A. Specific Aims. The 21st century has recorded the emergence of three highly pathogenic respiratory coronaviruses, including Severe Acute Respiratory Coronavirus (SARS-CoV) in 2003, the ongoing Middle East Respiratory Coronavirus (MERS-CoV) in 2012 and a recent SARS-like coronaviruses (SARS2) in Wuhan, China in Dec 2019. The scope of the current SARS2 pandemic is enormous, resulting in >200,000 cases in >100 countries and a ~3.4% mortality rate. Our renewal focus exclusively on the viral and host determinants that regulate the atomic-level interactions between the SARS2 S-glycoprotein and mammalian ACE2 receptors and other entry factors, like cellular proteases, revealing evolutionary mechanisms that guide SARS2 species specificity, tropism and pathogenesis. We will reveal the evolutionary pathways that regulate clade1 SARS2 virus emergence and growth in human primary cells and in mice, which is critical for countermeasure development.

To accomplish these goals, two highly experienced senior researchers apply an array of state of the art techniques which include a SARS2 recombinant viruses, zoonotic clade 1 SARS2-like precursors from bats and pangolins, SARS2 mouse adapted variants, and primary human airway epithelial cells developed in the Baric laboratory. In parallel, the Li laboratory builds upon an array of known crystal structures for SARS-CoV and related strains, coupled with new empirical studies that reveal the atomic level determinants regulating epidemic SARS2 and zoonotic SARS2-related S glycoprotein-receptor interactions across key intermediate species. His laboratory also provides novel biochemical assays that quarry key entry determinants that will reveal new insights into SARS2-like CoV cross species transmission mechanisms in vitro and pathogenesis in vivo. The overall premise is that that SARS-like clade 1 SARS2 epidemic and zoonotic precursor strains use novel atomic level interaction networks to regulate entry, cross species transmission and pathogenesis in mammals.

Specific Aim 1: Receptor Recognition and cross-species transmission of SARS2 How does receptor usage contribute to cell infection and cross-species transmission of the clade 1 SARS2 like CoV? We hypothesize that receptor usage regulates the infectivity and host range of SARS2 and related SARS2 strains. Our pioneering structural and molecular studies on the receptor recognition and cross-species transmission of SARS-CoV laid the foundation for novel research on SARS2. Recently we also determined the structural basis for the human receptor recognition by SARS2, shedding light on the cross-species transmission mechanisms. Here we will investigate the molecular and structural basis for: (**i**) bat receptor recognition by SARS-CoV-2; (ii) human and bat receptor recognition by RaTG13 (a SARS2-related bat CoV); and (**iii**) recognition of ACE2 from other animal species by SARS2. In parallel, we will elucidate the atomic level structures and evolutionary pathways that regulate SARS2 and SARS2-like ACE2 binding, host range and emergence mechanisms.

Specific Aim 2: Cell entry and cross-species transmission of SARS2. SARS2 has a unique collection of predicted protease cleave sites in the S glycoprotein, leading to the premise that it uses unique collections of cellular proteases to regulate infectivity, tropism and host range. Our previous research in SARS-CoV and MERS-CoV, which proved that host specific protease activities regulate cross species transmission, uniquely position us to compare the *role* of species-specific proteases in SARS2 entry, tropism and host range. We will investigate the: (**i**) physiological triggers that reprogram SARS2 and RaTG13 spike conformations required for human cells entry; (**ii**) electron microscopic studies of the conformational changes of SARS2 and RaTG13 spike proteins; (**iii**) processing programs of SARS2 and RaTG13 by human and bat proteases. This aim will elucidate the mechanisms by which host cellular proteases function to program infection and host range in vitro.

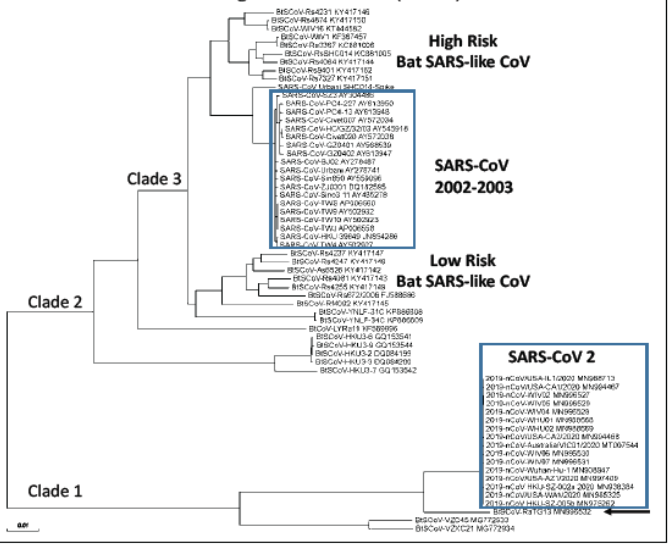
Specific Aim 3: Pathogenesis of SARS2 in Mouse Models of Human Disease. SARS2 can use the human (hACE2), but not mouse (mACE2) receptor for entry. The overall goal is to identify key interaction networks that regulate SARS2 and zoonotic SARS2-like CoV RaTG13 and PaDCoV cross species transmission and pathogenesis in mice and to recapitulate a robust lethal mouse model of human disease. Our underlying hypothesis is that microvariation with the clade 1 SARS2-like CoV S glycoprotein regulates alternative mammalian ACE2 receptor binding profiles and protease usage, altering human tissue tropism for different human primary lung cells in vitro and in vivo pathogenesis in mice. In part 1, we will identify the mouse ACE2 receptor residues that impede SARS2 and SARS2-like CoV binding and entry in vitro. In part 2, we evaluate SARS2 and SARS2-like CoV replication and pathogenesis in mice encoding CRISPR gene-edited chimeric mACE2 receptors that encode permissive hACE2 residues. In part 3, we evaluate the outcomes of in vitro evolved and structure-guided mutations from Aims 1 and 2 on virus replication and pathogenesis in vivo

Implication beyond SARS-CoV-2. How should we prevent a disaster like the current SARS2 outbreak from occurring again? The implication of this proposed research goes beyond SARS2 because it aims to understand the evolutionary drivers of zoonotic SARS2 CoV cross species transmission, emergence and pathogenesis, while producing robust mouse models of human disease, which will be used to evaluate countermeasure performance against SARS2 and related strains of the future.

Baric R: Cell entry, cross-species transmission and pathogenesis of novel CoV, SARS2 from Wuhan

B. Introduction. The underlying hypothesis posits that the clade 1 group 2b SARS2 and SARS2-like CoV represent existing and future threats to human health and the global economy. The program uses integrated approaches that couple biochemistry, structural biology, experimental evolution and virus reverse genetics to reveal novel insights into clade 1 SARS2 and related virus-receptor interactions, cross species transmission mechanisms, animal model development and pathogenesis in vivo. The program not only provides a conceptual framework that reveals new insights into the global threat potentials of clade 1 SARS2 like viruses, but also develops key resources needed for diagnostics and countermeasure development. We note that the previous funding cycle resulted in 29 manuscripts, including high profile journals like PNAS, Nature Biotechnology, Nature Microbiology, mBIO, Plos Pathogens and Nature Review Microbiology. Some emerging CoV highlights include new animal models of MERS-CoV disease (*1, 2*), hmAB countermeasure discovery (*3*), host proteases regulate group 2C cross species transmission potential (*4-6*), novel CoV S structures (*7*), and isolation of high risk group 2c MERS-like bat CoV (*8, 9*). By focusing the renewal of SARS2, we have changed the title of the proposal as it is designed to make significant contributions to SARS2 global health preparedness, animal model and countermeasure development.

Figure 1. Phylogeny of the group 2b SARS-like CoV. The high risk SARS-like Bat CoV are encoded within clade 1 (SARS2) and clade 3 (SARS-CoV, SHC014, WIV1, ec.). The closest relative to SARS2 is a bat CoV designated RaTG13 (arrow).



B.1. Overview of SARS2, SARS-CoV and related bat CoVs: Coronaviruses (CoV) have a long history of host shifting and most human CoVs likely originated from bats over the past ~600 years. Emerging respiratory viruses represent a huge threat to global health and the stability of global economies. SARS-CoV emerged in 2003 to cause ~8,000 infections and ~800 deaths (*10-13*). A number of closely related SARS-like bat viruses, identified between 2013-2016, were shown to use human ACE2 (hACE2) receptors in vitro and in vivo, predicting future emergence events in human populations (**Fig 1**)(*14, 15*). In 2019, a novel clade I SARS-like group 2b CoV, designated SARS-CoV2 (SARS2) emerged in Wuhan, China in Dec 2019, resulting in >205,000 infections and >8,000 deaths in >100 countries; a global pandemic. The SARS2 mortality rate is ~3.5%, and the elderly experience mortality rates of ~14% (*16-20*). Over 7,000 cases have been reported in all US states. Many of the symptoms caused by SARS2, such as acute respiratory syndrome (ARDS), are similar to those caused by the 2003 SARS-CoV. SARS2 and SARS-CoV are also genetically related, but differ by ~22% (**Fig 1**). Alarmingly, SARS2-related CoVs also exist in bats and pangolins (*21*). Bat virus RaTG13, shares 96% genomic identity with SARS2, implicating RaTG13 as a potential threat (31). Although pangolins do not appear to be a reservoir to SARS2, pangolin virus is ~10% different and encode an S glycoprotein-receptor binding domain (RBD) that is primed to infect primate cells as well (**Table 1**). Consequently, the animal reservoirs, evolution and long-term threat potential of the epidemic and SARS2-like CoV remains unclear. We investigate the cell entry mechanisms, pathogenesis and cross-species transmissibility of the SARS2 and SARS2-related CoV. This research is critical for understanding the bio-threat level of these viruses, for developing robust mouse models of human disease, and for preventing and controlling future spread of SARS2 and SARS2-like CoVs in humans.

