PLOS PATHOGENS

COVID-19 pandemic: Insights into structure, function, and hACE2 receptor recognition by SARS-CoV-2

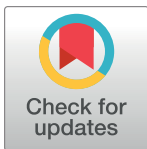**Anshumali Mittal[1]*, Kavyashree Manjunath[2], Rajesh Kumar Ranjan[3], Sandeep Kaushik[4], Sujeet Kumar[5], Vikash Verma[6]***

1 Sriram Sameeksha, Jalahalli, Bangalore, India, **2** Suraj Darshan, Bangalore, Karnataka, India, **3** Department of Environmental Science, Central University of South Bihar, Gaya, India, **4** EduBooster Learnings, Vasundhara, Ghaziabad, India, **5** Centre for Proteomics and Drug Discovery, Amity Institute of Biotechnology, Amity University, Maharashtra, India, **6** Biology Department, University of Massachusetts, Amherst, Massachusetts, United States of America

* mittalans@gmail.com (AM); vverma@umass.edu (VV)

Abstract

Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) is a newly emerging, highly transmissible, and pathogenic coronavirus in humans that has caused global public health emergencies and economic crises. To date, millions of infections and thousands of deaths have been reported worldwide, and the numbers continue to rise. Currently, there is no specific drug or vaccine against this deadly virus; therefore, there is a pressing need to understand the mechanism(s) through which this virus enters the host cell. Viral entry into the host cell is a multistep process in which SARS-CoV-2 utilizes the receptor-binding domain (RBD) of the spike (S) glycoprotein to recognize angiotensin-converting enzyme 2 (ACE2) receptors on the human cells; this initiates host-cell entry by promoting viral–host cell membrane fusion through large-scale conformational changes in the S protein. Receptor recognition and fusion are critical and essential steps of viral infections and are key determinants of the viral host range and cross-species transmission. In this review, we summarize the current knowledge on the origin and evolution of SARS-CoV-2 and the roles of key viral factors. We discuss the structure of RNA-dependent RNA polymerase (RdRp) of SARS-CoV-2 and its significance in drug discovery and explain the receptor recognition mechanisms of coronaviruses. Further, we provide a comparative analysis of the SARS-CoV and SARS-CoV-2 S proteins and their receptor-binding specificity and discuss the differences in their antigenicity based on biophysical and structural characteristics.

Introduction

Before 2003, only 2 human coronaviruses—Human Coronavirus (HCoV)-229E and HCo-V-OC43, causing mild illness—were known [1,2,3]. However, the emergence of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle East Respiratory Syndrome Coronavirus (MERS-CoV) changed the view worldwide because coronaviruses can cause life-threatening infections [4,5,6]. The ongoing pandemic of a novel strain of coronavirus, Severe

Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2), is posing an unforeseen public health and economic threats worldwide. As of June 27, 2020, SARS-CoV-2 has infected more than 9.65 million people, with 491,115 deaths reported from 215 countries and territories [7], of which there are 2,407,590 confirmed cases of COVID-19 and 124,161 deaths in the United States of America alone [8]. Recombination, mutator alleles, and mutational robustness are some of the evolutionary mechanisms [9] that make coronaviruses capable of expanding their host ranges, including humans. Therefore, understanding the virology of the coronaviruses at a structural level is of utmost importance because the health threats from these zoonotic viruses are constant and long-term.

Coronaviruses are large, enveloped, positive-stranded RNA viruses responsible for infecting a wide variety of mammalian and avian species [10]. These viruses contain spike-like projections of glycoproteins on their surface, which appear like a crown under the electron microscope; hence, they are referred to as coronaviruses. The coronavirus genome encodes several structural and nonstructural proteins. The structural proteins are responsible for host infection [11], membrane fusion [12], viral assembly [13], morphogenesis, and release of virus particles [14], among other functions, and the nonstructural proteins (nsps) facilitate viral replication and transcription [15,16]. The membrane (M), the envelope (E), and the spike protein (S) make up the structural proteins and are associated with the envelope. Among these structural proteins, the trimeric S proteins protrude from the virus envelope and are the key machinery that facilitates virus entry into the host cell [10,17].

The S proteins are clove-shaped, type-I transmembrane proteins and have 3 segments: a large ectodomain, a single-pass transmembrane, and an intracellular tail. The ectodomain of S proteins consist of the S1 subunit, containing a receptor-binding domain (RBD), and the membrane-fusion subunit (S2). The host-cell receptor recognition by the RBDs on S proteins is the initial step of viral infection, and the binding interactions between the coronavirus spike and its receptor is one of the most critical factors for host range and cross-species transmission. Human coronaviruses recognize a variety of host receptors; specifically, HCoV-229E recognizes human aminopeptidase N (hAPN) [18], MERS-CoV binds dipeptidyl peptidase-4 (DPP4) [19], HCoV-OC43 and HCoV-HKU1 bind certain types of O-acetylated sialic acid [20], and HCoV-NL63 and SARS-CoV recognize angiotensin-converting enzyme 2 (ACE2) [21,22]. Recent structures, along with functional studies, have suggested that the SARS-CoV-2 S proteins utilize ACE2 and Transmembrane Serine Protease 2 (TMPRSS2) for host-cell entry, which are very similar to the mechanisms exploited by SARS-CoV [23]. See the "Structure, function, antigenicity, and hACE2 receptor recognition by the SARS-CoV-2 S glycoprotein" section of this review for detailed information on the mechanism of coronavirus cell entry mediated by the viral S glycoproteins. The S proteins, common among all coronaviruses, are a major target for eliciting antibodies; therefore, structural and molecular details of S protein and its interactions with cognate receptors would be vital in developing vaccines and antiviral drugs against SARS-CoV-2.

In this review, we discuss the coronavirus classification, details of SARS-CoV-2 emergence, morphology, and key virulence factors. We specifically explain the structure of RNA-dependent RNA polymerase of SARS-CoV-2 and its significance in drug discovery. Further, the structure, function, and antigenicity of S glycoproteins and their interactions with human ACE2 (hACE2) receptor are discussed.

Emergence of SARS-CoV and SARS-CoV-2

In November 2002, SARS began spreading from the Guangdong province of Southern China, but its reservoir was unknown. In the past, Nipah and Hendra, both zoonotic viruses,

originated from bats, and this motivated researchers to find whether bats are the natural reservoirs of SARS-CoV [24,25]. In 2005, 2 research groups independently reported that bats (horseshoe bats in particular) are the natural host of genetically diverse coronaviruses and closely related to those responsible for the SARS outbreak [26,27]. These viruses were termed SARS-like coronaviruses, and they displayed considerable genetic similarities to SARS-CoV isolated from humans or civets. This suggested that the virus responsible for SARS outbreak was a member of the SARS-like coronaviruses group [26]. In Saudi Arabia, MERS-CoV emerged in 2012, when humans were infected through direct or indirect contacts with infected dromedary camels. However, genome analysis suggested that MERS-CoV might have also originated in bats and was transmitted to camels in the distant past [28] (Fig 1).

In December 2019, patients with severe pneumonia cases of unknown cause were reported in Wuhan, China, and a novel coronavirus strain was detected from the lower respiratory tract of 4 patients [29]. Viruses were isolated from these clinical samples, and their genomes were analyzed by deep sequencing [30,31,32]. Phylogenetic analysis of 2019-novel coronavirus (2019-nCoV) genomes and other coronaviruses were used to establish the evolutionary history and infection sources. Interestingly, this indicated that 2019-nCoV (GenBank: MN908947.3) shares about 96% nucleotide sequence identity to bat coronavirus RaTG13 (GenBank: MN996532.1), with 79.5% and 55% identity to SARS-CoV BJ01 (GenBank: AY278488.2) and MERS-CoV HCoV-EMC (GenBank: MH454272.1), respectively, and belongs to the same family of viruses that caused SARS and MERS (Fig 2 and S1 Fig). This suggests that bats are possibly the hosts of 2019-nCoV origin, and it might have been transmitted either directly from bats or through an unknown intermediate host to infect humans [29,33,34,35]. Despite high sequence similarities, a few notable and conserved variations arose in 2019-nCoV genomes that were not previously seen in betacoronaviruses. These notable features, which establish this virus as different from SARS-CoV and SARS-like coronaviruses, are (i) multiple mutations in the RBDs of S protein that may interact with ACE2 receptor, (ii) a polybasic furin-like protease site (RRAR/S) at the boundary of S1/S2 subunits rather than the single arginine observed in SARS-CoV, and (iii) the addition of 3 predicted O-linked glycans flanking the protease site [36,37]. Of note, a furin-like protease site is a signature of several highly pathogenic avian influenza viruses and pathogenic Newcastle disease virus [38,39].

Originally, this virus was called 2019-nCoV, but later the International Committee on Taxonomy of Viruses on February 11, 2020 officially named it SARS-CoV-2 because of its genetic similarity to SARS-CoV [33]. SARS-CoV-2 causes respiratory illness, and WHO named this illness COVID-19. It is contagious, primarily transmitted among people through respiratory droplets and contact routes [40,41], and more than 13 million COVID-19 cases are confirmed worldwide (as of July 14, 2020). Initially, WHO declared the COVID-19 outbreak a public health emergency of international concern and later confirmed it as a pandemic on March 11, 2020 [42].

