. Researchers are building on what has been done in the same direction as that aimed at SARS, and that is a good thing. However, a caveat is that early testing in small animals showed a good bit of allergenicity for the SARS vaccine candidate years ago. This complication will have to be circumvented. It has happened before. In the 1960s, a vaccine was developed against respiratory syncytial virus (ReSV), an important pathogen of infants and the elderly. When vaccines were challenged with the live virus later, the reaction to the vaccine had made the infection worse rather than preventing it. The vaccine had to be withdrawn.

The ReSV vaccine had elicited not just neutralizing antibodies but another type of antibody that was enhancing, by improving the binding of the virus to respiratory tissue. Some classes of antibody are cytophilic – they bind to cell receptors by their back end, while reacting to the virus with their front end. In drawing the virus to the cell surface, they do the opposite of what was intended. Given the extension of the spike protein far from the viral envelope surface, the prospect of a Coronavirus vaccine generating enhancing antibodies or allergenic antibodies is quite real.

A vaccinologist never knows exactly what they are getting into, no matter how good the science on which the vaccine is based. This must be kept in mind.

SARS came onto the scene 17 years ago, and there is still no SARS vaccine that is ready for prime time. Very good scientists have been at work on an HIV vaccine for a good 30 years, and still there is no HIV vaccine.

The Ebola vaccine distributed on a selective basis was rushed into production after the 2014 West African outbreak. The first version was too reactogenic, inducing significant side effects; the dosage had to be reduced to be tolerated. Protection afforded by that vaccine has not been perfect in field conditions. However, we are not letting the perfect be the enemy of the good. Plus, in the Democratic Republic of the Congo, tort lawyers are hard to come by.

After the 1957 flu pandemic and the 1968 flu epidemic, there was no development of a massively distributed flu vaccine. Only since 1977 has there been availability, and due to endless viral mutation, over the last decade the flu vaccine has never been more than 35 to 60% effective in keeping those vaccinated out of the hospital.

There is a lesson in these past histories. Virology was served by brilliant minds during those years, as brilliant or more brilliant than those in the hunt for a vaccine today. Most vaccines today are the product of those minds, and not recent developments. Yet they had checkered success then, and some near disasters at intervention.

That wizard behind the curtain, however brilliant, is making decisions, however well thought out, that may or may not work.

While development of a vaccine is probable over time, it is a future hope, not a certainty.

VI. GENERAL CONCLUSIONS

We are not about to reach for a crystal ball and make reckless predictions based on the unknowable. However, several major points can be made.

1. nCoV2019 is rapidly becoming a human virus, if it has not already done so. The current outbreak is expanding exponentially, with a frequency of approximately 20% severe infections requiring advanced respiratory supportive care. Mortality is at least 3%, and probably higher given the time lag between reported infections and deaths. Thus far, it is not as bad as SARS was 17 years ago, but it is spreading much faster. It is uncertain how effective lockdowns and shutdown of transport have been in curtailing spread.

2. The best defense against nCoV2019 now is common sense. To the greatest degree possible, shield from sneezes and coughs, keep more than 3 to 6 feet away from others as much as possible. Wash hands frequently, and also frequently use hand sanitizer. Avoid crowds when possible. While face masks are protective, too often they give a false sense of security and induce people to take chances they should not. Avoid touching your fingers to your nose and face. Practice these measures even though nCoV2019 is not present in your community, since they also protect you against the known danger of influenza virus already in your neighborhood.

3. Based on previous experience with SARS and Ebola, public health interdiction across international borders has been effective. Imported cases are being isolated and secondary spread in countries other than in China and its close neighbors has been curtailed. However, lockdowns and shutdowns cannot be sustained for long, as the flow of food and goods globally is now an important component of the global economy.

4. Certain proteins of nCoV2019, such as the S2 spike protein, the 3CL protease and the helicase, are virtually identical to those in SARS. Thus, studies done on SARS over the last 17 years are immediately pertinent to our understanding of nCoV2019. Other proteins, while less similar, may be close enough that approaches developed for SARS may prove useful in nCoV2019 in a far shorter time period than if the new Coronavirus were more divergent.