Table 1. SARS and SARS2 CoV RBD-ACE2 Binding Residues

SARS aa#	402	404	426	436	440	442	443	460	472	473	475	479	484	486	487	488	491
SARS-CoV	T	V	R	Y	Y	Y	L		L	N	Y	N	Y	T	T	G	Y
2019-nCoV aa#	415	417	439	449	453	455	456	473	486	487	489	493	498	500	501	502	505
2019-nCoV	T	K	N	Y	Y	L	F	Y	F	N	Y	Q	Q	T	N	G	Y
Pangolin CoV-GD	T	R	N	Y	Y	L	F	Y	F	N	Y	Q	H	T	N	G	Y
RaTG13	T	K	K	F	Y	L	F	Y	L	N	Y	Y	Y	T	D	G	H

Blue # in SARS-RBD are the identified residues contating ACE2 in Li et al., Science, 2005
Red # in 2019-nCoV-RBD are the identified residues contating ACE2 in Yan et al., Science, 2020

B.2. CoV-receptor interactions: Receptor recognition is an important determinant of the host range and tropism of CoVs (*22, 23*). An envelope-anchored spike protein mediates CoV entry into host cells by first binding to a host receptor through its S1 subunit and then fusing the host and viral membranes via its S2 subunit. A defined receptor-binding domain (RBD) in SARS-CoV spike S1 subunit binds human ACE2 protein with high affinity (*24-28*). Based on our previous structural studies of SARS-CoV RBD/ACE2 interactions, we have predicted that

SARS2 RBD also recognizes ACE2 as receptor (*29*), which has been validated by other studies (*30-33*). Moreover, we have just determined the crystal structure of SARS2 RBD complexed with human ACE2, revealing unique features of receptor recognition and new insights on the animal-to-human transmission of SARS2 (study accepted by Nature pending minor revision for publication) (**Fig 2**). Our previous research had identified critical residue changes in the SARS-CoV RBD that allowed SARS-CoV to transmit from animals to humans and from humans to humans (53-56). A 3.5 Å-resolution cryo-EM structure of the SARS-CoV-2 S trimer in the prefusion conformation has also been solved (*34*). However, receptor usage by SARS2, RaTG13 and PaDCoV strains remains unresolved, and will clearly identify novel RBD-ACE2 atomic level interaction networks that regulate infection potential in humans and animals.

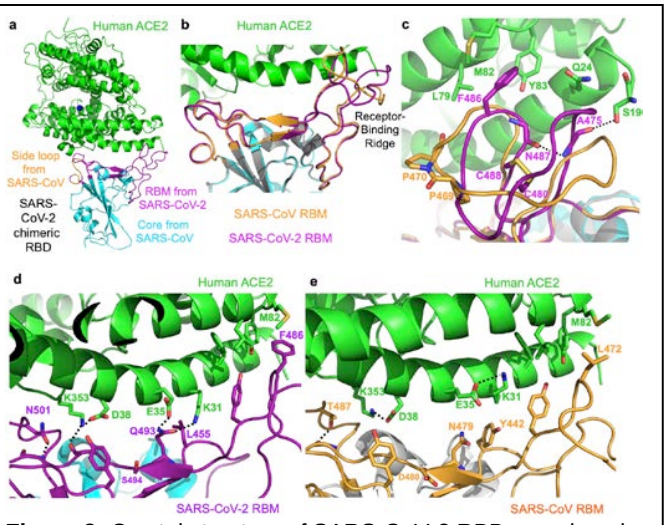
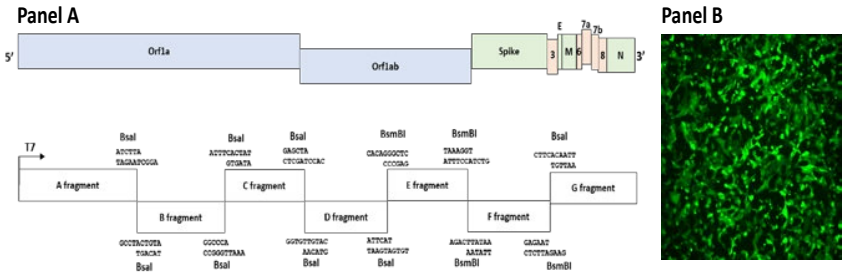


Figure 2: Crystal structure of SARS-CoV-2 RBD complexed with hACE2. (A) Overall structure. (B)-(E) Structural details of the SARS-CoV-2/hACE2 interface, showing critical interactions that determine the receptor recognition and cross-species transmission of SARS-CoV-2.

B.3. CoV entry into host cells: Coronavirus entry into host cells is limited not only by receptor recognition, but also by membrane fusion. To fuse host and viral membranes, coronavirus spikes need to be cleaved at two sites (one at the S1/S2 boundary and the other within S2) by host proteases (*35-39*). The availability of these host proteases to coronaviruses and the specificities of these host proteases on CoV spikes can contribute to the host range and tropism of coronaviruses. Coronavirus-spike-processing host proteases may include proprotein convertases from the virus-producing cells (e.g. furin), proteases from the extracellular environment (e.g. elastase), proteases on the cell surface of virus-targeted cells (e.g. TMPRSS2), and proteases in the endosomes of virus-targeted cells (e.g. cathepsins) (*35, 36*). The SARS2 CoV also encodes a novel furin cleavage site, whose function in replication and pathogenesis is studied in Aim 2 and 3 (*40*). We previously identified different protease activities as an important determinant for CoV host range and tissue specificity (*6, 41, 42*). Therefore, comparative study of the cell entry of SARS2, SARS-CoV, and SARS2 like bat CoV RaTG13 and pangolin CoV will reveal how SARS2 target humans, reveal likely reservoir species and the human threat potential of other zoonotic SARS2-CoV strains.

Figure 3. SARS-CoV-2 Recombinant Clone Design and Development of Reporter Viruses for Neutralization Assays. Panel A: SARS-CoV-2 Genome and Molecular Clone. Panel B: GFP expression from Reporter Virus.



B.4. SARS2 Reverse Genetics and Pathogenesis. The Baric laboratory is a global leader in the development of reverse genetic platforms for emerging CoV including human and animal strains, like SARS-CoV, MERS-CoV, WIV-1, WIV-16, SHC014, PEDV, SADS-CoV, HCoV NL63 and others (*14, 15, 43-45*). We have designed and recovered recombinant SARS2 viruses using the molecular clone and isolated reporter viruses encoding GFP and nLUC (**Fig 3**). Using reverse genetics, two different SARS2 recombinant virus strains are available, including the Seattle (MN985325) and a very early Dec 2019 Wuhan (MN996528) isolates. We are also generating full-length molecular clones for RaTG13 and pangolin CoV (PaDCoV), plus and minus indicator genes, which provide simple assays to evaluate the role of mammalian ACE2 receptors and proteases on entry and release. Reverse genetics will also support the recovery of mouse adapted strains (SARS2 MA), as described for SARS-CoV (*46, 47*).

B.5. Program Significance: The proposed research will elucidate the receptor usage, cell entry and cross-species transmission mechanisms, and pathogenesis of SARS2, PaDCoV and RaTG13. It will produce reverse genetic platforms and animal models for the SARS2-like CoV. This research will illuminate SARS2 tropism in primary human lung cells and the mechanisms by which SARS2 and SARS2-like bat CoVs adapt to human and other cells during cross-species transmission events. It will also provide valuable insights on the emerging

disease potentials of new SARS2-like CoVs. For example, the SARS2-like CoV animal reservoirs remain unknown and may change, as the virus samples new wildlife species regionally as it spreads across the globe.