Classification of coronaviruses

The coronavirus study group of the International Committee on Taxonomy of Viruses has classified coronaviruses under the family Coronaviridae, subfamily Coronavirinae. Based on genotypic and serological characterization, Coronavirinae is divided into 4 genera: *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus*, and *Deltacoronavirus* [43,44,45,46] (Fig 3A). Only 6 HCoV species that cause human disease were known before December 2019. Four of them cause common cold symptoms in immunocompromised individuals: these are HCoV-229E and HCoV-OC43, first identified in the mid-1960s [1,2,3]; HCoV-NL63, first identified in 2004 [47,48]; and HCoV-HKU1, first identified in 2005 [49]. The other 2 strains, which cause fatal illness, are SARS-CoV, first identified in 2003 [4,5], and MERS-CoV, first identified
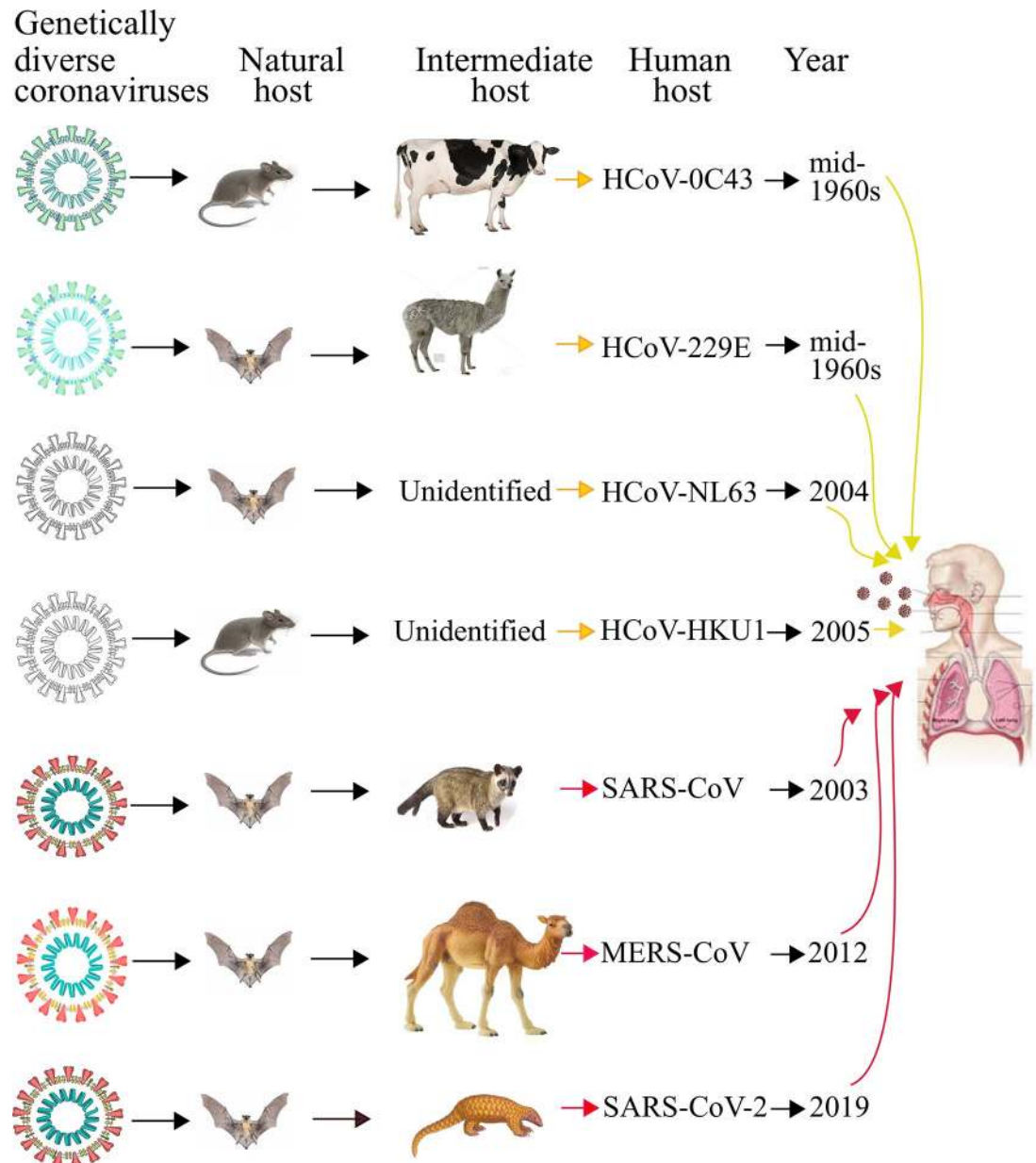
Fig 1. Origin and transmission of pathogenic HCoVs. Yellow and red arrows indicate mild and severe infections in humans, respectively. The figure is inspired from Jie Cui and colleagues [46], and the illustrations of coronaviruses (left) are adapted from "Desiree Ho, Innovative Genomics Institute," available at https://innovativegenomics.org/free-covid-19-illustrations/. HCoV, Human Coronavirus; MERS-CoV, Middle East Respiratory Syndrome Coronavirus; SARS-CoV, Severe Acute Respiratory Syndrome Coronavirus; SARS-CoV-2, SARS-CoV, Severe Acute Respiratory Syndrome Coronavirus-2.

https://doi.org/10.1371/journal.ppat.1008762.g001

in 2012 [6]. SARS-CoV-2 has 96% nucleotide sequence identity to bat coronavirus RaTG13, a SARS-like coronavirus; therefore, it belongs to *Betacoronavirus* genera (Fig 2 and S1 Fig).

Forsters and colleagues [50] performed a phylogenetic network analysis of 160 complete SARS-CoV-2 genomes sampled from across the world to understand the evolution of this virus in humans and infection sources. They sorted these closely related genomes into 3 lineages, namely A, B, and C, based on amino acid changes. Lineage A was named for the original bat coronavirus that caused COVID-19, but surprisingly, it was not the dominant virus type in
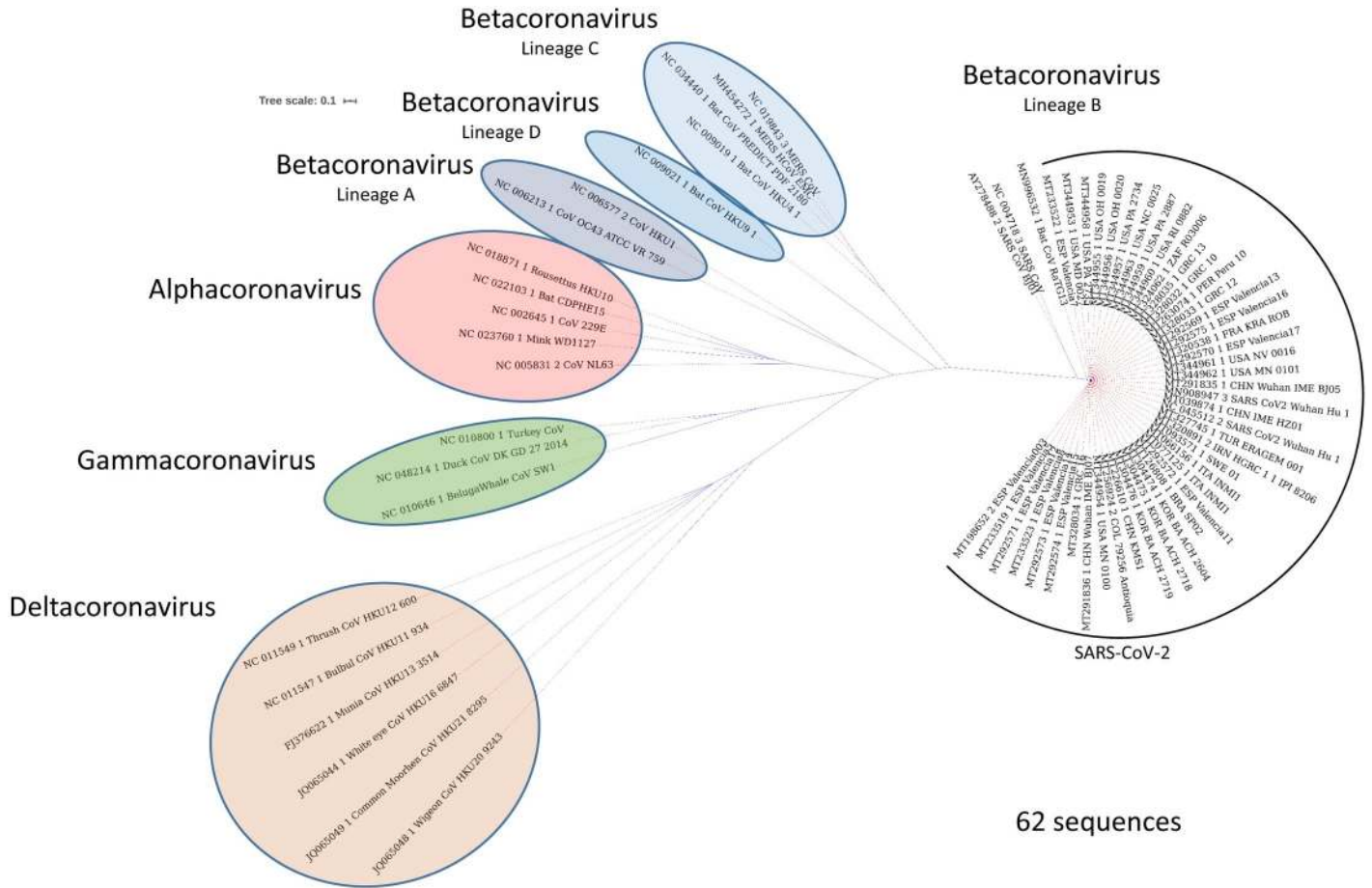
Fig 2. Phylogenetic relationships in the Coronavirinae subfamily. The subfamily is formed by 4 genera: *Alphacoronavirus*, *Betacoronavirus* (lineages A, B, C, and D), *Gammacoronavirus*, and *Deltacoronavirus*. We randomly picked 62 SARS-CoV-2 genome sequences, representing 15 different countries, together with other Coronavirinae subfamily members. The phylogenetic tree was created using NgPhylogeny.fr tool. The analysis indicates that SARS-CoV-2 has a close relationship with bat coronavirus RaTG13 and SARS-CoV; therefore, it is classified as a new member of the lineage B *Betacoronavirus*. SARS-CoV, Severe Acute Respiratory Syndrome Coronavirus; SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus-2.

Wuhan. The A and C types were found largely in the Americas and Europe, respectively, whereas the B type was mostly prevalent in East Asia and had acquired mutations before spreading outside East Asia. The lineage C differs from its parent lineage B by a mutation at amino acid position 26,144 and was prevalent in France, Italy, Sweden, England, California, Brazil, Singapore, Hong Kong, Taiwan, and South Korea but absent from mainland Chinese samples. This kind of phylogenetic classification has potential for accurately tracing the infection routes and will prove helpful in designing treatments and vaccine development [50].