5. Deployment of antiviral drugs will probably be essential to limit severity of illness and death. There are plenty of lead drug candidates of several classes, but none currently licensed or readily deployable without more data. The old antiretroviral protease inhibitor Lopinavir is being deployed, but its impact is not yet publicly known. If successful, the good news will travel fast. We would note it has now been days since first used.

6. A vaccine for nCoV2019 does not exist. Indeed, a vaccine for its predecessor SARS also does not exist, and would not be useful if it did. The key targets of a vaccine appear different enough that

immune reagents developed against SARS are not likely to be effective against nCoV2019. The use of an identical receptor, ACE2, and identification of the binding motif on nCoV2019 make for an excellent start. While a vaccine is promised by May, it will not blunt the impact of the current outbreak. Given the vagaries of vaccine development, deployment of a successful vaccine with low side effects in that time frame is uncertain.

7. Developing and actually licensing broadly inhibiting antiviral drugs and vaccines to every known family of human viruses is in the national and international interest as an essential component to national security.
   When Ebola hit in Africa in 1995, constituting and immediate and present danger to the whole world, we were not ready; 19 years after the virus had been discovered. When it hit again in 2014, another 19 years later, we still were not ready.
   When SARS hit in 2002, we were not ready, and still are not ready; nearly 40 years after human coronaviruses had been discovered.
   When nCoV2019 has hit, 17 years later, we are not ready.
   Emergence happens; we are at great risk of a global pandemic of major proportions, someday, and yet, for all our virological and technological expertise, we are never, ever ready. We need to be ready.

VII. ACKNOWLEDGEMENTS

We are greatly indebted to the Chinese scientific community that made the selfless and highly unusual decision to post the genomic sequence of Wuhan Coronavirus nCoV2019 as soon as it was available. This commentary, in the detail presented, would not have been possible otherwise.

We wish them well, and pray for the safety of them, their families and the Chinese people. As they hunker down in their homes or spend endless hours in their laboratories, let the viral outbreak pass over and away from them when they emerge.

We compliment Innophore for their transparency in publishing their process for screening and selecting possible nCoV2019 inhibitors using massive commitment of computer resources.

We are, as always, extraordinarily grateful our good friend and colleague of 35 years, Robert F. Garry of Tulane. We have produced 14 different publications together since 1989. We thank him for passing us information on nCoV2019, lively interchange, and reviewing the manuscript. However, we take complete responsibility for this article.

We thank Fay Gallaher and Denise Gorton for their support, as they watch us sitting in the glow of our computer screens.

We thank Wesley Terry for his able assistance in preparing the manuscript to match the format requirements of Amazon Kindle.

The National Center for Biotechnology Information, with its suite of databases, search programs and utilities, has been indispensable to

preparation of this article. Likewise, Wikipedia has responded positively to the challenge of updating the public on the profile and statistics of the outbreak. Virological.org has been an important outlet for communication. Without such resources, it would be much more difficult to assemble the information needed to have any confidence in analyzing nCov2019 and the ongoing outbreak.

VIII. BIBLIOGRAPHY

Adedeji AO, Singh K, Calceterra NE, DeDiego ML, Enjuanes L, Weiss S, Sarafianos SG. 2012. Severe Acute Respiratory Syndrome Coronavirus Replication Inhibitor That Interferes with the Nucleic Acid Unwinding of the Viral Helicase. Antimicrob Agents Chemother. 56(9): 4718-4728.

Aras S, Aiyar A, Amedee AM, Gallaher WR. 2009. Molecular character of influenza A/H1N1 2009: Implications for spread and control. Indian J Microbiol. 49(4):339-47.

Baez-Santos YM, St. John SE, Mesecar AD. 2015. The SARS-Coronavirus papain-like protease: Structure, function and inhibition by designed antiviral compounds. Antiviral Res. 115:21–38.

Belouzard, S, Chu VC, Whittaker GR. 2009. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. PNAS. 106(14): 5871-5876.