C. Innovation

The primary innovation of the program stems from the proposed comparative analysis of the receptor recognition, cell entry, cross-species transmission, and pathogenesis mechanisms of SARS2, RaTG13 and SARS-CoV in human and bat cells. Although bats are the likely natural reservoir of all or most of human CoVs, the cell entry mechanisms and pathogenesis of bat CoVs are poorly understood. Thus, we propose to study both SARS2 and SARS2-related bat CoV like RaTG13, and SARS-CoV 2003. Moreover, instead of only investigating SARS2 entry and pathogenesis in human cells, we propose to study SARS2 in cells from other species, particularly bats and mice. These approaches will fill in critical missing links in our knowledge about the cell entry mechanisms and pathogenesis of bat CoVs, identify potentially novel molecular mechanisms by which bat CoVs enter host cells, and discover key host factors in bat cells that are critical for CoV infections but are unknown to virologists. The above knowledge about bat CoVs and bat cells, when coupled with the entry mechanisms used by SARS2 infection in human and bat cells, will reveal a comprehensive picture regarding the molecular mechanisms governing the group 2b SARS-like CoV trans-species movement to humans, either directly or through potential intermediate hosts. Importantly, SARS2 global spread raises the spectra for new reservoir hosts in North America and beyond, which may play a critical role in maintaining endemic persistence across the globe. Therefore, the proposed research will go beyond the direct implications for preventing and controlling the spread of SARS2 infections in humans, and will establish the fundamental virologic principles that govern cell entry, pathogenesis, and cross-species transmission potential of the 3 group 2b SARS like CoV.

The second program innovation is reflected in the combination of in vitro and in vivo experimental approaches to study CoV entry and pathogenesis, leading to the development of mouse models for SARS2, RaTG13 and PaDCoV pathogenesis and heterologous challenge viruses. As the RaTG13 S gene encodes seven amino acid changes in the RBD, it may recognize distinct mammalian ACE2 receptors and proteases, as compared to SARS2. For example, RaTG13 and PaDCoV might use murine ACE2 receptors. The pangolin CoV spike is reported to program growth in Vero cells, but importantly, encodes a more heterogeneous S glycoprotein (~90 identity), providing a heterologous SARS2-like challenge virus (***doi:*** *https://doi.org/10.1101/2020.02.17.951335*). We hypothesize that these challenge viruses will produce distinct tropisms in primary human airway lung cells, promote replication and disease in standard laboratory, and/or chimeric h/mACE2 transgenic mice, or use different animal ACE2 receptors. They will be critical reagents for determining the breadth of SARS2 vaccines and therapeutic antibodies. Moreover, they will encode unique adaptive mutations following recombinant virus experimental evolution in vitro, either using cells that ectopically over-express different mammalian ACE2 receptors or in mice in vivo, revealing new molecular interaction networks that can govern cross species movement between mammals.

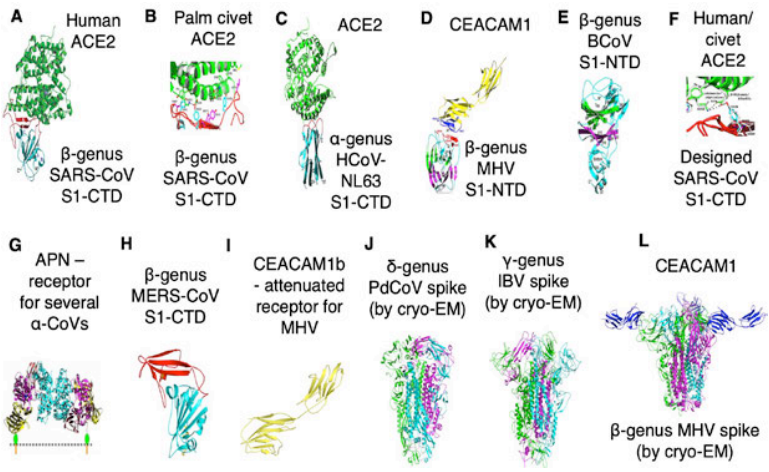


Figure 4. Representative Coronavirus Structures Determined by the Li Lab. (A)-(I) Crystal structures of coronavirus spike RBDs alone, RBD/receptor complexes, or receptors alone. (J)-(L) Cryo-EM structures of coronavirus spike proteins or their complex with receptor.

Dr. Li (see CV) has been working on the structural and molecular basis for the receptor recognition and cell entry mechanisms of SARS-CoV and MERS-CoV. He is recognized for pioneering discoveries on the animal origins and cross-species transmissions of SARS and MERS coronaviruses. He has determined many other structures of coronavirus spike proteins using either X-ray crystallography or cryo-EM (**Fig. 4**), and discovered many molecular mechanisms of coronavirus entry using protein biochemistry and molecular biology tools. His research has been one of the major driving forces in solving the atomic-level interactions that regulate receptor recognition and cell entry of coronaviruses. Previous research on CoVs has firmly demonstrated that it is critical to understand CoV entry and pathogenesis using both in vitro and in vivo approaches. Therefore, the close collaboration between the Baric and Li groups will be key to the success of the proposed research.

The third innovation of the program is the iterative framework of virus-host interaction research that has characterized our collaborations, which aim to identify high threat viruses, facilitate epidemic surveillance and

preparedness, predict species-specific receptor usage, and identify potential animal reservoir hosts and animal models of viruses. More specifically, our decade-long structural studies on the receptor recognition by SARS-CoV have identified key interactions between SARS-CoV spike protein and its host receptor angiotensin-converting enzyme 2 (ACE2), which were used to successfully predict the mammalian ACE2 receptor usage of the SARS2 2019 epidemic strain very early in the outbreak (**Fig 5**)(29). Another long-term goal is to establish a structure-function predictive framework for improved epidemic surveillance. As the group 2b SARS2-like CoV have novel RBD mutation sets and variation in the S glycoprotein, these viruses will reveal new atomic level interaction networks for receptor binding, predict receptor usage and host cell susceptibility and identify potential reservoir strains. By understanding protease usage across species and tissue types, these studies will allow us to incorporate the role of proteases in understanding species specificity and lung cell tropism in vitro and pathogenesis in vivo. An framework that reveals possible animal origins and animal models, based on predictive modeling of S gene sequence functions, provides a rapid approach to identify and cull animal reservoirs early in an epidemic and reveal mammalian species for animal models of human disease, a key bottleneck for countermeasure development. Additionally, structural studies on CoV spikes may reveal important neutralizing epitopes in the RBD, which guide the design of CoV type specific and broad based vaccines.

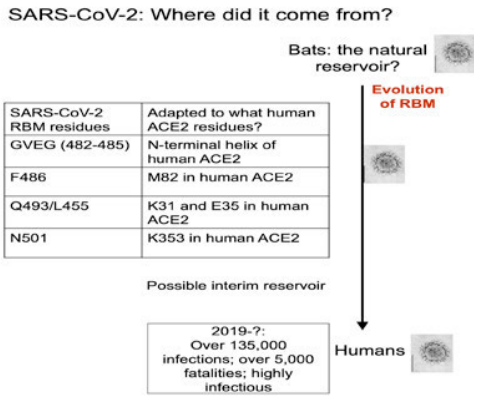


Figure 5. Summary of potential cross-species transmission RBM-hACE2 interactions of SARS2 bound to hACE2.

E. Experimental Approach.

E.1. Aim 1: Receptor recognition and cross-species transmission of SARS2.

Overview: We will investigate how SARS2 and RaTG13 interact with ACE2 from different host species, and how they have evolved to engage human ACE2. Although SARS2 was isolated from humans and RaTG13 from bats, the close sequence similarity between them suggests that SARS-CoV-2 likely originated from RaTG13 or a more closely related bat coronavirus. In this aim, we will identify the structural and molecular basis for ACE2 recognition by SARS2 and RaTG13, aiming to understand the critical structural determinants for SARS-2 to bind human ACE2 and for bat CoVs to transmit from bats to humans. We will also analyze how SARS2 adapts to ACE2 from other species, revealing potential host range and intermediate hosts. Overall, these studies were crafted to provide key data for understanding the emergence potential of the SARS2-like CoVs and for facilitating epidemic monitoring and the development of broadly cross-protective intervention control strategies. We will use a combination of molecular, biochemical and structural methods. Our previous research has demonstrated the importance of these approaches in understanding disease severity, animal origins and cross-species transmissions of viruses. Our expertise places us in a unique position to carry out these studies. *Overall hypothesis of Aim 1*: Receptor recognition by SARS2 is an important determinant of the cell infectivity and cross-species transmission (e.g., bat-to-human transmission, either directly or through intermediate hosts) of SARS2.

Preliminary data supporting Aim 1

Complementary approaches have shown that SARS-CoV species specificity is primarily determined by the affinity between the viral RBD and host ACE2 in the initial viral attachment step (*48-52*). In a span of about 10 years, we determined a series of crystal structures of SARS-CoV RBD complexed with ACE2. These studies included S glycoprotein RBDs derived from different species specific SARS-CoV-related strains bound to the ACE2 receptor orthologues from different animal species (*53-56*). These structures showed that SARS-CoV RBD contains a core structure and a receptor-binding motif (RBM), and that the RBM binds to the outer surface of the claw-like structure of ACE2 (*54*). Importantly, we identified two virus-binding hotspots on human ACE2 (*53, 55*). A number of naturally selected RBM mutations occurred near these two virus-binding hotspot and these residues largely determined the host range of SARS-CoV. Because of the sequence similarity between 2003 SARS-CoV and 2019 SARS2 spikes, we recently predicted that SARS2 also uses human ACE2 as its receptor (*29*), which has been validated by us 9 (see below) and in other studies (*30-33*).