Morphology, genomic structure, and key viral factors of SARS-CoV-2

Coronaviruses are nonsegmented, enveloped viruses with single-stranded RNA (ssRNA) ranging between 26 to 32 kb in length. At this length, the coronavirus genome is the largest among RNA viruses. Electron microscopy (EM) of negative-stained SARS-CoV-2 particles revealed their spherical shape, with the diameter ranging from 60–140 nm and an outer surface studded with distinctive 9- to 12-nm–long spikes that gave virions the appearance of a solar corona [29] (Fig 3B). The observed morphology of SARS-CoV-2 is consistent with other members of the Coronaviridae family.
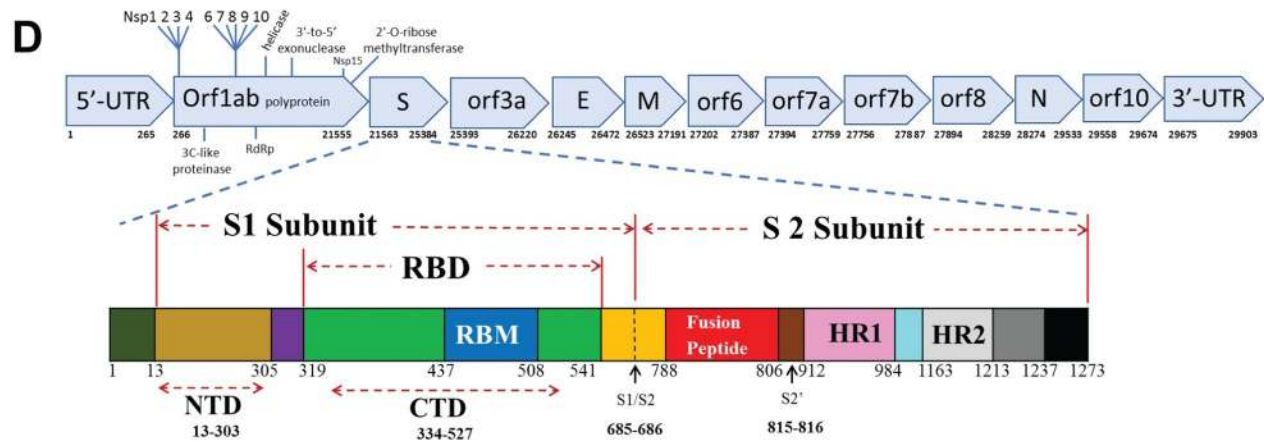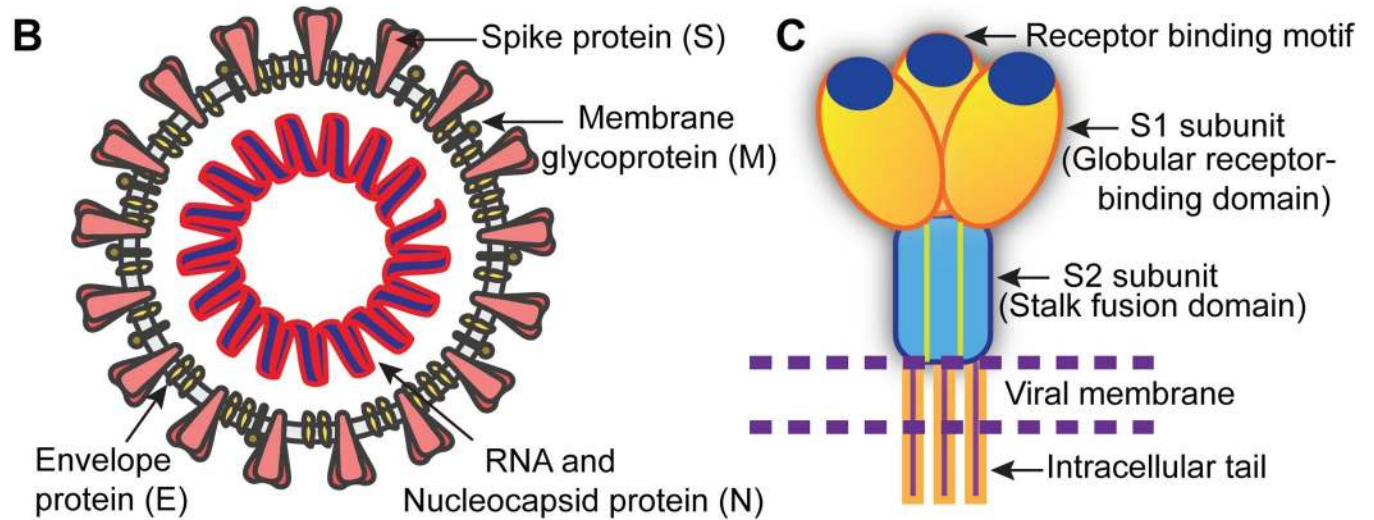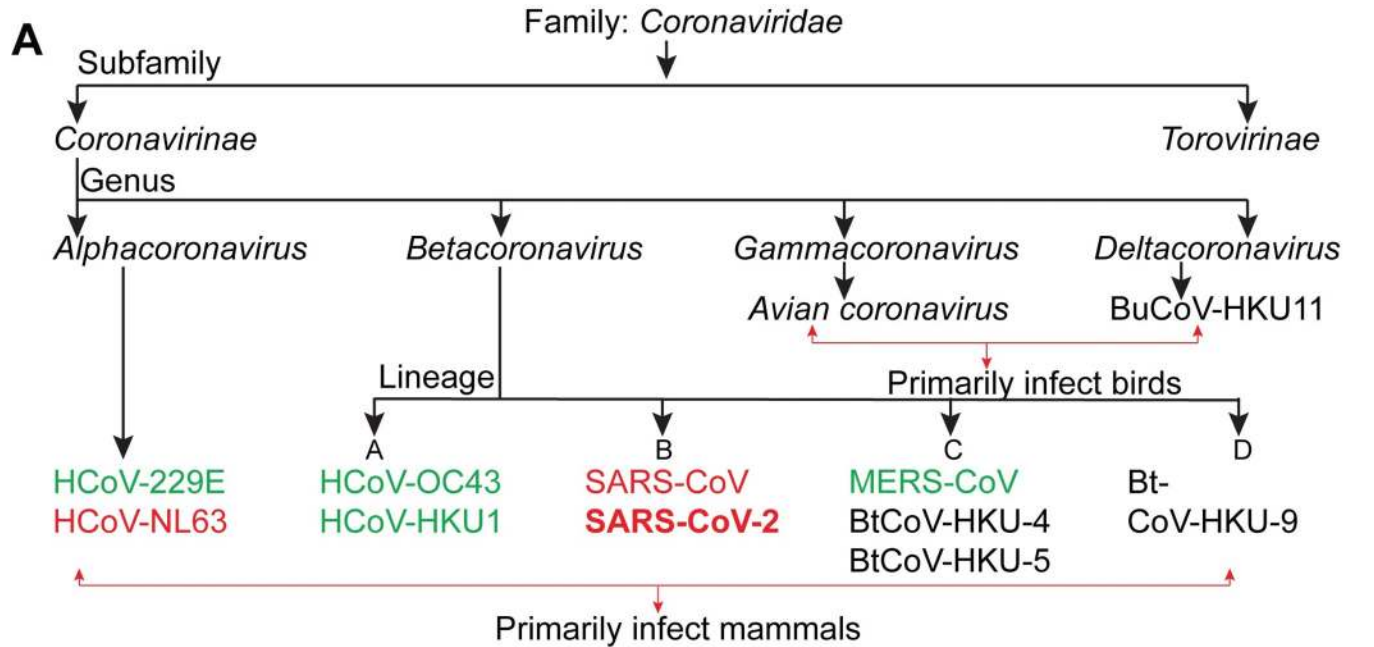
Fig 3. Classification and structure of coronavirus. (A) Classification of coronaviruses: the 7 known HCoVs are shown in green and red. HCoVs in red bind the host receptor ACE2. (B) Schematic of the SARS-CoV-2 structure; the illustration of the virus is adapted from "Desiree Ho, Innovative Genomics Institute," available at https://innovativegenomics.org/free-covid-19-illustrations/. (C) Cartoon depicts key features and the trimeric structure of the SARS-CoV-2 S protein. (D) Schematic of SARS-CoV-2 genome (top) and S protein (bottom); annotations are adapted from NCBI (NC_045762.2) and Expasy (https://covid-19.uniprot.org/uniprotkb/P0DTC2), respectively. ACE2, angiotensin-converting enzyme 2; CTD, C-terminal domain; E, envelope; HCoV, Human Coronavirus; HR1/2, heptad repeat 1/2; M, membrane; N, nucleocapsid; Nsp, nonstructural protein; NTD, N-terminal domain; orf, open reading frame; RBD, receptor-binding domain; RBM, receptor-binding motif; RdRp, RNA-dependent RNA polymerase; S protein, spike protein; SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus-2; UTR, untranslated region.

https://doi.org/10.1371/journal.ppat.1008762.g003

SARS-CoV-2 Wuhan-Hu-1 isolate (GenBank: MN908947.3) was among the first complete genome of the virus strains to be sequenced and comprises a 29,903-bp–long RNA. It is 5′-capped and 3′-polyadenylated, consists of 2 flanking untranslated regions (UTRs), and contains several open reading frames (orfs) that encode multiple proteins. The genome is arranged in the order of a noncoding 5′-UTR–replicase genes (orf1ab)–structural proteins (S, E, M, and N) and accessory proteins–noncoding 3′-UTR [51] (Fig 3D). Notably, it lacks the hemagglutinin-esterase gene, which is a common feature of lineage A *Betacoronaviruses* [31]. The orf1a/b, located at the 5′-end of the genome, is the largest orf, and it encodes 15 nsps (nsp1–10 and nsp12–nsp16) [52]. Briefly, the orf1a/b has overlapping orfs and produces 2 polypeptides, pp1a and pp1ab, because of ribosomal frameshifting. The virus genome encodes 2 cysteine proteases, a papain-like protease (PL2pro) or nsp3 and a 3C-like protease (3CLpro) or nsp5. These proteases cleave pp1a and pp1ab polypeptides into 15 nsps. Specifically, PL2pro is responsible for cleaving between nsp1|2, nsp2|3, and nsp3|4 sites, and the 3CLpro cleaves at the LQ↓SAG sites to produce nsp4 through nsp16 [31,53]. RNA-dependent RNA polymerase (nsp12) in complex with nsp7, nsp8, helicase (nsps13), and exonuclease (nsp14) are critical enzymes among these nsps responsible for the transcription and viral RNA replication.

