bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Structure, function and antigenicity of the SARS-CoV-2 spike glycoprotein

Alexandra C. Walls[1,5], Young-Jun Park[1,5], M. Alejandra Tortorici[1,2], Abigail Wall[3], Andrew T. McGuire[3,4] and David Veesler[1*]

[1] Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA.

[2] Institute Pasteur & CNRS UMR 3569, Unité de Virologie Structurale, 75015, Paris, France.

[3] Vaccines and Infectious Diseases Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98195, USA.

[4]Department of Global Health, University of Washington, Seattle, Washington 98195, USA

[5] These authors contributed equally

*Correspondence: dveesler@uw.edu

SUMMARY

The recent emergence of a novel coronavirus associated with an ongoing outbreak of pneumonia (Covid-2019) resulted in infections of more than 72,000 people and claimed over 1,800 lives. Coronavirus spike (S) glycoprotein trimers promote entry into cells and are the main target of the humoral immune response. We show here that SARS-CoV-2 S mediates entry in VeroE6 cells and in BHK cells transiently transfected with human ACE2, establishing ACE2 as a functional receptor for this novel coronavirus. We further demonstrate that the receptor-binding domains of SARS-CoV-2 S and SARS-CoV S bind with similar affinities to human ACE2, which correlates with the efficient spread of SARS-CoV-2 among humans. We found that the SARS-CoV-2 S glycoprotein harbors a furin cleavage site at the boundary between the S_1/S_2 subunits, which is processed during biogenesis and sets this virus apart from SARS-CoV and other SARS-related CoVs. We determined a cryo-electron microscopy structure of the SARS-CoV-2 S ectodomain trimer, demonstrating spontaneous opening of the receptor-binding domain, and providing a blueprint for the design of vaccines and inhibitors of viral entry. Finally, we demonstrate that SARS-CoV S murine polyclonal sera potently inhibited SARS-CoV-

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

2 S-mediated entry into target cells, thereby indicating that cross-neutralizing antibodies targeting conserved S epitopes can be elicited upon vaccination.

KEYWORDS

Coronavirus, SARS-CoV-2, SARS-CoV, spike glycoprotein, antibodies, neutralizing antibodies, viral receptor

INTRODUCTION

Three coronaviruses have crossed the species barrier to cause deadly pneumonia in humans since the beginning of the 21st century: severe acute respiratory syndrome coronavirus (SARS-CoV) (Drosten et al., 2003; Ksiazek et al., 2003), Middle-East respiratory syndrome coronavirus (Zaki et al., 2012) (MERS-CoV) and SARS-CoV-2 (Huang et al., 2020; Zhu et al., 2020). SARS-CoV emerged in the Guangdong province of China in 2002 and spread to five continents through air travel routes, infecting 8,098 people and causing 774 deaths. In 2012, MERS-CoV emerged in the Arabian Peninsula, where it remains a major public health concern, and was exported to 27 countries, infecting a total of ~2,494 individuals and claiming 858 lives. A previously unknown coronavirus, named SARS-CoV-2, was discovered in December 2019 in Wuhan, Hubei Province of China and was sequenced and isolated by January 2020 (Zhou et al., 2020; Zhu et al., 2020). SARS-CoV-2 is associated with an ongoing outbreak of atypical pneumonia (Covid-2019) that has affected over 72,436 people and killed more than 1,868 of them in 29 countries as of February 15th 2020. On January 30th 2020, the World Health Organization declared the SARS-CoV-2 epidemic as a public health emergency of international concern.

MERS-CoV was suggested to originate from bats but the reservoir host fueling spillover to humans is unequivocally dromedary camels (Haagmans et al., 2014; Memish et al., 2013). Both SARS-CoV and SARS-CoV-2 are closely related to each other and originated in bats which most likely serve as reservoir host for these two viruses (Ge et al., 2013; Hu et al., 2017; Li et al., 2005b; Yang et al., 2015a; Zhou et al.,