In a preliminary study, we determined the crystal structure of SARS2 RBD (engineered to facilitate crystallization) in complex of hACE2. Compared with SARS-CoV RBD, a hACE2-binding ridge in SARS2 RBD takes a more compact conformation; moreover, several residue changes in SARS2 RBD stabilize two virus-binding hotspots at the hACE2/RBD interface (**Figs 2, 5**). These structural features of SARS2 RBD enhance its hACE2 binding affinity over 2003 SARS-CoV. Additionally, we showed that RaTG13, a bat coronavirus closely

related to SARS2, also uses hACE2 as its receptor. The differences among SARS2, SARS-CoV and RaTG13 in hACE2 recognition shed light on potential animal-to-human transmission modes of SARS2. This study provides guidance for intervention strategies (e.g., hmAB) targeting receptor recognition by SARS2.

E.1.1 Subaim 1 - Molecular and structural basis for bat ACE2 recognition by SARS2

Hypothesis: SARS2 RBD can recognize bat ACE2; differences between human ACE2 (hACE2) and bat ACE2 (bACE2) in their SARS2-interacting residues form a critical cross-species barrier for bat to human or human to bat transmission events, which may seed new reservoir species as this virus circumvents the globe.

Rationale: Because SARS2 likely originated from bats, we test the hypothesis that it still uses bat ACE2 as its receptor. Our structural analysis also suggests that the SARS2 RBD binds bACE2, but that its bACE2-binding affinity is likely lower than its hACE2-binding affinity. Here we will investigate whether bACE2 is the host receptor for SARS2, and characterize the binding affinity between SARS2 RBD and several Asian, African and North American bACE2 (see below). Further, we will determine the crystal structure of SARS2 RBD complexed with bACE2 and identify mutations that may further improve SARS2 RBD's binding to bACE2.

Experimental designs and expected results: (**i**) We will investigate whether SARS2 spike recognizes bACE2. First, we will express and purify bACE2, and examine the binding interactions between SARS2 RBD and bACE2 using surface plasmon resonance (SPR). Second, we will carry out SARS2 Spike-mediated pseudovirus entry into bACE2-expressing cells. Specifically, we will prepare retroviruses pseudotyped with SARS2 spike (i.e., SARS2 pseudoviruses) and use them to enter bACE2-expressing HEK293T cells. For all of these assays, we will use hACE2 as a comparison to select bACE2 and 2003 SARS-CoV spike as a comparison to SARS2 spike. We expect that SARS2 binds bACE2 in a species specific manner, but its bACE2-binding affinity is likely lower than its hACE2-binding affinity. From different continents, we note that ~17 bat species ACE2 gene sequences are available. Representative (e.g., Asian, N. American) bACE2 receptors that promote SARS2 or RaTG13 pseudovirus entry will be prioritized (https://www.ncbi.nlm.nih.gov/gene/59272/ortholog/?scope=40674).

(**ii**) We will crystallize SARS2 RBD complexed with select bACE2. To this end, the SARS2 RBD/bACE2 complex will be purified on gel filtration chromatography and subjected to crystallization screens. X-ray data collection will be performed at synchrotron beamlines. The structure will be determined by molecular replacement using the structure of the SARS2 RBD/hACE2 complex as the search template. We expect that the overall structure of the SARS2 RBD/bACE2 complex is similar to that of SARS2 RBD/hACE2, but the structural details at the interaction interface will differ and account for SARS2's different usages of hACE2 and bACE2.

(**iii**) Based on the crystal structure, we will identify the key residue differences between hACE2 and bACE2 that account for SARS2's different usages of hACE2 and representative bACE2. To evaluate the impact of these key residue differences on SARS2 binding and entry, we will introduce the corresponding key hACE2 residues to bACE2, and will measure how the mutant bACE2 mediates SARS-CoV-2 RBD binding and SARS-CoV-2 pseudovirus entry. We expect that the key residue changes in bACE2 can increase SARS-CoV-2's binding affinity for bACE2 and enhance SARS2 entry into bACE2-expressing cells.

(**iv**) Using recombinant SARS2, RaTG13 and PaDCoV viruses encoding GFP, we will ectopically express the panel of bACE2 and hACE2 receptors in nonpermissive cells and reaffirm live virus entry and replication outcomes. If growth is minimal, serial virus passage should select for mutations that enhance poor bACE2 receptor usage. In parallel, we will also transfect various mutant bACE2 encoding key hACE2 residues into nonpermissive cells and record SARS2, RaTG13, PaDCoV and 2003 SARS-CoV GFP expression and growth.

Potential pitfalls and alternative approaches: We have extensive experience with all of these experimental approaches and hence we do not expect any major technical obstacles. For the alternative hypothesis (i.e., SARS2 RBD does not recognize bat ACE2), we will analyze the sequences of bACE2 and hACE2, identify potential key differences in their SARS2-interacting residues, and introduce mutations to bACE2 until the mutant bACE2 becomes an effective receptor for SARS2 RBD. We will then determine the crystal structure of the SARS2 RBD/mutant bACE2 complex. We anticipate that bACE2 receptors from some bats, which live outside of Asia, will also support SARS2 growth, potentially identifying new reservoir hosts across the globe. If not, we can test ACE2 genes from other N. American mammals (armadillos, possum, raccoon, etc), as SARS2 receptors.

E.1.2 Subaim 2 - Molecular and structural basis for human and bat ACE2 recognition by RaTG13

Hypotheses: RaTG13 RBD is less well adapted to hACE2 as compared with SARS2 RBD; moreover, RaTG13 RBD is better adapted to bat ACE2 than SARS-CoV-2 RBD.

Rationale: In preliminary studies, we already showed that RaTG13 uses hACE2 as its receptor, suggesting that RaTG13 has the potential to infect humans directly, in the absence of intermediate hosts. Hence, it is important to evaluate the hACE2- and bACE2-binding interactions for RaTG13. Although speculative, SARS2 may have been circulating in humans for a certain period of time before the current outbreak, and hence it may

have become well adapted to hACE2. On the other hand, RaTG13 was isolated directly from bats and should not have adapted to hACE2. In this subaim, we will characterize biochemically, the binding interactions between RaTG13 RBD and hACE2 or bACE2. We will further compare RaTG13 RBD to SARS2 RBD and SARS-CoV RBD in the hACE2- and bACE2-binding affinities. Moreover, we will determine the crystal structures of RaTG13 RBD complexed with hACE2 or bACE2, and identify key residue changes between RaTG13 RBD and SARS2 RBD or SARS-CoV RBD that account for their differential usages of hACE2 and bACEC2.

Experimental designs and expected results: (**i**) We will measure the binding affinities between RaTG13 RBD and hACE2 or bACE2 using SPR. We will also measure the efficiencies of RaTG13 pseudovirus entry into hACE2-expressing cells or bACE2-expressing cells, and further compare them to those of SARS2 and SARS-CoV pseudoviruses. We expect that RaTG13 RBD binds to hACE2 less strongly than SARS2 or SARS-CoV RBD does, and that RaTG13 RBD binds to bACE2 more strongly than its binds hACE2. Finally, we predict that bACE2 receptors, derived from some bats across the globe, will support RaTG13 growth.

(**ii**) We will crystallize RaTG13 RBD complexed with hACE2, and determine the structures of the complexes by molecular replacement using the structure of SARS2 RBD/hACE2 complex. We expect that key residue changes in RaTG13 RBD and SARS2 RBD will account for their differential usage of hACE2.

(**iii**) Based on the above crystal structure, we will identify key RBD residue differences among RaTG13, SARS2 and 2003 SARS-CoV, and evaluate the impact of these RBD differences on hACE2 binding and hACE2-dependent pseudovirus entry. We will introduce these key residue differences to RaTG13 RBD and RaTG13 pseudoviruses, and will measure the hACE2 binding affinity and hACE2-dependent pseudovirus entry. We expect that the key residue changes in RaTG13 RBM can increase RaTG13's binding affinity for hACE2 and enhance RaTG13 entry into hACE2-expressing cells.

(**iv**) Similarly, we will determine the crystal structure of RaTG13 RBD complexed with select bACE2. We expect that key residue changes between bACE2 and hACE2 will account for the differential usages of hACE2 and bACE2 by RaTG13 RBD.

(**v**) We will introduce the key bat ACE2 interaction residue differences to hACE2, and measure how the mutant hACE2 support RaTG13 RBD binding and RaTG13 pseudovirus entry. We expect that the key residue changes can render hACE2 to become a more efficient receptor for RaTG13 docking and entry.

(**vi**) We will introduce and test select key hACE2 RBD-enhancing mutations, using live attenuated RaTG13 that encode indicator genes for growth in cell culture (**Fig 3**). Recognizing the P3C0 potential of these experiments, we will first attenuate RaTG13 by introducing changes that ablate the RaTG13-GFP nsp16 2'*O*-methyltransferase (2'*O*-MTase), as described previously by our group (*57, 58*). Then, we will ectopically over-express various hACE2 and bACE2 receptors and evaluate RaTG13-GFP-2OMT^{-} virus growth.