Berry et al. 2015. Human coronavirus OC43 3CL protease and the potential of ML188 as a broad-spectrum lead compound: Homology modelling and molecular dynamic studies. BMC Structural Biology 15:8

Bousse et al., 1994, Regions on the Hemagglutinin-Neuraminidase Proteins of Human Parainfluenza Virus Type-1 and Sendai Virus Important for Membrane Fusion, Virology 204:506-514.

Bousse et al., 1995. A single amino acid change enhances the fusion promotion activity of human parainfluenza virus type 1 hemagglutinin-neuraminidase glycoprotein, Virology, 209:654-657.

Bratt MA, Gallaher WR. 1969. Preliminary analysis of the requirements for fusion from within and fusion from without by Newcastle disease virus. Proc Natl Acad Sci U.S.A. 64(2):536-43.

Carr CM, Kim PS. 1993. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. Cell. 73:823-32.

Centers for Disease Control and Prevention. 2009. Swine Influenza A (H1N1) Infection in Two Children --- Southern California, March--April 2009. MMWR Morb Mortal Wkly Rep. 58(15):400-2.

Chambers et al., 1990, Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins, J. Gen. Virology 71:3075-3080.

Chen, 1994, Functional role of the zipper motif region of human immunodeficiency virus type 1 transmembrane protein gp41, J. Virology 68:2002-2010.

Cianciolo GJ, Copeland TD, Oroszlan S, Snyderman R. 1985. Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins. Science. 230:453-5.

Collins et al.1984 Nucleotide Sequence of the Gene Encoding the Fusion (F) Glycoprotein of Human Respiratory Syncytial Virus, Proceedings of the National Academy of Sciences, 81:7683-7687.

Corver J, Broer R, van Kasteren P, Spaan. 2009. Mutagenesis of the transmembrane domain of the SARS coronavirus spike glycoprotein: refinement of the requirements for SARS coronavirus cell entry. Virol J 6:230.

de Groot RJ et al. 1987. Evidence for a coiled-coil structure in the spike proteins of coronaviruses. J Mol Biol. 196:963-6.

Du L, He Y, Zhou Y, Liu S, Zheng BJ, Jiang S. 2009. The spike protein of SARS-CoV--a target for vaccine and therapeutic development. Nat Rev Microbiol. 7:226-36.

Eckert DM, Kim PS. 2001. Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. Proc Natl Acad Sci U S A. 98(20):11187-92.

Eckert DM, Kim PS. 2001. Mechanisms of viral membrane fusion and its inhibition. Annu Rev Biochem. 70:777-810.

Eckert DM, et al. 1991. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. Cell. 99:103-15.

Elshabrawy HA, Coughlin MM, Baker SC, Prabhakar BS. 2012. Human Monoclonal Antibodies against Highly Conserved HR1 and HR2 Domains of the SARS -CoV Spike Protein Are More Broadly Neutralizing. Plos One. 7(11), e50366.

Epand RM. 2003. Fusion peptides and the mechanism of viral fusion. Biochimica et Biophysica Acta. 1614:116-121.

Gait et al. 1995. Progress in anti-HIV structure-based drug design, Trends in Biotechnology, 13:430-438.

Gallaher WR, Levitan DB, Blough HA. 1973. Effect of 2-deoxy-D-glucose on cell fusion induced by Newcastle disease and herpes simplex viruses. Virology. 55(1):193-201.

Gallaher WR. 1987. Detection of a fusion peptide sequence in the transmembrane protein of human immunodeficiency virus. Cell. 50(3):327-8.

Gallaher WR, Ball JM, Garry RF, Griffin MC, Montelaro RC. 1989. A general model for the transmembrane proteins of HIV and other retroviruses. AIDS Res Hum Retroviruses. 5(4):431-40.

Gallaher WR, Ball JM, Garry RF, Martin-Amedee AM, Montelaro RC. 1995. A general model for the surface glycoproteins of HIV and other retroviruses. AIDS Res Hum Retroviruses. 11(2):191-202.

Gallaher WR. 1996. Similar structural models of the transmembrane proteins of Ebola and avian sarcoma viruses. Cell. 85(4):477-8.