The 3′-terminus of the SARS-CoV-2 genome contains 4 structural proteins that are responsible for virus–host cell receptor binding, virion assembly, morphogenesis, and release of virus particles from the host cell. The E protein of SARS-CoV-2 is the smallest of all structural proteins found in the viral membrane and localizes to the endoplasmic reticulum and Golgi complex in the host cells [54]. The E protein, along with M and N, is known to facilitate virus-like particle formation [14]. The M glycoprotein is a transmembrane protein located in the viral membrane and is the most abundant structural protein in a virion, almost a hundred times more abundant than E protein. The M protein plays a major role in the viral assembly along with E and N proteins [13,14,55]. The N protein is responsible for packaging the viral genome RNA (gRNA) into a helical ribonucleocapsid (RNP). SARS-CoV-2 also has 8 accessory proteins derived from subgenomic RNA: 3a, 3b, 6, 7a, 7b, 8b, 9b, and orf14 (based on the National Center for Biotechnology Information [NCBI] annotation NC_045512.2, and [52]), and they are distributed among the structural genes [51,52,56].

Phylogenetic-tree–based analysis of the whole genomes and individual genes suggest that SARS-CoV-2 is closer to SARS-like bat coronaviruses than to SARS-CoVs. Specifically, the S gene of SARS-CoV-2 is closer to SARS-like bat coronaviruses, although the 3a and 8b accessory genes are closer to SARS-CoVs [52,57]. In a recent study based on available genomic sequences, it was observed that the SARS-CoV-2 (106-sequence) genome has a much lower mutation rate and genetic diversity than SARS-CoV (39 sequences), and in particular, the S-protein–coding gene is relatively more conserved than other protein-encoding genes [58].

Structure of the SARS-CoV-2 RNA-dependent RNA polymerase complex

Coronaviruses use an RNA-dependent RNA polymerase (RdRp) complex for replication of their genome and transcription of their genes [16]. The SARS-CoV-2 RdRp complex is

composed of a catalytic subunit nsp12 and two accessory subunits nsp7 and nsp8, which increase RdRp template binding and processivity [59]. The mechanism of replication and inhibition of SARS-CoV-2 RdRp has been elucidated by several groups using cryo-EM structures of the RdRp–nps7–nsp8 complex [15], its complex with RNA [60], and Remdesivir [61]. The overall structure of the SARS-CoV-2 nsp12–nsp7–nsp8 complex highly resembles that of SARS-CoV, with a global root mean-square deviation (RMSD) of approximately 1 Å for the α-carbon atoms [15,62]. The SARS-CoV-2 RdRp complex structure reveals that the nsp12 core catalytic subunit is bound to a heterodimer of nsp7–nsp8 and an additional nsp8 subunit at a different binding site (Fig 4A–4C) [62]. The N-terminus of nsp12 contains nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain followed by an interface domain and a C-terminal RdRp domain (Fig 4B) [15]. The RdRp domain includes 7 conserved motifs (A–G), which are distributed in the finger, palm, and thumb subdomains (Fig 4A and 4C). The palm subdomain is formed by 5 conserved motifs A–E; motif C contains a critical SDD sequence ("Ser-Asp-Asp" residues 759–761), which forms the catalytic active center. Both D760 and D761 coordinate with 2 magnesium ions at the catalytic center. The F and G motifs are located within the finger subdomain and direct the template strand RNA into the active site, and the thumb subdomain intersects the extensions from the finger subdomain to hold the first turn of RNA [59,60,61,62]. The residues involved in RNA binding as well as forming the catalytic active site are highly conserved among different RNA viruses, which highlights the conserved mechanism of genome replication used by RdRp [61].

The RNA polymerase of the viruses is an established target for inhibiting the viral replication and has pre-established values for clinical engagements by the broad-spectrum nucleotides, such as prodrug remdesivir. These drugs have shown therapeutic efficacies against several viruses from different families, including Ebola, Nipah, MERS, and SARS-CoV [63,64]. The cell-based studies in Vero E6 cells (American Type Culture Collection [ATCC]-1586) have shown that Remdesivir is able to potently block SARS-CoV-2 viral infections at very low concentrations (EC_{50} = 0.77 μM) in vitro [65]. The cryo-EM structure of the RdRp–Remdesivir complex suggests that Remdesivir inhibits the viral RdRp activity through nonobligate RNA chain termination, a mechanism that converts the prodrug to the active drug in the triphosphate form [61]. Besides Remdesivir, Flavipiravir, Ribavirin, Galidesivir, and EIDD-2801 have been shown to inhibit SARS-CoV-2 replication in cell-based assays. Specifically, EIDD-2801 is 3–10 times more potent than Remdesivir in blocking SARS-CoV-2 replication [66]. The cryo-EM structure of the RdRp–Remdesivir complex (Protein Data Bank [PDB]: 7BV2) provides the mechanism of Remdesivir binding, as well as a blueprint for designing more potent antiviral therapeutics to combat the vicious infection of SARS-CoV-2 [61].

Structure, function, antigenicity, and hACE2 receptor recognition by the SARS-CoV-2 S glycoprotein

The S protein is a multifunctional molecular machine that plays key roles in the early steps of viral infection by interacting with host susceptibility factors, including receptors and proteases. These interactions subsequently infect human cells, which contain hACE2 transmembrane proteins [67]. The SARS-CoV-2-S is a transmembrane glycoprotein composed of S1 regions containing the NTD and CTD, S2, a transmembrane region, and a short cytoplasmic domain (Fig 3C and 3D). Both cryo-EM and crystallographic methods have been used to determine multiple structures of the SARS-CoV-2 S protein alone, including the ectodomain of S protein (SARS-CoV-2-S), RBD (SARS-CoV-2-RBD), or in complex with full-length hACE2 or soluble hACE2/B°AT1, in a very short time. These structural studies have enabled us to understand the molecular basis of SARS-CoV-2 entry into human cells displaying ACE2 receptors
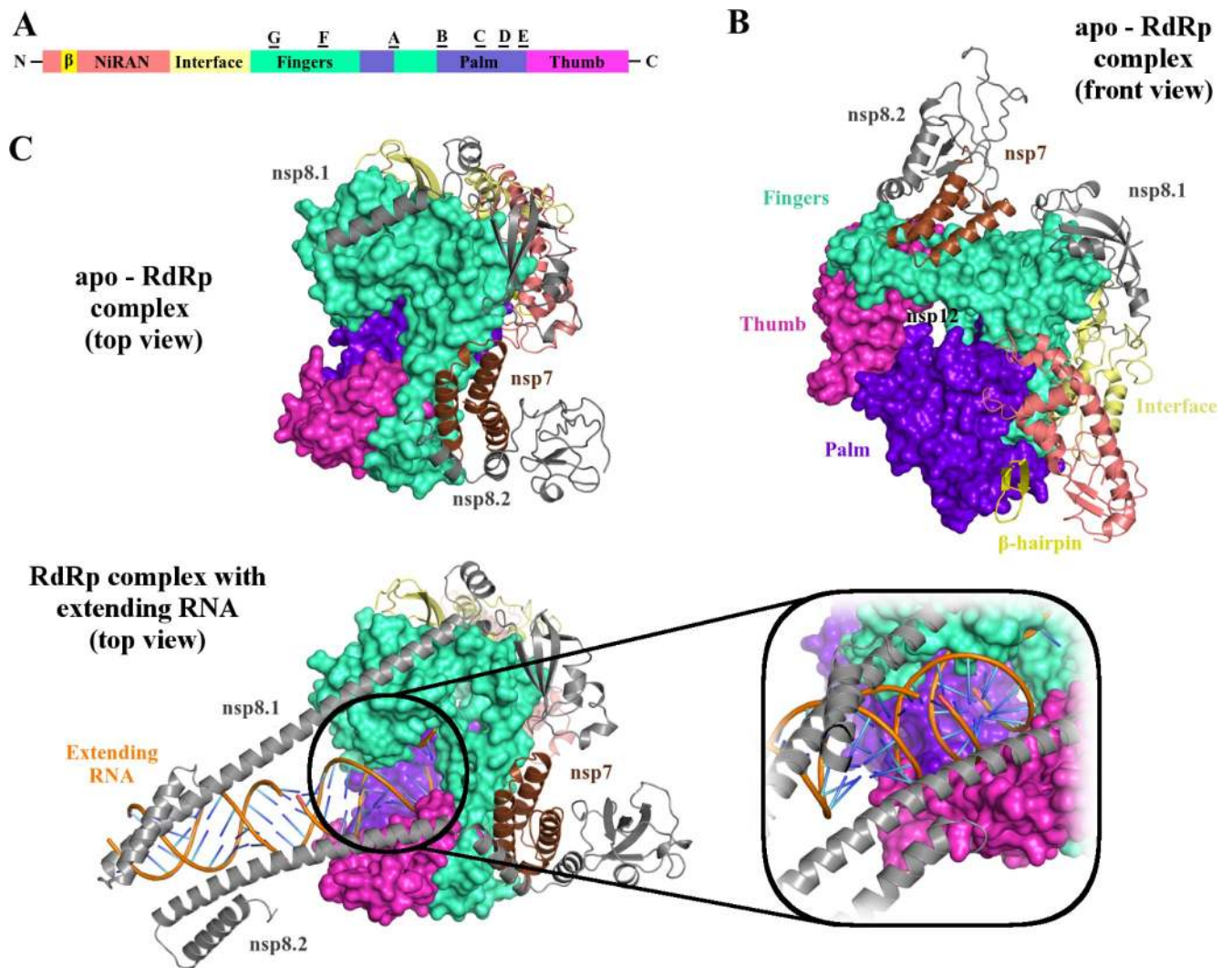
Fig 4. Cryo-EM structure of RdRp of SARS-CoV-2. (A) The domain architecture of RdRp or nsp12 of SARS-CoV-2 is subdivided into NiRAN, interface, fingers, palm, and thumb subdomains; A–G indicate conserved motifs. (B) The cryo-EM structure of apo-RdRp complex (shown as front view, PDB: 7BV1) consists of nsp12, nsp7 (brown), and 2 chains of nsp8 (nsp8.1 and nsp8.2, both in gray). The nsp8.1 interacts directly with nsp12, whereas the nsp8.2 binds to nsp7, which in turn interacts with nsp12. The RNA template is expected to enter the active site, which is formed by motifs A and C through a groove clamped by motifs F and G. Motif E and the thumb subdomain support the primer strand. The RdRp subdomain color scheme is according to Fig 4A. (C) The cryo-EM structure (in top view) of the RdRp complex bound to RNA (PDB: 6YYT) shows 2 chains of nsp8 stabilizing the extending RNA with their alpha helices. The apo-RdRp complex structure (PDB: 7BV1) is shown for comparison. The active site is expanded to show the RNA molecules coming out of the groove formed by the finger and the thumb subdomains. The figures were prepared using Pymol. cryo-EM, cryo-electron microscopy; NiRAN, nidovirus RdRp-associated nucleotidyltransferase; PDB, Protein Data Bank; RdRp, RNA-dependent RNA polymerase; SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus-2.