Potential pitfalls and alternative approaches: As described in Aim 2, cryo-EM structure of RaTG13 spike protein complexed with hACE2 or bACE2 will serve as an alternative for X-ray crystallography in understanding RaTG13/ACE2 interactions. We are well versed in these techniques.

E.1.3 Subaim 3 – Recognition of ACE2 from different host species (other than bats) by SARS-CoV-2.

Hypothesis: SARS2 RBD can recognize ACE2 from a diversity of animal species, except mice and rats; the binding affinities between SARS2 RBD and ACE2 from different animal species can shed light on the animal origins and potential reservoir species of SARS2.

Rationale: Based on the sequence alignment of ACE2 from different species at critical virus-binding residues, we predicted that SARS2 RBD can recognize ACE2 from a diversity of animal species, except mice and rats (*29*). Consequently, we will experimentally test these predictions using biochemical and structural approaches.

Experimental designs and expected results: (**i**) Using the list server of different species ACE2 genes shown above, we will express and purify ACE2 from different species, and measure the binding affinity between SARS2 RBD and ACE2 from different species. We will perform SARS2 pseudovirus entry into cells exogenously expressing ACE2 from different species (e.g., mice, pangolins, dogs, cats, pigs, bear, horse, armadillo, opossum, palm civets, etc), including some from N. America.

(**ii**) We will initially focus on ACE2 from pangolins and palm civets because pangolins have been proposed as the likely intermediate hosts for SARS2 and palm civets were the intermediate hosts for 2003 SARS-CoV. We will determine the structures of SARS2 RBD complexed with each of these ACE2 molecules and identify key residue differences between these ACE2 molecules. Then we interrogate ACE2 from other species.

(**iii**) We will identify key residue differences among select ACE2 molecules that prevent their function as receptors for SARS2. To this end, we will then introduce these key residues into hACE2 and analyze how these mutations affect SARS2 RBD/hACE2 binding and SARS2 pseudovirus entry.

(**iv**) We will test mutant ACE2 to support virus growth using SARS2, RaTG13 and PaDCoV viruses encoding GFP in cell culture. Briefly, we will ectopically express each mutant and wildtype ACE2 receptor in nonpermissive cells and record virus replication and GFP expression. If attenuated growth occurs, serial passage should select for viral RBD mutations that enhance mutant ACE2 receptor usage.

Potential pitfalls and alternative approaches: As described in Aim 2, host proteases may serve as another cross-species barrier for SARS-CoV-2. PaDCoV can also be studied in detail as appropriate.

E.2 Specific Aim 2: Cell Entry and Cross-species Transmission of SARS2

Overview: We will investigate how host cellular proteases determine cell entry and cross-species transmission of SARS2. To fuse membranes, CoV spikes must be proteolytically activated, programming dramatic structural changes to transition from the pre-fusion to post-fusion conformation (*35, 36*). These CoV-entry-activating proteases include proprotein convertases (PPCs) such as furin, cell-surface proteases such as TMPRSS2, lysosomal proteases such as cathepsins, and extracellular proteases such as trypsin-like proteases (*35, 36*). Among all of the potential protease activators, only TMPRSS2 has been shown to activate SARS2 entry (*59*). Thus the cell entry mechanism of SARS2 and RaTG13 is still poorly defined, hindering antiviral battles against SARS2. Moreover, receptor binding may also play an important role in the final structural change of CoV spikes. Aim 2 will investigate the roles of protein activators and receptor binding in the structural change of SARS2 and RaTG13 spikes and its role in SARS2 like CoV cell entry.

Importantly, the role of bat cellular proteases in the cell entry of SARS2 and RaTG13 is unknown. Our previous research on the cross-species transmissions of MERS-CoV and the closely related bat coronavirus HKU4 showed that there exist intrinsic differences between human and bat cellular proteases that contribute to the differential processing of the HKU4 spike in human and bat cells. We will investigate whether the intrinsic difference in human and bat proteases are a key factor in bat-to-human transmission of SARS2 and RaTG13.

Overall hypothesis of Aim 2: (i) Both host protease activators and receptor binding play critical roles in the structural changes of SARS2 and RaTG13 spikes and in the cell entry of SARS2 and RaTG13; (ii) intrinsic differences between human and bat cellular proteases represent another critical cross-species for the potential bat-to-human transmission of SARS2 and RaTG13.

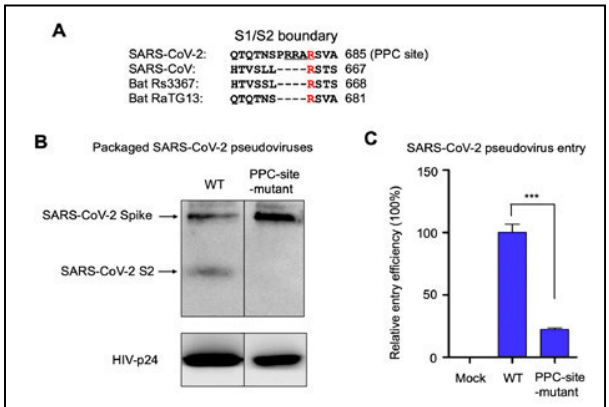


Figure 6: Proprotein convertase motif in SARS2 spike. (A) Sequence comparison of the spike proteins from SARS2, SARS-CoV, and two bat SARS-like coronaviruses in a region at the S1/S2 boundary. (B) Cleavage state of SARS2 spike on the surface of pseudoviruses. (C) SARS2 pseudovirus entry into hACE2-expressing Hela cells. The two types of pseudoviruses correspond to the pseudoviruses in panel (B).

Preliminary data supporting Aim 2

Through examining the sequence of SARS2 spike protein, we identified a putative cleavage site for PPCs (e.g., furin) at the S1/S2 boundary (**Fig. 6A**). Curiously, this putative PPC site is absent in the spikes of 2013 SARS-CoV spike and other SARS-like bat CoV. We investigated the role of PPCs in SARS2 entry. First, to detect the cleavage state of the SARS2 spike on the surface of pseudoviruses, we packaged SARS2 pseudoviruses in HEK293T cells and performed Western blot on the pseudoviruses. The result showed that SARS2 spike had been cleaved during viral packaging (**Fig.6B**). We then mutated the putative PPC site in SARS2 spike to the corresponding sequence in SARS-CoV spike; the mutant SARS2 spike was not cleaved during viral packaging (**Fig.6B**). Further, we performed pseudovirus entry experiments using both wild type SARS2 pseudovirues and PPC-site-mutant SARS2 pseudoviruses. The results showed that SARS2 pseudoviruses efficiently entered hACE2-expressing Hela cells (**Fig. 6C**). In contrast, the mutant SARS2 pseudoviruses demonstrated significantly reduced efficiency entry into the same cells (**Fig. 6C**). Therefore, we have identified and confirmed the PPC cleavage site in SARS2 spike, and shown that PPC cleavage of SARS2 spike is critical for SARS2 entry.

E.2.1 Subaim 1 – Physiological Triggers that Reprogram SARS2 and RaTG13 Spike Conformations required for Entry into Human Cells. Hypothesis: Both host protease activators and receptor binding play critical roles in the structural changes of SARS2 and RaTG13 spikes and in cell entry of SARS2 and RaTG13.

Rationale: In preliminary studies, we found that PPC cleavage of SARS2 spike is important for cell entry. Our previous studies also showed that lysosomal proteases and trypsin play critical roles in CoV entry and tropism. Moreover, for CoV spikes to undergo final structural changes, the spikes need to be cleaved at two sites, one between S1 and S2 (i.e., S1/S2 site) and within S2 (i.e., S2' site). The formation of S2' fragments is strongly

associated with the final structural changes of CoV spikes. We have recently performed in-depth analysis of the triggers and structural changes of mouse CoV (MHV) spike protein (*60*). In this subaim, we will systematically examine the roles of proteases and receptor binding in human cell entry to SARS2 and RaTG13 using pseudovirus entry assay, protease sensitivity assay and negative-stain EM assay.

Experimental designs and expected results: (**i**) We will examine the role of lysosomal proteases in the cell entry of SARS2. To this end, we will perform SARS2 pseudovirus entry assay in the presence of lysosomal protease inhibitors. We will test endosomal acidification inhibitors and specific inhibitors that target either cathepsins in general or individual cathepsins (e.g., cathepsin L and cathepsin B). We expect that cathepsin inhibitors will reduce the cell entry efficiency of SARS2 and RaTG13 pseudoviruses. We also expect that one or more specific cathepsins be involved in activating SARS2 and RaTG13 pseudovirus entry.

(**ii**) We will examine the role of trypsin in the cell entry of SARS2. To this end, we will perform SARS2 and RaTG13 pseudovirus entry assay in the presence of trypsin. We expect that trypsin activates SARS2 and RaTG13 pseudovirus entry, as evidenced by increased expression of indicator genes.