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

2020). Whereas palm civets and racoon dogs have been recognized as an intermediate host for zoonotic transmission of SARS-CoV between bats and humans (Guan et al., 2003; Kan et al., 2005; Wang et al., 2005), the intermediate host of SARS-CoV-2 remains unknown. The recurrent spillovers of coronaviruses in humans along with detection of numerous coronaviruses in bats, including many SARS-related coronaviruses (SARSr-CoVs), suggest that future zoonotic transmission events may continue to occur (Anthony et al., 2017; Ge et al., 2013; Hu et al., 2017; Li et al., 2005b; Menachery et al., 2015; Menachery et al., 2016; Yang et al., 2015a; Zhou et al., 2020). In addition to the highly pathogenic zoonotic pathogens SARS-CoV, MERS-CoV and SARS-CoV-2, all belonging to the β-coronavirus genus, four low pathogenicity coronaviruses are endemic in humans: HCoV-OC43, HCoV-HKU1, HCoV-NL63 and HCoV-229E.To date, no therapeutics or vaccines are approved against any human-infecting coronaviruses.

Coronavirus entry into host cells is mediated by the transmembrane spike (S) glycoprotein that forms homotrimers protruding from the viral surface (Tortorici and Veesler, 2019). S comprises two functional subunits responsible for binding to the host cell receptor (S_1 subunit) and fusion of the viral and cellular membranes (S_2 subunit). For many CoVs, S is cleaved at the boundary between the S_1 and S_2 subunits which remain non-covalently bound in the prefusion conformation (Belouzard et al., 2009; Bosch et al., 2003; Burkard et al., 2014; Kirchdoerfer et al., 2016; Millet and Whittaker, 2014, 2015; Park et al., 2016; Walls et al., 2016a). The distal S_1 subunit comprises the receptor-binding domain(s), and contributes to stabilization of the prefusion state of the membrane-anchored S_2 subunit that contains the fusion machinery (Gui et al., 2017; Kirchdoerfer et al., 2016; Pallesen et al., 2017; Song et al., 2018; Walls et al., 2016a; Walls et al., 2017b; Yuan et al., 2017). For all CoVs, S is further cleaved by host proteases at the so-called S_2' site located immediately upstream of the fusion peptide (Madu et al., 2009; Millet and Whittaker, 2015). This cleavage has been proposed to activate the protein for membrane fusion *via* extensive irreversible conformational changes (Belouzard et al., 2009; Heald-Sargent and Gallagher, 2012; Millet and Whittaker, 2014, 2015; Park et al., 2016; Walls et al., 2017b). As a result, coronavirus

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

entry into susceptible cells is a complex process that requires the concerted action of receptor-binding and proteolytic processing of the S protein to promote virus-cell fusion.

Different coronaviruses use distinct domains within the S_1 subunit to recognize a variety of attachment and entry receptors, depending on the viral species. Endemic human coronaviruses OC43 and HKU1 attach *via* their S domain A (S^A) to 5-N-acetyl-9-*O*-acetyl-sialosides found on glycoproteins and glycolipids at the host cell surface to enable entry into susceptible cells (Hulswit et al., 2019; Tortorici et al., 2019; Vlasak et al., 1988). MERS-CoV S, however, uses domain A to engage non-acetylated sialosides as attachment receptors (Li et al., 2017; Park et al., 2019) and promote subsequent binding of domain B (S^B) to the entry receptor, dipeptidyl-peptidase 4 (Lu et al., 2013; Raj et al., 2013). SARS-CoV and several SARS-related coronaviruses (SARSr-CoV) interact directly with angiotensin-converting enzyme 2 (ACE2) *via* S^B to enter target cells (Ge et al., 2013; Kirchdoerfer et al., 2018; Li et al., 2005a; Li et al., 2003; Song et al., 2018; Yang et al., 2015a).

As the coronavirus S glycoprotein is surface-exposed and mediates entry into host cells, it is the main target of neutralizing antibodies (Abs) upon infection and the focus of therapeutic and vaccine design. S trimers are extensively decorated with N-linked glycans that are important for proper folding (Rossen et al., 1998) and to modulate accessibility to host proteases and neutralizing antibodies (Walls et al., 2016b; Walls et al., 2019; Xiong et al., 2017; Yang et al., 2015b). We previously characterized potent human neutralizing Abs from rare memory B cells of individuals infected with SARS-CoV (Traggiai et al., 2004) or MERS-CoV (Corti et al., 2015) in complex with SARS-CoV S and MERS-CoV S to provide molecular-level information of the mechanism of competitive inhibition of S^B attachment to the host receptor (Walls et al., 2019). The S230 anti-SARS-CoV antibody also acted by functionally mimicking receptor-attachment and promoting spike fusogenic conformational rearrangements through a ratcheting mechanism that elucidated the unique nature of the coronavirus membrane fusion activation (Walls et al., 2019).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