Gallaher WR, DiSimone C, Buchmeier MJ. 2001. The viral transmembrane superfamily: possible divergence of Arenavirus and Filovirus glycoproteins from a common RNA virus ancestor. BMC Microbiol. 1:1.

Gallaher WR. 2009. Towards a sane and rational approach to management of Influenza H1N1. Virol J. 6:51. https://virologyj.biomedcentral.com/articles/10.1186/1743-422X-6-51

Gallaher WR, Garry RF. 2015. Modeling of the Ebola virus delta peptide reveals a potential lytic sequence motif. Viruses. 7(1):285-305.

Gething MJ, Bye J, Skehel J, Waterfield M. 1980. Cloning and DNA sequence of double-stranded copies of hemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza to assure that this and future threats of pandemic influenza can be met. Nature. 287:301-306.

Giannecchini et al., 2003, Antiviral Activity and Conformational Features of an Octapeptide Derived from the Membrane-Proximal Ectodomain of the Feline Immunodeficiency Virus Transmembrane Ectodomain, J. Virol. 77:3724-3733.

Goba A, Khan SH, Fonnie M, Fullah M, Moigboi A, Kovoma A, Sinnah V, Yoko N, Rogers H, Safai S et al. 2016. An Outbreak of Ebola Virus Disease in the Lassa Fever Zone. J Infect Dis. 214(suppl 3):S110-S121.

Gralinski LE, Menachery VD. 2020. Return of the Coronavirus: 2019-nCoV. Viruses MDPI Open Access Journals. 12:135.

He J, Melnik LI, Komin A, Wiedman G, Fuselier T, Morris CF, Starr CG, Searson PC, Gallaher WR, Hristova K et al. 2017. Ebola Virus Delta Peptide Is a Viroporin. J Virol. 91(16), e00438-17.

He Y, Zhou Y, Liu S, Kou Z, Li W, Farzan M, Jiang S. 2004. Receptor-binding domain of SARS-CoV spike protein induces highly potent neutralizing antibodies: implication for developing subunit vaccine. Biochem Biophys Res Commun. 324:773-81

Hsu, M.C. et al. 1981. Activation of the Sendai virus Fusion protein (F) involves a conformational change with exposure of a new amino terminus. Virology 104, 294-302.

Jiang S, Lin K, Strick N, Neurath AR. 1993. HIV-1 inhibition by a peptide. Nature 1993, 365, 113.

Kilbourne ED. 2006. Influenza pandemics of the 20th century. Emerg Infect Dis. 12:9-14.

Kowalski, M., et al. 1987 Functional Regions of the envelope glycoprotein of human immunodeficiency virus Type 1. Science 237, 1351-1355.

Ksiazek et al. 2003. A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome, N Engl J Med 348:1953-66

Lambert DM, et al. 1996. Peptides from conserved regions of paramyxovirus fusion (F) proteins are potent inhibitors of viral fusion. Proc Natl Acad Sci U S A. 93:2186-91.

Li H, Papadopoulos V. 1998. Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern. Endocrinology. 139:4991-7.

Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M. 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature. 426:450-4.

Luo et al., 1999, Amino Acid Substitutions within the Leucine Zipper Domain of the Murine Coronavirus Spike Protein Cause Defects in Oligomerization and the Ability to Induce Cell-to-Cell Fusion, J.Virol. 73: 8152-8159.

Lupas et al., Predicting coiled coils from protein sequences, Science 252:1162-1165.

Mathewson AC, Bishop A, Yao Y, Kemp F, Ren J, Chen H, Xu X, Berkhout B, van der Hoek L, Jones IM. 2008. Interaction of severe acute respiratory syndrome-coronavirus and NL63 coronavirus spike proteins with angiotensin converting enzyme-2. J Gen Virol. 2008. 89:2741-5

Medinas RJ, et al. 2002. C-Terminal gp40 peptide analogs inhibit feline immunodeficiency virus: cell fusion and virus spread. J Virol. 76:9079-86.

Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, Keith JC Jr, McCoy JM. 2000. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403:785-9.

Miller, F.A., et al. 1968. Antiviral activity of carbobenzoxy di- and tripeptides on Measles virus. Applied Microbiology 16, 1489-1496.