(**iii**) We will investigate the role of the identified protease activators and receptor binding in the final structural changes of SARS2 and RaTG13 spikes, using protease sensitivity assays. We will express and purify the protease activators that we will have identified, like furin or one of the cathepsins. We will perform the protease sensitivity assay on the SARS2 and RaTG13 spikes in the presence of the proteases and/or hACE2. We expect that the combination of one or more of the proteases plus hACE2 binding will allow SARS2 and RaTG13 spikes to become more sensitive to proteolysis and be degraded to S2' fragment. The appearance of S2' fragment will strongly suggest that the final structural changes of the SARS2 and RaTG13 spikes have taken place (*60-63*).

Potential pitfalls and alternative approaches: We have had extensive experience with all of the above assays on coronavirus spike proteins, including in-depth protease sensitivity assay and negative-stain EM assay (*60*). Therefore, we do not expect any major technical problems. As an alternative, ectopic overexpression, siRNA knockdowns or CRISPR gene ablation of target proteases can be performed in permissive cells, allowing us to evaluate the role of proteases on live SARS2, RaTG13 and PaDCoV entry and replication.

E.2.2 Subaim 2 – EM studies of the Conformational Changes of SARS2 and RaTG13 spikes

Hypothesis: (i) In the presence of the triggers identified in subaim 1, SARS2, RaTG13 and 2003 SARS-CoV spikes transition from pre-fusion to post-fusion structures; (ii) comparison of the tertiary structures of these prefusion and postfusion structures will inform cell entry mechanisms and identify potential targets for antivirals.

Rationale: In this subaim, we aim to visualize the conformational changes of SARS2 and RaTG13 spikes using negative-stain EM and cryo-EM. Furthermore, we will determine the tertiary structures of these spikes in different conformations. Although the cryo-EM structures of pre-fusion SARS2 spike alone or in complex of hACE2 have been determined (*33, 34, 64*), the cryo-EM structure of pre-fusion RaTG13 spike as well as the post-fusion conformation of the SARS2 and RaTG13 spikes are unknown.

Experimental designs and expected results:

(i) We aim to directly view the final structural changes of SARS-CoV-2 and RaTG13 spikes. To this end, we will express and purify SARS2 and RaTG13 spike trimers. To prepare these spike trimers in the pre-fusion state, we will remove the C-terminal transmembrane anchor and intracellular tail of these spikes and replace them with a GCN4 trimerization tag and a His_6 tag. We will express these spike trimers in either insect cells or mammalian cells (we have extensive experience with both protein expression systems). Finally, we will perform negative-stain EM on the SARS2 and RaTG13 spikes in the presence of triggers as identified above. We expect that after being treated with the triggers (e.g., one or more of the protease activators plus hACE2 binding), spike trimers will transition to their post-fusion conformation (i.e., rod-like structures).

(ii) We will determine the cryo-EM structures of RaTG13 spike proteins. To this end, we will collect cryo-EM data on the sample and calculate the cryo-EM density map. We will build the structural model of RaTG13 spikes using the structures of SARS2 spike as the template. We expect that the overall structure of RaTG13 spike will be similar to those of SARS-CoV and SARS2 spikes, but there may be critical differences in RBD conformations and inter-subunit packing that may be related to the functions of the spikes.

(iii) We will determine the cryo-EM structures of pre-fusion RaTG13 spike trimer in complex of hACE2 or bACE2. We will purify the complexes of RaTG13 spike trimer and hACE2 or bACE2, and determine their cryo-EM structures. The structures will reveal detailed interactions between RaTG13 spike and hACE2 or bACE2, which will cross validate the crystal structures of RaTG13 RBD/ACE2 complexes, as determined in Aim 1.

(iv) We will determine the cryo-EM structures of SARS2 and RaTG13 spikes in the post-fusion conformation. First, we will prepare the S2' fragment from the full-length spike ectodomain. To this end, the SARS2 and RaTG13 spikes will be treated with the triggers identified in subaim 1. Then we will purify S2' fragment in the

post-fusion conformation from the rest of the spike ectodomains using gel filtration chromatography. Second, we will directly express and purify SARS2 and RaTG13 S2 or S2' fragment in insect or mammalian cells. Based on previous research (*65*), MHV S2 fragment was in the post-fusion conformation if expressed by themselves. After obtaining the S2 or S2' fragment in the post-fusion conformation, we will determine their cryo-EM structures using the structure of MHV S2 fragment as the template. These post-fusion spikes may reveal new drug targets (e.g., pockets in the post-fusion structure).

Potential pitfalls and alternative approaches: As we have determined the cryo-EM structures of several CoV spike proteins either alone or in complex with receptor (*7, 60, 66*), we do not expect any major technical problems. If the physiological triggers do not cause 100% (or majority) of the pre-fusion molecules to transition to the post-fusion conformation, we can use low concentrations of urea (which is a denaturant) to facilitate the process. We have shown that urea can trigger 100% of MHV spikes to transition to the post-fusion conformation (*60*). Urea is removed by dialysis before subjecting the protein samples for cryo-EM experiments.

E.2.3 Subaim 3 – Differential processing of SARS2 and RaTG13 spikes by human and bat proteases. Hypothesis: In addition to receptor binding, intrinsic differences between human and bat cellular proteases function as a critical cross-species barrier for the potential bat-to-human transmission of SARS2.

Rationale: In previous studies (*6, 42*), we showed that MERS-CoV spike, but not HKU4 spike, can be processed by human PPCs and human lysosomal proteases. However, HKU4 spike can be processed by the corresponding bat cellular proteases and proteases are key for virus recovery (*5*). Thus, intrinsic differences between human and bat cellular proteases form a critical cross-species barrier for group 2C MERS-like CoV transmission from bats to humans. We further showed that different lysosomal proteases are a critical determinant of coronavirus tropism (*4*). In this subaim, we will examine the functional differences between the bat and human cellular proteases that lead to potential differential processing of SARS2 and RaTG13 spikes.

Experimental designs and expected results: (**i**) We will investigate whether bat furin can process SARS2 spike. First, we will express SARS2 spike on the cell surface and detect its cleavage state using Western blot. Second, we will package SARS2 pseudoviruses and detect its cleavage state. Third, we will express and purify recombinant bat furin and examine its capacity to cleave pre-fusion SARS2 spikes. Based on our previous research (*4, 6, 42*), we expect that bat and human furin will differentially process SARS2 spikes.

(**ii**) We will also investigate whether bat cathepsins will process SARS2 spikes. First, we will compare whether SARS2 pseudoviruses enter bat cells more efficiently than human cells, after excluding the impact of other cellular proteases on pseudovirus entry. We have listed detailed descriptions on how to exclude these other cellular proteases in earlier publications (*4*). Second, we will express and purify bat cathepsins (whose human counterparts trigger the conformational changes of SARS2 spike and pseudotype cell entry) and treat SARS2 spike with the recombinant bat cathepsins. Based on our previous research (*4, 6, 42*), we expect that bat cathepsins are more efficient than human cathepsins in processing SARS2 and RaTG13 spikes.

(**iii**) We will identify the cleavage sites in SARS2 and RaTG13 spikes targeted by human and bat cathepsins. Using the above cathepsin-cleaved fragments from recombinant SARS2 and RaTG13 spikes, we will perform N-terminal protein sequencing to identify the proteolysis sites of the human and bat cathepsins. We will compare the sequences of these cleavages sites in SARS2 and RaTG13 spikes, and identify critical sequence differences between bat cathepsin sites and human cathepsin sites that may account for the potentially different processing of these spikes by human and bat cathepsins. We expect that human and bat cathepsins will cleave SARS2 and RaTG13 spikes at different sites. Because the signature cleavage sites for human cathepsins are still not well defined and the signature cleavage sites for bat cathepsins are unclear, these experiments will reveal the activities of human and bat cathepsins as determinants of the cross-species transmission for SARS2. The information will reveal whether evolution has taken place from the RaTG13 spike to SARS2 spike, making the pandemic spike more sensitive to human cathepsin cleavages during putative bat-to-human transmission.

(**iv**) We will introduce targeted mutations that ablate or introduce human or bat cathepsin cleavage sites, introduce (RaTG13) or ablate (SARS2) proprotein convertase furin cleavage sites (**Fig 6**), or ablate TMPRS2 cleavage sites from the SARS2 and RaTG13 spike glycoprotein (59), and recover recombinant viruses using reverse genetics in vitro. Virus growth will be evaluated in Calu3 and Vero cells in vitro as well as in primary human airway epithelium and other primary human lung cells in **Aim 3**. We will also evaluate virus growth in Vero cells over-expressing furin or TMPRSS2, available in our laboratory.