We report here that ACE2 could mediate SARS-CoV-2 S-mediated entry into cells, establishing it as a functional receptor for this newly emerged coronavirus. The SARS-CoV-2 S^B engages human ACE2 (hACE2) with comparable affinity than SARS-CoV S^B from viral isolates associated with the 2002-2003 epidemic (i.e. binding with high affinity to hACE2). Tight binding to hACE2 could explain the efficient transmission of SARS-CoV-2 in humans, as was the case for SARS-CoV. We identified the presence of an unexpected furin cleavage site at the S_1/S_2 boundary of SARS-CoV-2 S, which is cleaved during biosynthesis, a novel feature setting this virus apart from SARS-CoV and SARSr-CoVs. Abrogation of this cleavage motif moderately affected SARS-CoV-2 S-mediated entry into VeroE6 or BHK cells but may contribute to expand the tropism of this virus, as reported for several highly pathogenic avian influenza viruses and pathogenic Newcastle disease virus (Klenk and Garten, 1994; Steinhauer, 1999). We determined a cryo-electron microscopy structure of the SARS-CoV-2 S ectodomain trimer and reveal that it adopts multiple S^B conformations that are reminiscent of previous reports on both SARS-CoV S and MERS-CoV S. Finally, we show that SARS-CoV S mouse polyclonal sera potently inhibited entry into target cells of SARS-CoV-2 S pseudotyped viruses. Collectively, these results pave the way for designing vaccines eliciting broad protection against SARS-CoV-2, SARS-CoV and SARSr-CoV.

RESULTS

ACE2 is an entry receptor for SARS-CoV-2

The SARS-CoV-2 S glycoprotein shares ~80% amino acid sequence identity with the SARS-CoV S Urbani and with bat SARSr-CoV ZXC21 S and ZC45 S glycoprotein. The latter two SARSr-CoV sequences were identified from *Rinolophus sinicus* (Chinese horseshoe bats), the species from which SARSr-CoV WIV-1 and WIV-16 were isolated (Ge et al., 2013; Yang et al., 2015a). Furthermore, Zhou *et al* recently reported that SARS-CoV-2 is most closely related to the bat SARSr-CoV RaTG13 with which it forms a distinct lineage from other SARSr-CoVs, and that their S glycoproteins share 98% amino acid sequence identity (Zhou et al., 2020). SARS-CoV recognizes its entry receptor human ACE2 (hACE2) at the surface of type II pneumocytes, using S^B which shares ~75% overall amino acid sequence identity with SARS-CoV-2 S^B and 50%