[17,68,69,70]. Several structures of SARS-CoV-2-S were observed in multiple states (the prefusion, closed, and partially open conformations and in complex with hACE2 receptor) with the RBDs either in an "up" or "down" conformation (Fig 5A and 5B). Of note, to engage the hACE2 receptor, the RBDs of S1 undergo hinge-like movements that either expose or hide the receptor-binding regions and these conformations are referred to as "up" (receptor-accessible) or "down" (receptor-inaccessible) conformations, respectively. SARS-CoV-2-S structures show that the protein adopts a clover-shaped homotrimeric structure, with 3 S1 heads that recognize a cognate cell-surface receptor and a membrane-anchored trimeric S2 stalk that
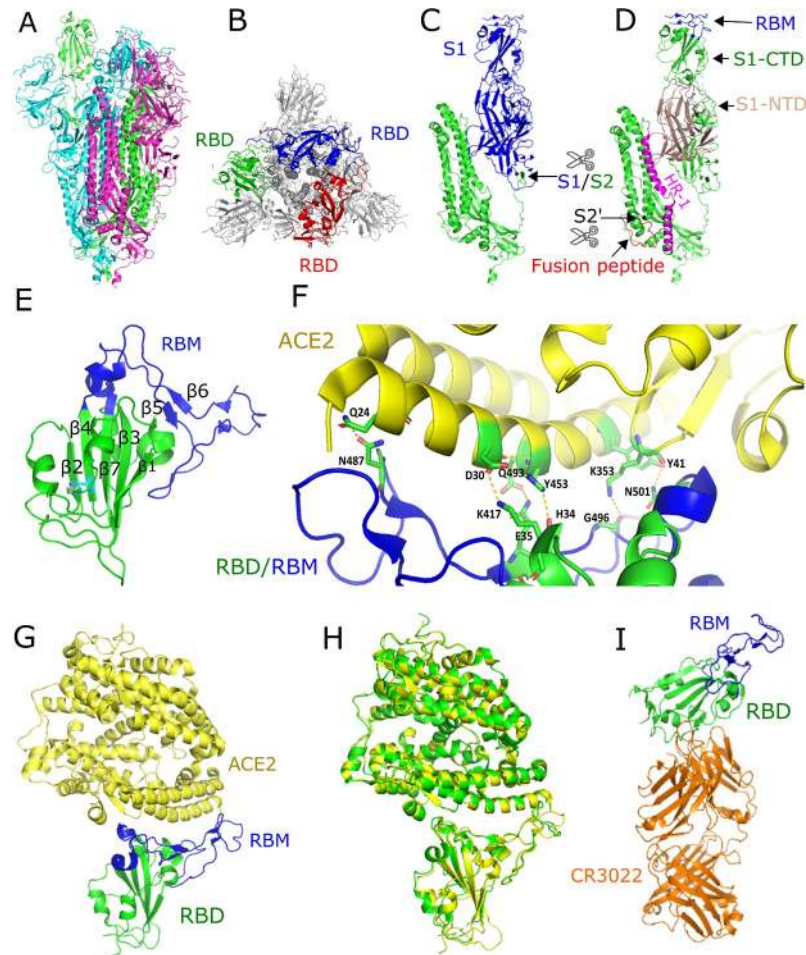
Fig 5. Structure of the SARS-CoV-2 S protein alone and in complex with ACE2 receptor. (A) Side view of the trimeric SARS-CoV-2 S ectodomain in the prefusion state (PDB: 6VSB). The protomer in green is in the "up" conformation, and the other 2 protomers in red and cyan are in "down" conformation. (B) Top view of the trimeric S protein showing RBDs in red, blue, and green on each protomer. (C) Structure of a single protomer showing the receptor-binding subunit S1 (blue) and the membrane-fusion subunit S2 (green). The furin-like protease site at the boundary of S1/S2 subunits is depicted. (D) The S1 subunit showing the RBM in the CTD region (blue) and the NTD region (brown). The S2 subunit showing the fusion peptide (red), second cleavage site S2′ (black), and HR1 (pink). (E) Structure of the RBD, core subdomain (green), and RBM (blue) (PDB: 6LZG). (F) SARS-CoV-2-RBD:ACE2 receptor polar interface shown by specific residues. (G) Structure of the SARS-CoV-2-RBD in complex with ACE2 receptor (PDB: 6LZG). (H) Structural similarity between the SARS-CoV-RBD:hACE2 (green) and SARS-CoV-2-S-CTD:hACE2 (yellow) complexes. (I) Crystal structure of the SARS-CoV-2-RBD (green) in complex with a monoclonal antibody CR3022 (orange). The RBM and CR3022 binding sites do not overlap and are distantly located on the RBD (PDB: 6W41). The figures were prepared using Pymol. ACE2, angiotensin-converting enzyme 2; CTD, C-terminal domain; hACE2, human ACE2; HR1, heptad repeat 1; NTD, N-terminal domain; PDB, Protein Data Bank; RBD, receptor-binding domain; RBM, receptor-binding motif; S protein, spike protein; SARS-CoV, Severe Acute Respiratory Syndrome Coronavirus; SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus-2.

contains the fusion machinery and is primarily α-helical [17] (Fig 5C and 5D). In the prefusion conformation of SARS-CoV-2-S protein, the RBDs rest above the trimeric S2 stalk, exhibiting 2 protomers in the "down" conformation and 1 protomer in the "up" conformation, which is a receptor-accessible state required for binding to an hACE2 receptor [17]. Overall, the SARS-CoV-2-S ectodomain resembles the closely related SARS-CoV-S structure with an RMSD of 3.8 Å over 959 Cα atoms, with a high degree of structural homology when individual domains of SARS-CoV-S and SARS-CoV-2-S were aligned [17].

SARS-CoV-2-RBD interactions with hACE2 receptor

Multiple structures of SARS-CoV-2-RBD in complex with either full-length hACE2 or soluble hACE2 have shown that the extracellular peptidase domain (PD) of ACE2 recognizes the RBDs of S protein mainly through polar interactions [68,69]. Similar to other betacorona-viruses, the SARS-CoV-2-RBD structure suggested that it contains 2 subdomains: a core sub-domain containing a twisted 5-stranded antiparallel β-sheet (β1, β2, β3, β4, and β7) with a conserved disulfide bond between β2–β4. The other subdomain is the receptor-binding motif (RBM), located between β4 and β7 strands as an extended insertion (Fig 5E).

The RBM forms a gently concave surface that accommodates the N-terminal α-helix of the hACE2 and a series of hydrophilic residues that form a solid network of H-bond and salt bridge interactions (Fig 5F). In brief, strong polar contacts include CTD residues A475, N487, E484, and Y453 that interact with S19, Q24, K31, and H34 of the α1 helix of hACE2, respectively [11]. In addition, residues Q498, T500, and N501 on the bulged loop form a network of H-bonds with Y41, Q42, K353, and R357 from hACE2 [68]. Thus, overall virus–receptor inter-actions are dominated by polar contacts mediated by hydrophilic residues [11,68,69] (Fig 5G).

Comparison of the SARS-CoV-2-RBD and SARS-CoV-RBD interactions with hACE2 receptor

The majority of the secondary structure elements shared between SARS-CoV-RBD (PDB: 2AJF) and SARS-CoV-2-RBD (PDB: 6LZG, 6M17) are well superimposed, with an RMSD of 0.475 Å over 128 Cα atoms, with the exception of the receptor-binding loop. Interestingly, these structures revealed that the majority of binding sites of SARS-CoV-RBD in hACE2 also overlap with the SARS-CoV-2-RBD–binding sites, suggesting that the SARS-CoV-2-RBD: hACE2 complex is strikingly similar to the SARS-CoV-RBD:hACE2 structure with an RMSD of 0.431 Å over 669 Cα atoms (Fig 5G and 5H). However, despite the overall similarity, a num-ber of sequence variations were observed at the binding interface that may account for the dif-ference in the affinities for hACE2 receptors. The detailed comparison of the receptor-binding interfaces suggested that the SARS-CoV-2-RBD:hACE2 complex (PDB: 6VW1, 6M17) has larger buried surface areas (1,773 Å^2 versus 1,686 Å^2), has additional contacts (21 versus 17) and more van der Waals interactions (288 versus 213), as well as more H-bonds (16 versus 1), than the SARS-CoV-RBD:hACE2 (PDB: 2AJF) complex [69]. Additionally, residue F486 in SARS-CoV-2-RBD forms stronger aromatic–aromatic interactions with Y83 of hACE2 than I472 of SARS-Co-V-RBD. Residue E484 in the SARS-CoV-2-RBD also forms stronger ionic interactions with K31 compared to P470 of SARS-CoV-RBD [69]. A SARS-CoV-2 sample col-lected from the state of Kerala in India on January 27, 2020, revealed an Arg408 → Ile408 mutation in the SARS-CoV-2-S protein (GenBank: MT012098.1), which otherwise is a strictly conserved residue in SARS-CoV, SARS-CoV-2, and bat SARS-like CoVs. Residue R408 is located near to the binding interface of both the SARS-CoV-2-RBD:hACE2 (PDB: 6VW1) and SARS-CoV-RBD:hACE2 (PDB: 2AJF) complexes but appears not to be interacting directly with hACE2 in either case. However, R408 does form an H-bond (3.3 Å) with the glycan attached to N90 from hACE2, thus potentially contributing to the higher affinities observed for SARS-CoV2-RBD:hACE2 interactions than the SARS-CoV-RBD:hACE2 complex, in which the corresponding R395 is located relatively further away (6.1 Å) from N90 of hACE2. The Arg408 → Ile408 mutation that emerged in SARS-CoV-2 strain (GenBank: MT012098.1) suggested that the loss of H-binding capacity could potentially reduce hACE2 binding affinity. The equilibrium dissociation constants (K_D) for hACE2 interacting with the S proteins have indicated that the binding affinity of SARS-CoV-2-S is several-fold higher than that of SARS-CoV [11,17,69].