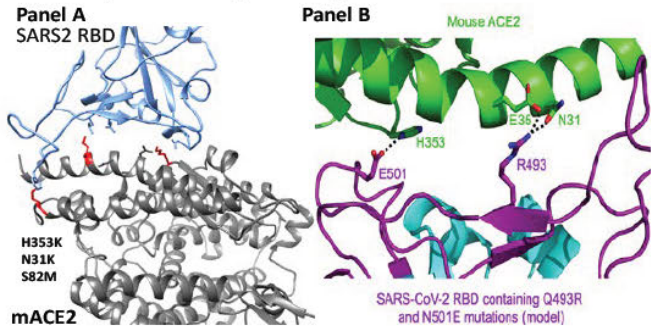
Expected results, potential pitfalls and alternative approaches: Our focus is the cellular proteases that are main players in processing CoV spikes and enhance CoV entry (e.g. furin, cathepsin L, cathepsin B). However, if they do not process the SARS2 spike, we will focus on other PPCs and lysosomal proteases (*4*).

E3. Aim 3. Pathogenesis of SARS2 and SARS2-like CoV in Mouse Models of Human Disease.
Rationale: The overall goal is to identify key interaction networks that regulate SARS2 and zoonotic SARS2-like CoV RaTG13 and PaDCoV cross species transmission and pathogenesis in mice, to recapitulate a robust lethal mouse model of human disease and identify host determinants that regulate primary human cell infection. Our underlying hypothesis is that microvariation with the clade 1 SARS2-like CoV S glycoprotein regulates alternative mammalian ACE2 receptor binding profiles and protease usage, altering human tissue tropism for different human primary lung cells in vitro and in vivo pathogenesis in mice. Given the importance, parallel approaches using CRISPR gene editing and SARS2 S-RBD structure-guided mouse adapted virus design are proposed.

E3a. Identifying Mutation Sets that Promote Clade I Group 2b SARS2 Entry in vitro and in vivo. i) *Introduction of Human Residues in mACE2 Receptors*. Our modeling has predicted key interaction residues in the mACE2 that impede SARS2 binding (*29*). Using our previously described approach to map key species specific interaction residues that govern mouse and human MERS dipeptidyl peptidase receptor usage (*67, 68*), we will systematically introduce 1 to 3 human ACE2 interface residues into the mACE2 receptor (**Fig 7a**). Using ectopic over-expression in non-permissive cells, we will evaluate clade 1 SARS2-like CoV entry and growth by plaque assay and nLUC or GFP indicator gene expression. We anticipate that changes at mACE2 residues 31 and 353 will promote efficient SARS2 entry replication and expression in vitro. As the PaDCoV RBD is near identical with SARS2, its replication is anticipated to follow similar phenotypes. However, we anticipate that the novel variation patterns in SARS2-like bat CoV RaTG13 RBD will efficiently program virus replication in unaltered mACE2, hACE2 or chimeric humanized mACE2 receptors. Alternatively, the hACE2 exons that support SARS2 entry will be introduced by CRISPR gene editing into the mACE2 receptor for in vitro stdies as well. The goal is to identify a candidate chimeric h/mACE2 for knockin into a mouse which program improved SARS2 and other SARS2-like CoV entry and replication (*54*). We note that transgenic knockin h/mACE2 mice provide pathogenesis and countermeasure studies without having to resort to SARS2 MA viruses.

E3b. Developing CRISPR-Gene Edited Chimeric mACE2 Transgenic Mice. *i) CRISPR Gene Edited Transgenic Mice*. We have previously used CRISPR-CAS9 gene editing to produce a robust mouse model for MERS-CoV pathogenesis in vivo (*1*). Using established approaches, we will initially use the CRISPR-CAS9 system to program the use of homologous recombination mediated repair rather than NHEJ repair, thereby increasing the frequency with which specific hDPP4 residues are introduced into the mACE2 gene. We will work with the UNC transgenic mouse facility to produce chimeric mACE2 receptors encoding human residues. Briefly, we will inject the ACE2 targeting crRNA and humanized ACE2 homologous DNA sequence into mouse embryos using the UNC Transgenic Microinjection Facility. Using standard PCR approaches, we will identify MERS DPP4 transgenic mice (*1*). We will screen founder mice for human mACE2 mutations (m/hACE2 mice) and backcross them at least twice on C57BL6 mice to remove spurious mutations. Homozygous breeders will then be established to expand the final homozygous knockin m/hACE2 transgenic mouse colony.

Figure 7. Mouse Adaptive Mutation Sets. Panel A: Human mutations in mACE2 Predicted to Enhance SARS2 mACE2 usage in vitro and in vivo. Blue= SARS2 RBD, Red= human ACE2 residues in mACE2. Panel B: Mutations in the SARS2 RBD predicted to use the mACE2 receptor for docking and entry in vitro and in vivo.



ii) In vivo Pathogenesis Studies. Wildtype SARS-CoV replicates but does not produce clinical disease in standard laboratory mice (*69*). In contrast, SARS2 does not replicate in standard laboratory mice. Given its similar S-RBD, it seems likely that PaDCoV will not replicate in mice while the variation within the RaTG13 S-RBD, which encodes some similar residues as SARS-CoV, is less clear but may program low-level replication in standard laboratory mice. Briefly, groups of 15 female BALB/c and 18 female m/hACE2 transgenic mice/each will be infected with 1 x 10^5 SARS2, rRaTG13 or PaDCoV intransally, and 5 controls inoculated with PBS. Subgroups of 6 animals/ will be sacrificed on days 0 (controls), or infected animals at 2, 4, 7 days post-infection. We will also test the 2003 wildtype SARS-CoV under identical treatment conditions. Clinical disease will be followed by weight loss, morbidity, mortality and plethysmography to evaluate respiratory function over the disease course. We will collect lungs, brain and liver and RNA at each designated time point, and evaluate virus titers by plaque assay. Tissues will be scored for pathologic changes and viral antigen in the lung using immunohistochemistry and an American Thoracic Society Lung Injury Scoring

system, as described by our group (*70*). Virus tropism will be determined by immunohistochemistry using antisera targeting the SARS2 nucleocapsid protein as previously described by our group (*71*). If serious disease is noted in transgenic animals (e.g., acute lung injury, significant weight loss (>20%) or death), flow cytometry will evaluate changes in leukocyte populations on days 4 and 7 post-infection as described previously (72). If replication models are observed following infection of transgenic h/mACE2 mice, we will independently passage each virus in two h/mACE2 transgenic mice at 2 day intervals for 15 serial passages, selecting for pathogenic variants. These viruses will be isolated, sequenced and rederived using the SARS2, PaDCoV and RaTG13 molecule clones and pathogenic outcomes reevaluated as described above.

E3c. Impact of in vitro evolved and structure-guided mutations on viral pathogenesis. ***Rational***: A goal is to make a mouse adapted version of SARS2. We recognize the possibility of two distinct animal models being available for our downstream studies, allowing us to evaluate targeted mutations in the background either of a mouse-adapted strain or in humanized knockin mice. If mouse adapted viruses are available, we will prioritize GOF experiments in these genetic backbones, as these viruses will likely poorly recognize hACE2 receptors and hence minimize P3C0 concerns. In contrast, LOF mutants will be constructed in wildtype SARS2 and RaTG13 genetic backgrounds, as these will also minimize P3C0 concerns and be tested in chimeric h/mACE2 mice.

i) SARS2 Mouse adapted Mutations. Guided by studies by the Li laboratory in Aim 1, we will introduce mutations that are predicted to promote mACE2 binding by the SARS2 RBD (*55*)(PMC3308800). Using the atomic level structures in **Fig 2**, we predicted several changes in the SARS2 RBD, like those shown in **Fig 7b**. Virus growth will be characterized on nonpermissive cells expressing mACE2 receptors. If replication competent, groups of six female mice/timepoint will be inoculated with $1x10^3$ to 1×10^5 SARS2 MA recombinant virus, intranasally, and followed for 7-10 days to determine maximal pathogenic doses. Using a lethal dose, 10 wk old females (n-40) will then be infected and subgroups of eight animals/each will be sacrificed on days 0, 2, 4, 7 and 15 post-infection. The experiment will also be repeated with male mice. Clinical disease will be followed by percent weight loss, morbidity, mortality and respiratory function by plethysmography (1,69). We will collect lungs, liver, spleen, draining lymph nodes and evaluate virus titers by plaque assay. Tissues will be scored for pathologic changes and viral antigen in the lung using immunohistochemistry (71). In the lung, we will use an American Thoracic Society Lung Injury Scoring system that creates an aggregate score for the following phenotypes: neutrophils in the alveolar and interstitial space, hyaline membranes, proteinaceous debris filling the air spaces, and alveolar septal thickening, as described by our group (*70*). We note that mouse adapted SARS2 strains provide the global research community with opportunities to use standard laboratory mice for countermeasure development, speeding bench to bedside translation of products, while also revealing novel mutation networks that regulate ACE2 usage.