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

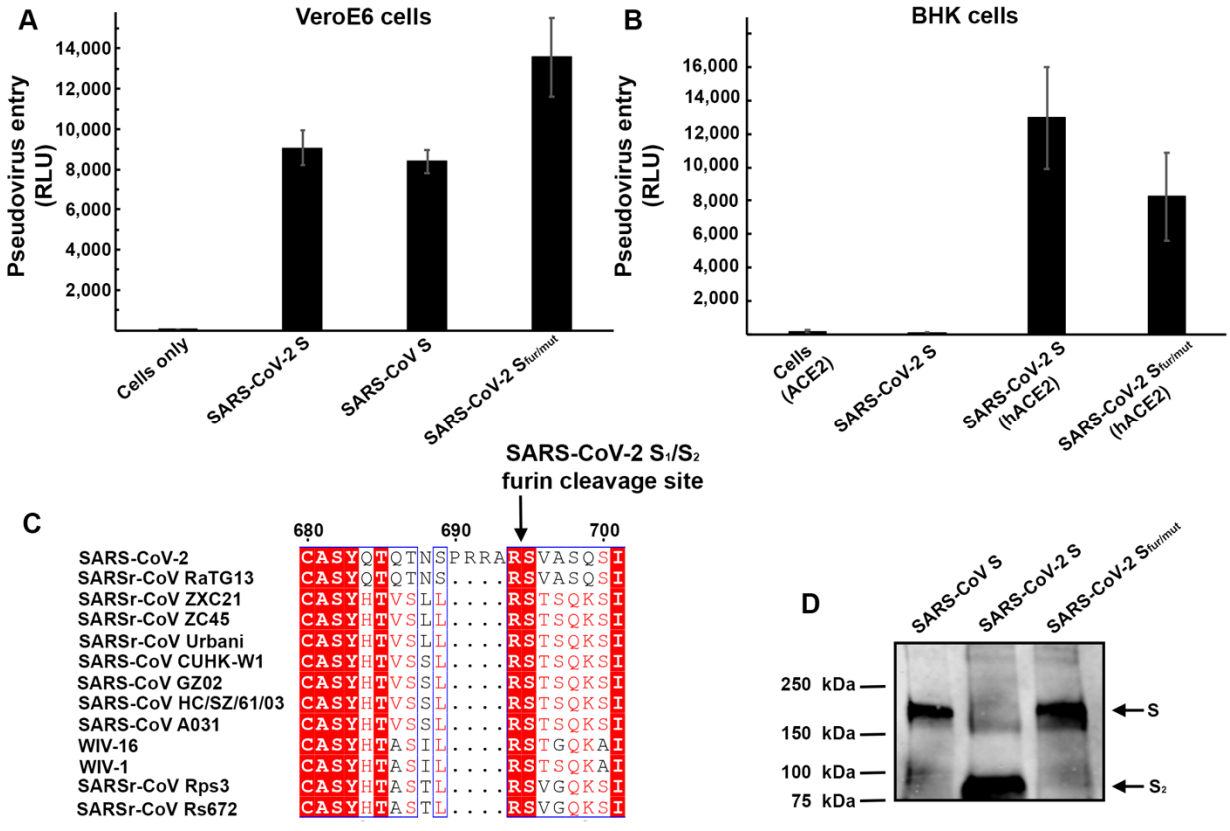
identity within their receptor-binding motifs (RBMs) (Li et al., 2005a; Li et al., 2003; Li et al., 2005c; Wan et al., 2020). Previous studies also showed that the host proteases cathepsin L and TMPRSS2 prime SARS-CoV S for membrane fusion through cleavage at the S_1/S_2 and at the S_2' sites (Belouzard et al., 2009; Bosch et al., 2008; Glowacka et al., 2011; Matsuyama et al., 2010; Millet and Whittaker, 2015; Shulla et al., 2011).

We set out to investigate the functional determinants of S-mediated entry into target cells using a murine leukemia virus (MLV) pseudotyping system (Millet and Whittaker, 2016). To assess the ability of SARS-CoV-2 S to promote entry into target cells, we first compared transduction of SARS-CoV-2 S-MLV and SARS-CoV S-MLV into VeroE6 cells, that are known to express ACE2 and support SARS-CoV replication (Drosten et al., 2003; Ksiazek et al., 2003). Both pseudoviruses entered cells equally well **(Fig. 1 A)**, suggesting that SARS-CoV-2 S-MLV could potentially use African green monkey ACE2 as entry receptor. To confirm these results, we evaluated entry into BHK cells and observed that transient transfection with hACE2 rendered them susceptible to transduction with SARS-CoV-2 S-MLV **(Fig. 1 B)**. These results demonstrate hACE2 is a functional receptor for SARS-CoV-2, in agreement with recently reported findings (Hoffmann et al., 2020; Letko and Munster, 2020; Zhou et al., 2020).

Sequence analysis of SARS-CoV-2 S reveals the presence of a four amino acid residue insertion at the boundary between the S_1 and S_2 subunits compared to SARS-CoV S and SARSr-CoV S **(Fig. 1 C)**. This results in the introduction of a furin cleavage site, a feature conserved among the 103 SARS-CoV-2 isolates sequenced to date but not in the closely related RaTG13 S (Zhou et al., 2020). Using Western blot analysis, we observed that SARS-CoV-2 S was virtually entirely processed at the S_1/S_2 site during biosynthesis in HEK293T cells, presumably by furin in the Golgi compartment **(Fig. 1 D)**. This observation contrasts with SARS-CoV S which was incorporated into pseudovirions largely uncleaved **(Fig. 1 D)**. To study the influence on pseudovirus entry of the SARS-CoV-2 S_1/S_2 furin cleavage site, we designed an S mutant lacking the four amino acid residue insertion and the furin cleavage site by mutating Q_{686}T**NSPRR**AR↓SV_{696} (wildtype SARS-CoV-2 S) to Q_{686}T**IL**R↓SV_{692} (SARS-CoV-2