Using cryo-EM, the structure of full-length hACE2 in complex with SARS-CoV-2-RBD and B°AT1 (neutral amino acid transporter) was determined. This structure revealed that the hACE2:B°AT1 complex is assembled as a dimer of heterodimers, in which the collectrin-like domain of hACE2 drives homodimerization (PDB: 6M17) [68]. The SARS-CoV-2-RBD is recognized by the extracellular PD of ACE2 as described previously. Further, it demonstrates that a homodimeric ACE2 can accommodate 2 S protein trimers, each through a monomer of ACE2 [68]. Interestingly, a superimposition of the ternary complex on the RBD in the "down" conformation has indicated that the PD clashes with the S protein, whereas in the "up" conformation (PDB: 6VSB), no clashes are observed. This suggests that the "up" conformation of the RBD is a receptor-accessible state and is essential for the ACE2-receptor binding. Taken together, the overall interface between SARS-CoV2-RBD:hACE2 is very similar to the previously known SARS-CoV-RBD:hACE2 interface and is dominated by the polar interactions, as reported by different investigations [11,68,69]. These observations further suggest that SARS-CoV-2-RBD has increased atomic interactions with hACE2, which results in higher affinities compared with the SARS-CoV-RBD:hACE2 complex, which might be one of the reasons for enhanced human-to-human transmission of SARS-CoV-2.

SARS-CoV-2 exhibits distinct epitope features on the RBD from SARS-CoV

Numerous binding and neutralization epitopes have been identified on the S protein of coronaviruses, making the S protein an essential target for vaccine design [71,72,73]. Soon after the emergence of COVID-19 pandemic, some of the initial efforts were focused on screening SARS-CoV-S–specific antibodies to find neutralizing antibody/antibodies for vaccine and drug development against SARS-CoV-2. The hypothesis behind these studies was based on significant sequence as well as structural similarities, and moreover, both viruses bind to the same receptor with overlapping epitopes. Therefore, it was expected that SARS-CoV–specific antibody/antibodies alone or in combination can interfere or even inhibit SARS-CoV-2 and hACE2 receptor interactions.

It has been shown in vitro, as well in animal models, that monoclonal antibodies such as 80R [74], CR3014 [75], S230.15 [76], and m396 [76] can block binding of the S1 domain and hACE2 receptors by potently neutralizing SARS-CoV. However, CR3022 [77] alone did not show neutralization, but the mixture of CR3022 and CR3014 both showed neutralization of SARS-CoV in a synergistic fashion by recognizing different epitopes on the RBDs [75]. Of note, a report suggests that CR3022 can also neutralize SARS-CoV alone [78]. Interestingly, researchers from China tested several published SARS-CoV–specific monoclonal antibodies and found that CR3022 can bind to the RBDs of SARS-CoV-2 with a K_D of 6.3 nM, whereas other antibodies such as m396, CR3014, and S230.15 failed to bind to the SARS-CoV-2-S protein [17,79]. However, a low level of binding to SARS-CoV-2-S was observed with a SARS-CoV-S1–specific polyclonal antibody T62 (#40150-T62, Sino Biological Inc., Beijing, China), and it could poorly neutralize SARS-CoV-2-S-protein–mediated virus entry. Further analysis revealed that the epitope for T62 was likely located on the RBDs of SARS-CoV-2-S, but detailed information is lacking [34]. In an exciting study, the Wilson laboratory determined the crystal structure of CR3022 antibody in complex with SARS-CoV-2-RBD (PDB: 6W41) and revealed that CR3022 binds a highly conserved epitope that is distantly located from the receptor-binding site, which enables cross-reactive binding, but could not neutralize SARS-CoV-2 in vitro [78] (Fig 5I). However, whether CR3022 can synergize with other SARS-CoV-2-RBD–binding antibodies for neutralization requires further evaluation and study.

The SARS-CoV (GenBank: AY278488.2) and SARS-CoV-2 (GenBank: MN908947.3) S proteins share about 76% amino acid sequence identity, suggesting that the remaining 24% amino

acid sequences, which are nonconserved, might be responsible for antigenic differences between these 2 proteins. In the quest to find novel antibody binding epitopes on S proteins, Zheng and colleagues performed antibody epitope analysis and surface epitope accessibility using bioinformatic tools to identify both weak and strong epitopes, which might be otherwise experimentally ignored [80]. Their analysis identified 5 shared epitopes, along with 40 and 29 unique epitopes, on the S proteins of SARS-CoV and SARS-CoV-2, respectively. Among these unique epitopes, 92.7% originated from nonconserved regions, which might explain why most of the SARS-CoV–specific antibodies discussed in this review did not bind to the S protein of SARS-CoV-2 [80]. Taken together, these results suggest the necessity of developing SARS-CoV-2–specific antibodies and vaccine candidates.

ACE2-independent receptors in viral pathogenesis

It is now established that both SARS-CoV and SARS-CoV-2 exploit hACE2 receptor to gain a host-cell entry [10,23]; however, some studies indicate that in addition to ACE2, SARS-CoV might exploit other host factors such as vimentin (a cytoskeleton protein) and lectins (a glycoprotein) to mediate viral entry [81,82]. It is important to note that the precise role of lectins in SARS-CoV infection has not been explored extensively, and the topic remains controversial. Jeffers and colleagues reported that SARS-CoV might use both ACE-2 and CD209L, a C-type lectin, to invade host cells [82]. In contrast, Zhou and colleagues and others indicated that mannose-binding lectins interfere with viral entry, potentially by blocking other interactions [83,84].

The S protein seems to be heavily glycosylated; however, the role of glycosylation in SARS-CoV-2 infection also remains an unexplored area. A recent in vitro study reported interactions between SARS-CoV-2 S protein and C-type lectins as well as sialic-acid binding lectins; however, a major limitation of this study is that it does not provide any cell-based in vivo data, and proper controls are missing [85]. SARS-CoV-2 seems to infect a diverse range of cell types; therefore, it is reasonable to speculate that ACE2-independent interactions might provide an alternate route for viral invasion. Given the importance of this topic and its massive impact on human lives, future studies will have to carefully evaluate whether non-ACE2 interactions compete with ACE2 to inhibit viral entry or ACE2-independent interactions produce a synergistic effect with ACE2-mediated entry to exacerbate the symptoms of COVID-19.

Conclusions

The recent global outbreak of COVID-19 has killed almost 570,000 people (as of July 14, 2020) [7] and threatened the global economy, causing economic hardships to millions of people. Extensive progress has been made in terms of understanding the structure and function of the S glycoproteins. Specifically, decade-long structural studies on the S proteins of SARS-CoVs have designated 6 key residues (Y442, L472, N479, D480, T487, and Y491 for SARS-CoV) [67] in the RBDs that are critical for the host-cell ACE2 receptor binding, as well as for playing important roles in the cross-species transmission. Notably, 5 out of these 6 residues differ between the RBDs of SARS-CoV and SARS-CoV-2 S proteins, which have exhibited enhanced binding between the RBDs of SARS-CoV-2 and ACE2 receptors. This might be one of the reasons behind widespread human-to-human transmission of SARS-CoV-2. In addition, there are likely to be other factors that contribute to the infectivity and pathogenicity of SARS-CoV-2 that need to be investigated.

The trimeric prefusion structure of the SARS-CoV-2 S protein was obtained in an asymmetric conformation in which 1 protomer was observed in the "up" and the other 2 in the "down" conformations. This phenomenon, known as protein "breathing," was observed in the

S1 domain while determining the trimeric prefusion structure, which suggested the mechanism used by the CR3022 antibody to access a cryptic epitope on the trimeric S protein that is otherwise not possible. Interestingly, a similar breathing phenomenon identified unique and conserved epitopes in the trimeric interface of influenza hemagglutinin protein recently. The antibodies binding to these cryptic epitopes did not inhibit viral infection in vitro but conferred in vivo protection [86,87]. A similar phenomenon was observed in case of the CR3022 monoclonal antibody; therefore, further in vivo studies are required as soon as a suitable animal model is established for SARS-CoV-2 studies. In the course of writing this review, 2 exciting reports became available: (i) an antibody 47D11 that is reported to neutralize SARS-CoV-2, as well as SARS-CoV, in cell culture through an unknown mechanism that is different from the virus neutralization process [88]; and (ii) an inactivated novel coronavirus vaccine (PiCoV-acc) that is able to induce SARS-CoV-2–specific neutralizing antibodies in mice, rats, and non-human primates. Additionally, data demonstrate that the PiCoVacc vaccine provides partial to complete protection in macaques against SARS-CoV-2 challenge [89]. Future investigations are required to understand the mechanism of neutralization by these antibodies.

Last but not the least, glycosylation has been an important measure of virus antigenic properties and plays a critical role for the manufacturing of effective vaccines against HIV and influenza. Notably, the SARS-CoV-2 S protein is densely decorated by host-derived heterogenous N-linked glycans, as indicated by a site-specific glycosylation analysis undertaken by mass spectrometry. Specifically, each SARS-CoV-2 S trimer displays 66 N-linked glycosylation sites with an elevation in oligomannose- and hybrid-type glycans compared with typical host-derived glycoproteins [90]. Finally, glycan profiling will be an important addition to measure antigen quality and should be examined while producing glycoprotein-based vaccine candidates for COVID-19.

Though it is observed that SARS-CoV-2 binds to its receptor on the host cells with higher affinities than SARS-CoV, the fatality rate caused by SARS-CoV-2 (3.4%) is significantly less than the reported rate of SARS-CoV (9%–11%, WHO). The reasons behind these differences remain elusive, and future research will shed light on these variations. Recent sequencing data indicate that the SARS-CoV-2 mutation rate is approximately 25 mutations per year. If these mutations enable more efficient virus spread and increased pathogenicity, then vaccine development will be a challenging task. Hopefully, future studies will be able to resolve these questions and come up with medications as well as vaccines against this deadly virus. Even with the vaccine and medications against this virus, future outbreaks of similar viruses and pathogens are likely to continue. Therefore, apart from curbing this outbreak, government policies and efforts should be made to formulate thorough measures to prevent future outbreaks of viruses and bacteria (there is already a significant threat from antibiotic-resistant bacteria).