LOF and GOF Mutations	Target Function	Recombinant Virus Backbone
Furin Cleavage site KO	LOF	SARS2 (Aim 2)
Furin Cleavage Site	GOF	RaTG13 (Aim 2)
TMPRSS2 KO	LOF	SARS2 (Aim 2)
Cathepsin Site KO	LOF	SARS2 or RaTG13
Bat Cleavage Site	GOF	SARS2 (Aim 2)
Human Cleavage Site	GOF	RaTG13 (Aim 2)
Targeted ACE2 Mutations	LOF	SARS2 (Aim 1)
Targeted ACE2 Mutations	GOF	RaTG13 (Aim 1)

Table 2. Other Representative Priority S glycoprotein Mutations for in vitro Replication in Primary Human Cells and for in vivo Pathogenesis Studies. We target mutant SARS2 or RaTG13 RBD recombinant viruses that demonstrate altered usage of different mammalian ACE2 receptors or growth in vitro (Aim 1). In parallel, we target interesting cleavage site mutant recombinant viruses that display altered in vitro growth phenotypes in Vero or Calu3 cells or in furin or TMPRSS2 expressing Vero cells in Aim 2.

ii) In vitro Studies. We will work closely with Dr. Fangs laboratory to introduce targeted S-RBD and protease site mutations that were identified to give interesting phenotypes (e.g., altered ACE2 receptor usage, altered protease usage, etc.) in Aims 1 and 2 into the SARS2 and RaTG13 molecular clones and recover recombinant viruses in the presence or absence of nLUC indicator genes. A variety of priority mutations will be targeted as guided by outcomes in Aims 1 and 2, as shown in **Table 2 and Fig 7b**. Recombinant viruses will have be isolated and evaluated for growth on nonpermissive cells expressing the human, bat, civet, mouse and other select mouse ACE2 receptors or in the presence/absence of various ectopically expressed proteases (Aims 1 and 2). In this subaim, we will also evaluate virus growth on primary human lung cells in vitro using primary human airway epithelial cells (HAE), small airway epithelium (SAE), type II pneumocytes, lung endothelial cells, and lung fibroblasts purchased from the Cystic Fibrosis Core at the University of North Carolina at Chapel Hill as previously described by our group (*44, 71*). We will evaluate virus growth by nLUC assays and by virus growth over 72 hrs using established techniques in the lab. We anticipate that select S-RBD and protease site mutations will alter in vitro tissue tropism and replication efficiency across some but rarely all of the primary cells in the panel. We anticipate that loss of furin cleavage sites will attenuate virus growth, while gain of furin cleavage sites

may enhance in vitro growth, as evidenced by increased nLUC expression in vitro. The goal is to identify key mutations that regulate the permissiveness of upper and lower airway epithelium to SARS2 and RaTG13 virus entry and replication, potentially identifying changes that contribute to serious lung injury following infection.

ii) In vivo Mouse Studies. We will also evaluate the impact of targeted S-RBD mutations either using other the mouse-adapted viruses or the transgenic h/mACE2 mice, including control C57BL/6 mice as described above. Moreover, we will evaluate the function of the novel proprotein convertase (e.g., furin) and TMPRSS2 protease cleavage site mutations on SARS2 and RaTG13 in vivo pathogenesis using transgenic h/mACE2 mice. We recognize that the introduction of select SARS2 RBD-ACE2 enhancing and novel furin cleavage site mutations into the wildtype RaTG13 background may constitute a P3C0 experiment, requiring approval from appropriate NIH committees. Briefly, recombinant viruses will be intranasally inoculated into standard laboratory or h/mACE2 transgenic mice (18 female mice and 6 controls) sacrificed in groups of six on days 2, 4 and 7 post infection for virus titer, immunohistochemistry and pathologic evaluation. Clinical disease outcomes are scored daily as noted above in **Section E3c**. We anticipate that cleavage site mutants (e.g., furin ko) will display altered virus tropisms in mice in vivo, attenuating in vivo pathogenic outcomes. Alternatively, we anticipate that introducing the furin cleavage site into RaTG13 may expand in vivo tropism for epithelial cells in the airway epithelium and parenchyma, and may enhance in vivo pathogenesis.

E3d. Expected Outcomes/Alternative Approaches. Given our previous success (*1*), we do not anticipate problems with isolating gene edited chimeric h/mACE2 receptors that program efficient replication and pathogenesis of SARS2 and RaTG13 infection in vivo. However, in vivo passage may be required to enhance viral pathogenesis, as previously shown for MERS-CoV and SARS-CoV (*1, 47*). If necessary, virus will be serially passaged at two-day intervals in vitro or in vivo for 15 passages in duplicate, and plaque purified viruses sequenced. Using two novel experimentally evolved variants, we will inoculate groups of six pilot mice/each to identify viruses that replicate efficiently and produce clinical disease (e.g., weight loss). We anticipate that recombinant viruses encoding structurally and biochemically defined S-RBD and protease site mutations will alter virus entry, pathogenesis and tissue tropism in primary cells in vitro and pathogenesis in vivo, identifying informative changes that regulate cross species movement of high risk human epidemic and bat SARS2 CoV. As appropriate, we can also study PaDCoV replication and pathogenesis in h/mACE2 knockin mice in vivo. Depending on outcomes, we will also evaluate SARS2 MA pathogenesis in 1 year aged mice, as aged mice oftentimes display increased ARDS-like phenotypes after group 2b SARS-like CoV infection (14,43).

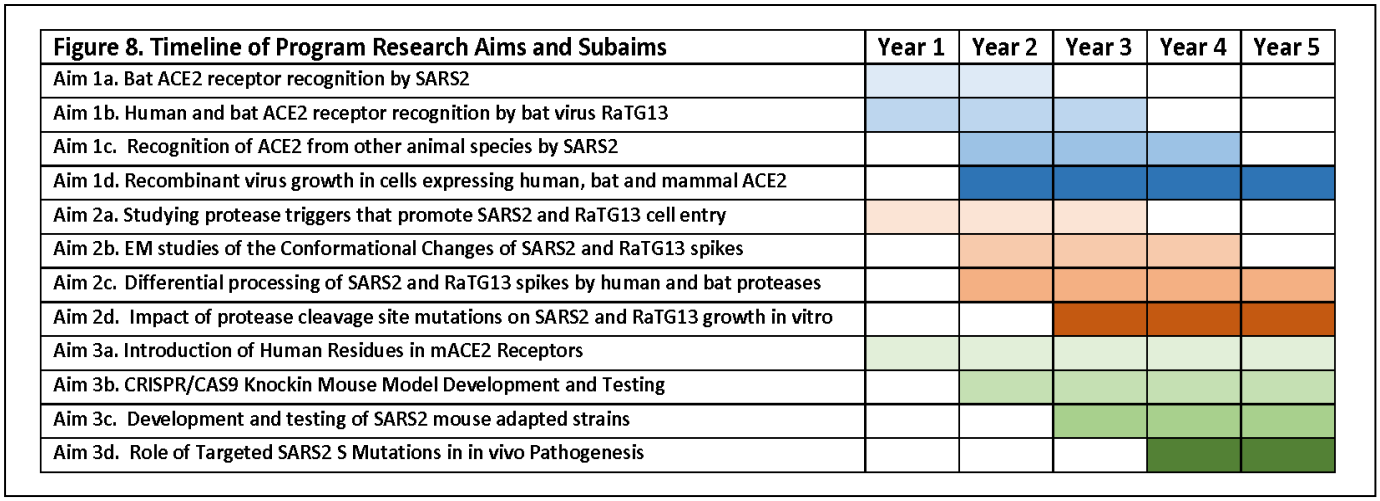


Figure 8. Timeline of Program Research Aims and Subaims	Year 1	Year 2	Year 3	Year 4	Year 5
Aim 1a. Bat ACE2 receptor recognition by SARS2					
Aim 1b. Human and bat ACE2 receptor recognition by bat virus RaTG13					
Aim 1c. Recognition of ACE2 from other animal species by SARS2					
Aim 1d. Recombinant virus growth in cells expressing human, bat and mammal ACE2					
Aim 2a. Studying protease triggers that promote SARS2 and RaTG13 cell entry					
Aim 2b. EM studies of the Conformational Changes of SARS2 and RaTG13 spikes					
Aim 2c. Differential processing of SARS2 and RaTG13 spikes by human and bat proteases					
Aim 2d. Impact of protease cleavage site mutations on SARS2 and RaTG13 growth in vitro					
Aim 3a. Introduction of Human Residues in mACE2 Receptors					
Aim 3b. CRISPR/CAS9 Knockin Mouse Model Development and Testing					
Aim 3c. Development and testing of SARS2 mouse adapted strains					
Aim 3d. Role of Targeted SARS2 S Mutations in in vivo Pathogenesis					

E4. Summary and Timeline. We will initiate all aims simultaneously (**Fig 8**). However, biochemical and atomic structures will influence the start periods for downstream subaims in Aims 1-3. We note that the research program will provide an array of animal models of human disease, molecular clones, and recombinant viruses encoding indicator genes, atomic level structures and detailed insights into the molecular and structural basis for SARS2-like bat coronavirus receptor usage and cross species transmission mechanisms. Consequently, the program bridges basic science investigations and translates these discoveries into key resource opportunities that support countermeasure development, which is critical for global health preparedness and response to the COVID-19 pandemic.