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

$S_{fur/mut}$). SARS-CoV-2 $S_{fur/mut}$ preserves only the conserved Arg residue at position 994 of wildtype SARS-CoV-2 S thereby mimicking the S_1/S_2 cleavage site of the related SARSr-CoV S CZX21 **(Fig. 1 D).** SARS-CoV-2 $S_{fur/mut}$ is therefore expected to undergo processing at the S_1/S_2 site upon encounter of a target cell, similarly to SARS-CoV S and SARSr-CoV S (i.e. via TMPRSS2 and/or cathepsin L). As expected, SARS-CoV-2 $S_{fur/mut}$-MLV harbored uncleaved S upon budding **(Fig. 1 D)**. The observed transduction efficiency of VeroE6 cells was higher for SARS-CoV-2 $S_{fur/mut}$-MLV than for SARS-CoV-2 S-MLV **Fig. 1 A)** whereas the opposite trend was observed for transduction of hACE2-expressing BHK cells **(Fig. 1 B)**. These results suggest that S_1/S_2 cleavage during S biosynthesis was not necessary for S-mediated entry in the conditions of our experiments **(Fig. 1 C-D)**. We speculate that the detection of a polybasic cleavage site in the fusion glycoprotein of SARS-CoV-2 could putatively expand its tropism and/or enhance its transmissibility, compared to SARS-CoV and SARSr-CoV isolates, due to the near-ubiquitous distribution of furin-like proteases and their reported effects on other viruses (Klenk and Garten, 1994; Millet and Whittaker, 2015; Steinhauer, 1999) .



bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Figure 1. hACE2 is a functional receptor for SARS-CoV-2 S. A. Entry of MLV pseudotyped with SARS-CoV-2 S, SARS-CoV-2 $S_{fur/mut}$ and SARS-CoV S in VeroE6 cells. **B.** Entry of MLV pseudotyped with SARS-CoV-2 S or SARS-CoV-2 $S_{fur/mut}$ in BHK cells transiently transfected with hACE2. The experiments were carried out in triplicate with two independent pseudovirus preparations and a representative experiment is shown. **C.** Sequence alignment of SARS-CoV-2 S with multiple related SARS-CoV and SARSr-CoV S glycoproteins reveals the introduction of an S_1/S_2 furin cleavage site in this novel coronavirus. Identical and similar positions are respectively shown with white or red font. The four amino acid residue insertion at SARS-CoV-2 S positions 690-693 is indicated with periods. The entire sequence alignment is presented in Fig. S1. **D.** Western blot analysis of SARS-CoV-2 S-MLV, SARS-CoV-2 $S_{fur/mut}$-MLV and SARS-CoV S-MLV pseudovirions using an anti-SARS-CoV S_2 antibody.

SARS-CoV-2 recognizes human ACE2 with comparable affinity than SARS-CoV

The binding affinity of SARS-CoV for hACE2 correlate with the overall rate of viral replication in distinct species, transmissibility and disease severity (Guan et al., 2003; Li et al., 2004; Li et al., 2005c; Wan et al., 2020). Indeed, specific S^B mutations enabled efficient binding to hACE2 of SARS-CoV isolates from the three phases of the 2002-2003 epidemic, which were associated with marked disease severity (Consortium, 2004; Kan et al., 2005; Li et al., 2005c; Sui et al., 2004). In contrast, SARS-CoV isolates detected during the brief 2003-2004 re-emergence interacted more weakly with hACE2, but tightly with civet ACE2, and had low pathogenicity and transmissibility (Consortium, 2004; Kan et al., 2005; Li et al., 2005c).