Montefiori, D. C., W. E. Robinson, Jr., S. S. Schuffman, and W. M. Mitchell. 1988. Evaluation of antiviral drugs and neutralizing antibodies to human immunodeficiency virus by a rapid and sensitive microtiter infection assay. J. Clin. Microbiol. 26**:**231-235

Morosov VA, Morozov AV, Sernaan, M, Denner. J. 2012. Single mutations in the transmembrane envelope protein abrogate the immunosuppressive property of HIV-1. Retrovirology 9:67.

Morrison TG. 2003. Structure and function of a paramyxovirus fusion protein. Biochimica et Biophysica Acta. 1614:73-84.

Moscona et al. 1992. Fusion Properties of Cells Infected with Human Parainfluenza Virus Type 3: Receptor Requirements for Viral Spread and Virus-Mediated Membrane Fusion, Journal of Virology, 66:6280-6287.

Müller C, Schulte FW, Lange-Grünweller K, Obermann W, Madhugiri R, Pleschka S, Ziebuhr J, Hartmann RK, Grünweller A. 2018. Broad-spectrum antiviral activity of the eIF4A inhibitor silvestrol against corona- and picornaviruses.Antiviral Res.150:123-129.

Muramatsu T et al. SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. Proc Nat Acad Sci USA 113:12997–13002

Nicolaides, E., et al. 1968. Potential antiviral agents. Carbobenzoxy di- and tripeptides active against Measles and herpes viruses. Journal of Medicinal Chemistry 11, 74-79.

Nieva JL, Aitziber A. 2003. Are fusion peptides a good model to study viral cell fusion? Biochimica et Biophysica Acta. 1614:104-115.

Norrby, E. 1971. The effect of a carbobenzoxy tripeptides on the biological Activities of measles virus. Virology 44, 599-608.

Pinon et al., 2003, An Antiviral Peptide Targets a Coiled-Coil Domain of the Human T-Cell Leukemia Virus Envelope Glycoprotein, J. Virol. 77:3281-3290.

Poutanen SM, et al. 2003. Identification of Severe Acute Respiratory Syndrome in Canada. N Engl J Med. 348:1995-2005.

Qureshi NM1, Coy DH, Garry RF, Henderson LA. 1990. Characterization of a putative cellular receptor for HIV-1 transmembrane glycoprotein using synthetic peptides. AIDS 4:553-8.

Renard M, Varela PF, Letzelter C, Duquerroy S, Rey FA, Heidmann T. 2005. Crystal structure of a pivotal domain of human syncytin-2, a 40 million years old endogenous retrovirus fusogenic envelope gene captured by primates. J Mol Biol. 352:1029-34.

Richardson, C.D., et al. 1980. Specific inhibition of Paramyxovirus and myxovirus replication by oligopeptides and amino acid Sequences similar to those at the N-termini of the F1 or HA2 viral polpeptides. Virology 105, 205-222.

Richardson, C.D., and Choppin, P.W. 1983. Oligopeptides that specifically Inhibit membrane fusion by paramyxoviruses: studies on the site of action. Virology 131, 518-532.

Rimsky et al., 1998, Determinants of Human Immunodeficiency Virus type 1 Resistance to gp41-derived Inhibitory Peptides, J. Virol. 72:986-993.

Root MJ, et al. 2001. Protein design of an HIV-1 entry inhibitor. Science. 291(5505):884-8.

Rota PA, et al. 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300:1394-9.

Sainz B JR, Rausch JM, Gallaher WR, Garry RF, Wimley WC. 2005. Identification and characterization of the putative fusion peptide of the severe acute respiratory syndrome-associated coronavirus spike protein. J Virol. 79(11):7195-206.

Sainz B Jr, Rausch JM, Gallaher WR, Garry RF, Wimley WC. 2005. The aromatic domain of the coronavirus class I viral fusion protein induces membrane permeabilization: putative role during viral entry. Biochemistry. 44(3):947-58.