Supporting information

S1 Fig. Phylogenetic tree of coronaviruses. The figure shows the phylogenetic tree drawn for 69 coronavirus genomic sequences, including the SARS-CoV-2 sequences. Sequences belonging to different Coronavirinae subfamilies are labeled. SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus-2.
(TIF)

Acknowledgments

References

1. Hamre D, Procknow JJ. A new virus isolated from the human respiratory tract. Proc Soc Exp Biol Med. 1966; 121: 190–193. https://doi.org/10.3181/00379727-121-30734 PMID: 4285768

2. Kahn JS, McIntosh K. History and recent advances in coronavirus discovery. Pediatr Infect Dis J. 2005; 24: S223–227, discussion S226. https://doi.org/10.1097/01.inf.0000188166.17324.60 PMID: 16378050

3. Tyrrell DA, Bynoe ML. Cultivation of viruses from a high proportion of patients with colds. Lancet. 1966; 1: 76–77. https://doi.org/10.1016/s0140-6736(66)92364-6 PMID: 4158999

4. WHO. Coronavirus never before seen in humans is the cause of SARS [Internet]. 2003 [cited 2020 Jun 27]. Available from: https://www.who.int/mediacentre/news/releases/2003/pr31/en/

5. Holmes KV. SARS-associated coronavirus. N Engl J Med. 2003; 348: 1948–1951. https://doi.org/10.1056/NEJMp030078 PMID: 12748314

6. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med. 2012; 367: 1814–1820. https://doi.org/10.1056/NEJMoa1211721 PMID: 23075143

7. WHO. Coronavirus Disease (COVID-19) Dashboard [Internet]. 2020 [cited 2020 Jun 27]. Available from: https://covid19.who.int/

8. WHO. Coronavirus Disease (COVID-19) Dashboard, United States of America [Internet]. 2020 [cited 2020 Jun 27]. Available from: https://covid19.who.int/region/amro/country/us

9. Peck KM, Burch CL, Heise MT, Baric RS. Coronavirus Host Range Expansion and Middle East Respiratory Syndrome Coronavirus Emergence: Biochemical Mechanisms and Evolutionary Perspectives. Annu Rev Virol. 2015; 2: 95–117. https://doi.org/10.1146/annurev-virology-100114-055029 PMID: 26958908

10. Li F. Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu Rev Virol. 2016; 3: 237–261. https://doi.org/10.1146/annurev-virology-110615-042301 PMID: 27578435

11. Lan J, Ge J, Yu J, Shan S, Zhou H, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 2020; 581: 215–220. https://doi.org/10.1038/s41586-020-2180-5 PMID: 32225176

12. Hofmann H, Pohlmann S. Cellular entry of the SARS coronavirus. Trends Microbiol. 2004; 12: 466–472. https://doi.org/10.1016/j.tim.2004.08.008 PMID: 15381196

13. Vennema H, Godeke GJ, Rossen JW, Voorhout WF, Horzinek MC, et al. Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. EMBO J. 1996; 15: 2020–2028. https://doi.org/10.1002/j.1460-2075.1996.tb00553.x PMID: 8617249

14. Siu YL, Teoh KT, Lo J, Chan CM, Kien F, et al. The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles. J Virol. 2008; 82: 11318–11330. https://doi.org/10.1128/JVI.01052-08 PMID: 18753196

15. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science. 2020; 368: 779–782. https://doi.org/10.1126/science.abb7498 PMID: 32277040

16. Snijder EJ, Decroly E, Ziebuhr J. The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. Adv Virus Res. 2016; 96: 59–126. https://doi.org/10.1016/bs.aivir.2016.08.008 PMID: 27712628

17. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020; 367: 1260–1263. https://doi.org/10.1126/science.abb2507 PMID: 32075877

18. Wentworth DE, Holmes KV. Molecular determinants of species specificity in the coronavirus receptor aminopeptidase N (CD13): influence of N-linked glycosylation. J Virol. 2001; 75: 9741–9752. https://doi.org/10.1128/JVI.75.20.9741-9752.2001 PMID: 11559807

19. Raj VS, Mou H, Smits SL, Dekkers DH, Muller MA, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature. 2013; 495: 251–254. https://doi.org/10.1038/nature12005 PMID: 23486063

20. Hulswit RJG, Lang Y, Bakkers MJG, Li W, Li Z, et al. Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. Proc Natl Acad Sci U S A. 2019; 116: 2681–2690. https://doi.org/10.1073/pnas.1809667116 PMID: 30679277

21. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature. 2003; 426: 450–454. https://doi.org/10.1038/nature02145 PMID: 14647384

22. Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, et al. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc Natl Acad Sci U S A. 2005; 102: 7988–7993. https://doi.org/10.1073/pnas.0409465102 PMID: 15897467

23. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell. 2020; 181: 271–280 e278. https://doi.org/10.1016/j.cell.2020.02.052 PMID: 32142651

24. Yob JM, Field H, Rashdi AM, Morrissy C, van der Heide B, et al. Nipah virus infection in bats (order Chiroptera) in peninsular Malaysia. Emerg Infect Dis. 2001; 7: 439–441. https://doi.org/10.3201/eid0703.010312 PMID: 11384522

25. Halpin K, Young PL, Field HE, Mackenzie JS. Isolation of Hendra virus from pteropid bats: a natural reservoir of Hendra virus. J Gen Virol. 2000; 81: 1927–1932. https://doi.org/10.1099/0022-1317-81-8-1927 PMID: 10900029

26. Li W, Shi Z, Yu M, Ren W, Smith C, et al. Bats are natural reservoirs of SARS-like coronaviruses. Science. 2005; 310: 676–679. https://doi.org/10.1126/science.1118391 PMID: 16195424

27. Lau SK, Woo PC, Li KS, Huang Y, Tsoi HW, et al. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc Natl Acad Sci U S A. 2005; 102: 14040–14045. https://doi.org/10.1073/pnas.0506735102 PMID: 16169905

28. WHO. Middle East respiratory syndrome coronavirus (MERS-CoV) [Internet]. 2019 [cited 2020 Jun 27]. Available from: https://www.who.int/emergencies/mers-cov/en/

29. Zhu N, Zhang D, Wang W, Li X, Yang B, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020; 382: 727–733. https://doi.org/10.1056/NEJMoa2001017 PMID: 31978945

30. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020; 579: 270–273. https://doi.org/10.1038/s41586-020-2012-7 PMID: 32015507

31. Chan JF, Kok KH, Zhu Z, Chu H, To KK, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg Microbes Infect. 2020; 9: 221–236.

32. Lu R, Zhao X, Li J, Niu P, Yang B, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020; 395: 565–574. https://doi.org/10.1016/S0140-6736(20)30251-8 PMID: 32007145

33. WHO. Naming the coronavirus disease (COVID-19) and the virus that causes it. 2020 [cited 2020 Jun 27]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it

34. Ou X, Liu Y, Lei X, Li P, Mi D, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. Nat Commun. 2020; 11: 1620. https://doi.org/10.1038/s41467-020-15562-9 PMID: 32221306

35. Guo YR, Cao QD, Hong ZS, Tan YY, Chen SD, et al. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak—an update on the status. Mil Med Res. 2020; 7: 11.

36. Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, et al. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antiviral Res. 2020; 176: 104742. https://doi.org/10.1016/j.antiviral.2020.104742 PMID: 32057769

37. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nat Med. 2020; 26: 450–452. https://doi.org/10.1038/s41591-020-0820-9 PMID: 32284615

38. Klenk HD, Garten W. Host cell proteases controlling virus pathogenicity. Trends Microbiol. 1994; 2: 39–43. https://doi.org/10.1016/0966-842x(94)90123-6 PMID: 8162439

39. Steinhauer DA. Role of hemagglutinin cleavage for the pathogenicity of influenza virus. Virology. 1999; 258: 1–20. https://doi.org/10.1006/viro.1999.9716 PMID: 10329563

40. Chan JF, Yuan S, Kok KH, To KK, Chu H, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 2020; 395: 514–523. https://doi.org/10.1016/S0140-6736(20)30154-9 PMID: 31986261

41. Liu J, Liao X, Qian S, Yuan J, Wang F, et al. Community Transmission of Severe Acute Respiratory Syndrome Coronavirus 2, Shenzhen, China, 2020. Emerg Infect Dis. 2020; 26: 1320–1323. https://doi.org/10.3201/eid2606.200239 PMID: 32125269

42. BBC. Coronavirus: Confirmed global cases pass one million [Internet]. Apr 3, 2020 [cited 2020 Jun 27]. Available from: https://www.bbc.com/news/world-52144390

43. Coronaviridae Study Group of the International Committee on Taxonomy of V The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020; 5: 536–544. https://doi.org/10.1038/s41564-020-0695-z PMID: 32123347

44. Fung TS, Liu DX. Human Coronavirus: Host-Pathogen Interaction. Annu Rev Microbiol. 2019; 73: 529–557. https://doi.org/10.1146/annurev-micro-020518-115759 PMID: 31226023

45. Woo PC, Lau SK, Lam CS, Lau CC, Tsang AK, et al. Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alpha-coronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. J Virol. 2012; 86: 3995–4008. https://doi.org/10.1128/JVI.06540-11 PMID: 22278237

46. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol. 2019; 17: 181–192. https://doi.org/10.1038/s41579-018-0118-9 PMID: 30531947

47. Fouchier RA, Hartwig NG, Bestebroer TM, Niemeyer B, de Jong JC, et al. A previously undescribed coronavirus associated with respiratory disease in humans. Proc Natl Acad Sci U S A. 2004; 101: 6212–6216. https://doi.org/10.1073/pnas.0400762101 PMID: 15073334

48. van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJ, et al. Identification of a new human coronavirus. Nat Med. 2004; 10: 368–373. https://doi.org/10.1038/nm1024 PMID: 15034574

49. Lau SK, Woo PC, Yip CC, Tse H, Tsoi HW, et al. Coronavirus HKU1 and other coronavirus infections in Hong Kong. J Clin Microbiol. 2006; 44: 2063–2071. https://doi.org/10.1128/JCM.02614-05 PMID: 16757599

50. Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2 genomes. Proc Natl Acad Sci U S A. 2020; 117: 9241–9243. https://doi.org/10.1073/pnas.2004999117 PMID: 32269081