To understand the contribution of receptor interaction to the infectivity of SARS-CoV-2, we characterized engagement of hACE2 by SARS-CoV-2 S^B and SARS-CoV S^B side-by-side. We used biolayer interferometry to study binding kinetics and affinity of the purified hACE2 ectodomain to SARS-CoV-2 S^B and SARS-CoV S^B immobilized at the surface of biosensors. We found that hACE2 bound to SARS-CoV-2 S^B and SARS-CoV S^B with respective equilibrium dissociation constants of 2.9 nM **(Fig. 2 A)** and 7.7 nM **(Fig. 2 B)**, and comparable kinetic rate constants although the off-rate was slightly

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

higher for SARS-CoV S^B **(Table 1)**. Previous structural work identified 14 positions that are key for binding of SARS-CoV S^B to hACE2: T402, R426, Y436, Y440, Y442, L472, N473, Y475, N479, Y484, T486, T487, G488 and Y491 (Li et al., 2005a). Analysis of the 104 SARS-CoV-2 genome sequences available from GISAID (Elbe and Buckland-Merrett, 2017) shows that 8 out of these 14 positions are strictly conserved in SARS-CoV-2 S^B whereas the other 6 positions are (semi)-conservatively substituted: $R426_{SARS\text{-}CoV}N448_{SARS\text{-}CoV\text{-}2}$, $Y442_{SARS\text{-}CoV}L464_{SARS\text{-}CoV\text{-}2}$, $L472_{SARS\text{-}CoV}F495_{SARS\text{-}CoV\text{-}2}$, $N479_{SARS\text{-}CoV}Q502_{SARS\text{-}CoV\text{-}2}$, $Y484_{SARS\text{-}CoV}Q507_{SARS\text{-}CoV\text{-}2}$, $T487_{SARS\text{-}CoV}N510_{SARS\text{-}CoV\text{-}2}$ **(Fig. 2 C)**. The conservation of many key contact residues could explain the similar binding affinities of SARS-CoV-2 S^B and SARS-CoV S^B for hACE2. No mutations of residues predicted to contact hACE2 have been observed among SARS-CoV-2 S sequences available to date. However, we note that SARSr-CoV ZXC21 and ZC45, which harbor the most closely related S sequences after SARSr-CoV RaTG13, comprise two deletions in the RBD that could affect binding to hACE2 (between SARS-CoV-2 S residues 482-489 and 494-499) **(Fig. S 1)**. Collectively, these results suggest that SARS-CoV-2 is at least as well adapted to the hACE2 orthologue as the 2002-2003 epidemic strains of SARS-CoV, which could explain the efficient transduction efficiency mediated by their respective S glycoproteins **(Fig. 1 A-B)** and the current rapid SARS-CoV-2 transmission in humans.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