Sainz B Jr, Mossel EC, Gallaher WR, Wimley WC, et al. 2006. Inhibition of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) infectivity by peptides analogous to the viral spike protein. Virus Res. 120(1-2):146-55.

Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, Wang W, Eden JS, Shen JJ, Liu L, Holmes EC, Zhang YZ. 2018. The evolutionary history of vertebrate RNA viruses. Nature 556:197-202.

Sia SK, et al. 2002. Short constrained peptides that inhibit HIV-1 entry. Proc Natl Acad Sci U S A. 99:14664-9.

Song W, Gui M, Wang X, Xiang Y. 2018. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. PLoS Pathog. 14(8):e1007236.

Suarez et al., 2000, Membrane Interface-Interacting Sequences within the Ectodomain of the Human Immunodeficiency Virus type 1 Envelope Glycoprotein: Putative Role During Viral Fusion, J. Virol. 74:8038-8047.

Taubenberger JK, Morens DM. 2006. 1918 influenza: the mother of all pandemics. Emerg Infect Dis. 12:15-22.

Tsang KW, et al. 2003. A Cluster of Cases of Severe Acute Respiratory Syndrome in Hong Kong. N Engl J Med. 348:1977-85.

Ujike M et al. 2008. Heptad Repeat-Derived Peptides Block Protease-Mediated Direct Entry from the Cell Surface of Severe Acute Respiratory

Syndrome Coronavirus but Not Entry via the Endosomal Pathway. J Virol 82:588-592.

Vishwanathan SA, Thomas A, Brasseur R, Epand RF, Hunter E, Epand RM. 2008. Large changes in the CRAC segment of the gp41 of HIV does not destroy fusion activity if the segment interacts with cholesterol Biochemistry. 47:11869–11876.

White, J. M., 1992. Membrane Fusion. Science, 258:917-924.

Wild C, Oas T, McDanal C, Bolognesi D, Matthews T. 1992. A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. Proc Natl Acad Sci USA. 89:10537-10541.

Wild et al., 1994, Propensity for a Leucine Zipper-Like Domain of Human Immunodeficiency Virus Type 1 gp41 to Form Oligomers Correlates With a Role in Virus-Induced Fusion Rather Than Assembly of the Glycoprotein Complex, Proc. Natl. Acad. Sci. USA 91:12676-80.

Wilson, I.A., et al. 1981. Structure of the haemagglutinin Membrane glycoprotein of influenza virus at 3 A resolution. Nature 289, 366-373.

Yang H, Xie W, Xue X, Yang K, Ma J, et al. 2005 Design of wide-spectrum inhibitors targeting coronavirus main proteases. PLoS Biol 3(10): e324.

Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, Lu G, Wu Y, Yan J, Shi Y, Zhang X, Gao GF. 2017. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. Nat Commun 8:15092.

Xia S, Wei Xu, Qian Wang, Cong Wang, Chen Hua, Weihua Li, Lu Lu, Shibo Jiang. 2018. Peptide-Based Membrane Fusion Inhibitors Targeting HCoV-229E Spike Protein HR1 and HR2 Domains. Int J Mol Sci. 19:487.

About the Authors

William R. Gallaher Ph.D. and Andrew D. Gallaher

William R. Gallaher, Ph.D, is a senior scientist with 53 years of experience investigating virus infection. He is published in Newcastle Disease Virus, Herpes Simplex virus, HIV-1, Ebola, mouse hepatitis virus, SARS, influenza and now Wuhan nCoV2019. Retired as Professor Emeritus from LSU Health, New Orleans, he is Chief Scientific Officer at Mockingbird Nature Research Group, a bioinformatics and consulting corporation located in Pearl River, Louisiana.

Andrew D. Gallaher has been working with his father since 2016 on discovery of immunosuppressive mechanisms and cytotoxic motifs in human and animal viruses. Father and son presented together at the recent Annual Meeting of the American Society for Microbiology, South Central Branch. Appointed a Staff Scientist at Mockingbird in the summer of 2019, he consults in his free time from Arlington, Virginia, while serving his country at the Pentagon as a Master Sergeant in the United States Marine Corps. He will soon complete his 19th year of service as a veteran of the war on terror.