51. Wang C, Liu Z, Chen Z, Huang X, Xu M, et al. The establishment of reference sequence for SARS-CoV-2 and variation analysis. J Med Virol. 2020; 92(6): 667–674. https://doi.org/10.1002/jmv.25762 PMID: 32167180

52. Wu A, Peng Y, Huang B, Ding X, Wang X, et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. Cell Host Microbe. 2020; 27: 325–328. https://doi.org/10.1016/j.chom.2020.02.001 PMID: 32035028

53. Harcourt BH, Jukneliene D, Kanjanahaluethai A, Bechill J, Severson KM, et al. Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. J Virol. 2004; 78: 13600–13612. https://doi.org/10.1128/JVI.78.24.13600-13612.2004 PMID: 15564471

54. Nieto-Torres JL, Dediego ML, Alvarez E, Jimenez-Guardeno JM, Regla-Nava JA, et al. Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. Virology. 2011; 415: 69–82. https://doi.org/10.1016/j.virol.2011.03.029 PMID: 21524776

55. Voss D, Kern A, Traggiai E, Eickmann M, Stadler K, et al. Characterization of severe acute respiratory syndrome coronavirus membrane protein. FEBS Lett. 2006; 580: 968–973. https://doi.org/10.1016/j.febslet.2006.01.026 PMID: 16442106

56. Zumla A, Chan JF, Azhar EI, Hui DS, Yuen KY. Coronaviruses—drug discovery and therapeutic options. Nat Rev Drug Discov. 2016; 15: 327–347. https://doi.org/10.1038/nrd.2015.37 PMID: 26868298

57. Menachery VD, Yount BL Jr, Debbink K, Agnihothram S, Gralinski LE, et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. Nat Med. 2015; 21: 1508–1513. https://doi.org/10.1038/nm.3985 PMID: 26552008

58. Jia Y, Zhang Y, Huang KS, Ho HY, Hor WS, Yang CH, et al. Analysis of the mutation dynamics of SARS-CoV-2 reveals the spread history and emergence of RBD mutant with lower ACE2 binding affinity. bioRxiv 034942 [Preprint]. 2020 [cited 2020 Jun 27]. Available from: https://www.biorxiv.org/content/10.1101/2020.04.09.034942v1

59. Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, et al. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc Natl Acad Sci U S A. 2014; 111: E3900–3909. https://doi.org/10.1073/pnas.1323705111 PMID: 25197083

60. Wang Q, Wu J, Wang H, Gao Y, Liu Q, et al. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. Cell. 2020; 182(2): 417–428.e13. https://doi.org/10.1016/j.cell.2020.05.034 PMID: 32526208

61. Yin W, Mao C, Luan X, Shen DD, Shen Q, et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science. 2020; 368(6498):1499–1504. https://doi.org/10.1126/science.abc1560 PMID: 32358203

62. Peng Q, Peng R, Yuan B, Zhao J, Wang M, et al. Structural and Biochemical Characterization of the nsp12-nsp7-nsp8 Core Polymerase Complex from SARS-CoV-2. Cell Rep. 2020; 31: 107774. https://doi.org/10.1016/j.celrep.2020.107774 PMID: 32531208

63. Warren TK, Jordan R, Lo MK, Ray AS, Mackman RL, et al. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. Nature. 2016; 531: 381–385. https://doi.org/10.1038/nature17180 PMID: 26934220

64. de Wit E, Feldmann F, Cronin J, Jordan R, Okumura A, et al. Prophylactic and therapeutic remdesivir (GS-5734) treatment in the rhesus macaque model of MERS-CoV infection. Proc Natl Acad Sci U S A. 2020; 117: 6771–6776. https://doi.org/10.1073/pnas.1922083117 PMID: 32054787

65. Wang M, Cao R, Zhang L, Yang X, Liu J, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 2020; 30: 269–271. https://doi.org/10.1038/s41422-020-0282-0 PMID: 32020029

66. Sheahan TP, Sims AC, Zhou S, Graham RL, Pruijssers AJ, et al. An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. Sci Transl Med. 2020; 12(541): eabb5883. https://doi.org/10.1126/scitranslmed.abb5883 PMID: 32253226

67. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor Recognition by the Novel Coronavirus from Wuhan: An Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. J Virol. 2020; 94 (7): e00127–20. https://doi.org/10.1128/JVI.00127-20 PMID: 31996437

68. Yan R, Zhang Y, Li Y, Xia L, Guo Y, et al. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science. 2020; 367: 1444–1448. https://doi.org/10.1126/science.abb2762 PMID: 32132184

69. Wang Q, Zhang Y, Wu L, Niu S, Song C, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. Cell. 2020; 181: 894–904 e899. https://doi.org/10.1016/j.cell.2020.03.045 PMID: 32275855

70. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, et al. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell. 2020; 181: 281–292 e286. https://doi.org/10.1016/j.cell.2020.02.058 PMID: 32155444

71. Qiu T, Mao T, Wang Y, Zhou M, Qiu J, et al. Identification of potential cross-protective epitope between a new type of coronavirus (2019-nCoV) and severe acute respiratory syndrome virus. J Genet Genomics. 2020; 47: 115–117. https://doi.org/10.1016/j.jgg.2020.01.003 PMID: 32171450

72. Reguera J, Santiago C, Mudgal G, Ordono D, Enjuanes L, et al. Structural bases of coronavirus attachment to host aminopeptidase N and its inhibition by neutralizing antibodies. PLoS Pathog. 2012; 8: e1002859. https://doi.org/10.1371/journal.ppat.1002859 PMID: 22876187

73. Prabakaran P, Gan J, Feng Y, Zhu Z, Choudhry V, et al. Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody. J Biol Chem. 2006; 281: 15829–15836. https://doi.org/10.1074/jbc.M600697200 PMID: 16597622

74. Sui J, Li W, Murakami A, Tamin A, Matthews LJ, et al. Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association. Proc Natl Acad Sci U S A. 2004; 101: 2536–2541. https://doi.org/10.1073/pnas.0307140101 PMID: 14983044

75. ter Meulen J, van den Brink EN, Poon LL, Marissen WE, Leung CS, et al. Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. PLoS Med. 2006; 3: e237. https://doi.org/10.1371/journal.pmed.0030237 PMID: 16796401

76. Zhu Z, Chakraborti S, He Y, Roberts A, Sheahan T, et al. Potent cross-reactive neutralization of SARS coronavirus isolates by human monoclonal antibodies. Proc Natl Acad Sci U S A. 2007; 104: 12123–12128. https://doi.org/10.1073/pnas.0701000104 PMID: 17620608

77. Coughlin M, Lou G, Martinez O, Masterman SK, Olsen OA, et al. Generation and characterization of human monoclonal neutralizing antibodies with distinct binding and sequence features against SARS coronavirus using XenoMouse. Virology. 2007; 361: 93–102. https://doi.org/10.1016/j.virol.2006.09.029 PMID: 17161858

78. Yuan M, Wu NC, Zhu X, Lee CD, So RTY, et al. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. Science. 2020; 368: 630–633.

79. Tian X, Li C, Huang A, Xia S, Lu S, et al. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. Emerg Microbes Infect. 2020; 9: 382–385. https://doi.org/10.1080/22221751.2020.1729069 PMID: 32065055

80. Zheng M, Song L. Novel antibody epitopes dominate the antigenicity of spike glycoprotein in SARS-CoV-2 compared to SARS-CoV. Cell Mol Immunol. 2020; 17: 536–538. https://doi.org/10.1038/s41423-020-0385-z PMID: 32132669

81. Yu YT, Chien SC, Chen IY, Lai CT, Tsay YG, et al. Surface vimentin is critical for the cell entry of SARS-CoV. J Biomed Sci. 2016; 23: 14. https://doi.org/10.1186/s12929-016-0234-7 PMID: 26801988

82. Jeffers SA, Tusell SM, Gillim-Ross L, Hemmila EM, Achenbach JE, et al. CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. Proc Natl Acad Sci U S A. 2004; 101: 15748–15753. https://doi.org/10.1073/pnas.0403812101 PMID: 15496474

83. Zhou Y, Lu K, Pfefferle S, Bertram S, Glowacka I, et al. A single asparagine-linked glycosylation site of the severe acute respiratory syndrome coronavirus spike glycoprotein facilitates inhibition by mannose-binding lectin through multiple mechanisms. J Virol. 2010; 84: 8753–8764. https://doi.org/10.1128/JVI.00554-10 PMID: 20573835

84. Ip WK, Chan KH, Law HK, Tso GH, Kong EK, et al. Mannose-binding lectin in severe acute respiratory syndrome coronavirus infection. J Infect Dis. 2005; 191: 1697–1704. https://doi.org/10.1086/429631 PMID: 15838797

85. Chiodo F, Bruijns SCM, Rodriguez E, Li RJE, Molinaro A, et al. Novel ACE2-Independent Carbohydrate-Binding of SARS-CoV-2 Spike Protein to Host Lectins and Lung Microbiota. bioRxiv 092478 [Preprint]. 2020 [cited 2020 May 14]. Available from: https://www.biorxiv.org/content/10.1101/2020.05.13.092478v1

86. Bangaru S, Lang S, Schotsaert M, Vanderven HA, Zhu X, et al. A Site of Vulnerability on the Influenza Virus Hemagglutinin Head Domain Trimer Interface. Cell. 2019; 177: 1136–1152 e1118. https://doi.org/10.1016/j.cell.2019.04.011 PMID: 31100268

87. Watanabe A, McCarthy KR, Kuraoka M, Schmidt AG, Adachi Y, et al. Antibodies to a Conserved Influenza Head Interface Epitope Protect by an IgG Subtype-Dependent Mechanism. Cell. 2019; 177: 1124–1135 e1116. https://doi.org/10.1016/j.cell.2019.03.048 PMID: 31100267

88. Wang C, Li W, Drabek D, Okba NMA, van Haperen R, et al. A human monoclonal antibody blocking SARS-CoV-2 infection. Nat Commun. 2020; 11: 2251. https://doi.org/10.1038/s41467-020-16256-y PMID: 32366817

89. Gao Q, Bao L, Mao H, Wang L, Xu K, et al. Rapid development of an inactivated vaccine candidate for SARS-CoV-2. Science. 2020; 369(6499): 77–81. https://doi.org/10.1126/science.abc1932 PMID: 32376603

90. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. Science. 2020; 369(6501): 330–333. https://doi.org/10.1126/science.abb9983 PMID: 32366695