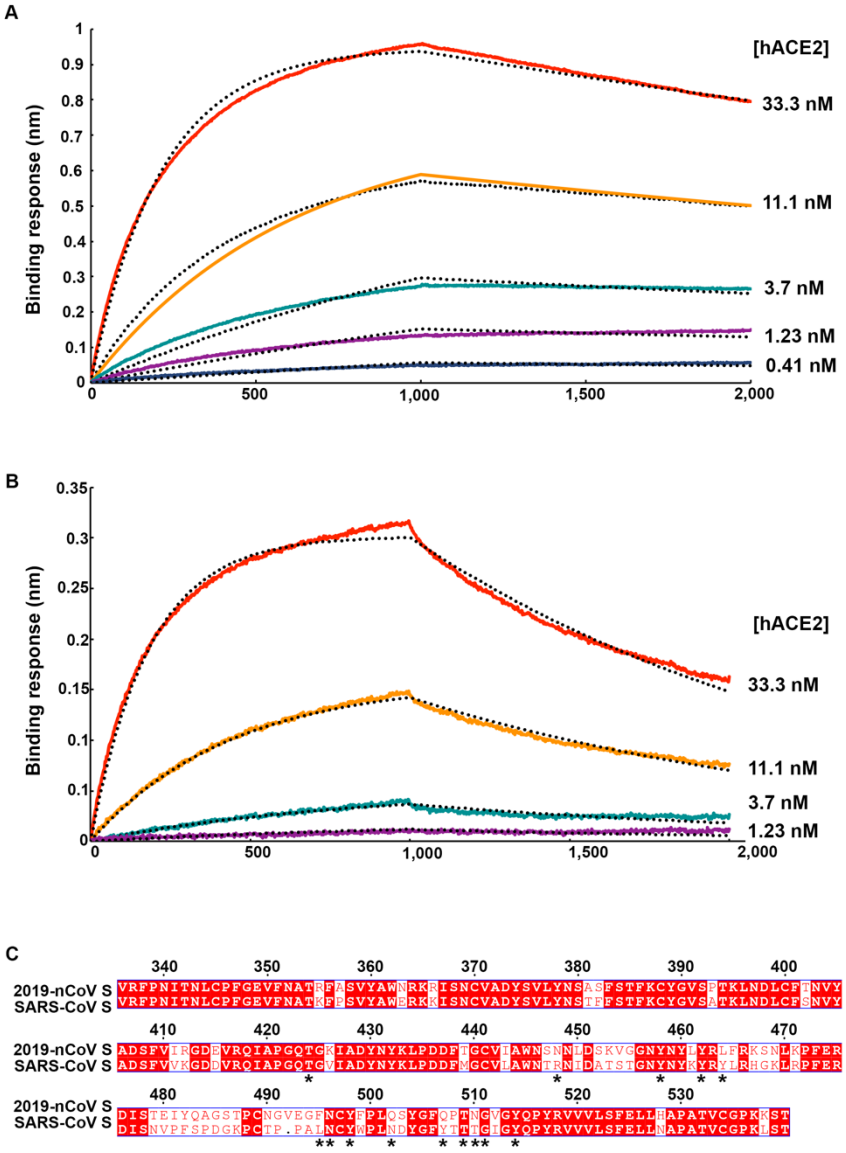


Figure 2. SARS-CoV-2 S recognizes hACE2 with comparable affinity to SARS-CoV S. A-B. Biolayer interferometry binding analysis of the hACE2 ectodomain to immobilized SARS-CoV-2 S^B (B) or SARS-CoV S^B (C). The experiments were performed in duplicates with different protein preparations and one representative set of curves is shown. Dotted lines correspond to a global fit of the data using a 1:1 binding model. **C.** Sequence alignment of SARS-CoV-2 S^B and SARS-CoV S^B Urbani (late phase of the 2002-2003 SARS-CoV epidemic). Identical and similar positions are

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

respectively shown with white or red font. The single amino acid insertion at position 492 of the SARS-CoV-2 S^B is indicated with a period at the corresponding SARS-CoV S^B position. The 14 residues that are key for binding of SARS-CoV S^B to hACE2 are labeled with a star.

Architecture of the SARS-CoV-2 spike glycoprotein trimer

To enable single-particle cryoEM study of the SARS-CoV-2 S glycoprotein, we designed a prefusion stabilized ectodomain trimer construct with an abrogated furin S_1/S_2 cleavage site (Tortorici et al., 2019; Walls et al., 2017a; Walls et al., 2016a; Walls et al., 2019), two consecutive proline stabilizing mutations (Kirchdoerfer et al., 2018; Pallesen et al., 2017) and a C-terminal foldon trimerization domain (Miroshnikov et al., 1998). 3D classification of the cryoEM data revealed the presence of multiple conformational states of SARS-CoV-2 S corresponding to distinct organization of the S^B domains within the S_1 apex. Approximately half of the particle images selected correspond to trimers harboring a single S^B domain opened whereas the remaining half was accounted for by closed trimers with the 3 S^B domains closed. The observed conformational variability of S^B domains is reminiscent of observations made with SARS-CoV S and MERS-CoV S trimers although we did not detect trimers with two S^B domains open and the distribution of particles across the S conformational landscape varies among studies (Gui et al., 2017; Kirchdoerfer et al., 2018; Pallesen et al., 2017; Song et al., 2018; Walls et al., 2019; Yuan et al., 2017).

We determined a reconstruction of the closed SARS-CoV-2 S ectodomain trimer at 3.3 Å resolution (applying 3-fold symmetry) and an asymmetric reconstruction of the trimer with a single S^B domain opened at 3.7 Å resolution **(Fig 3 A-H, Fig S2, Table 2)**. The S_2 fusion machinery is the best resolved part of the map whereas the S^A and S_B domains are less well resolved, presumably due to conformational heterogeneity. The final atomic model comprises residues 36-1157, with internal breaks corresponding to flexible regions, and lacks the C-terminal most segment (including the heptad repeat 2) which is not visible in the map, as is the case for all S structures determined to date.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Overall, the SARS-CoV-2 S ectodomain is a 160Å-long trimer with a triangular cross-section, resembling the closely related SARS-CoV S structure **(Fig S3 A-F)** (Gui et al., 2017; Kirchdoerfer et al., 2018; Song et al., 2018; Walls et al., 2019; Yuan et al., 2017).

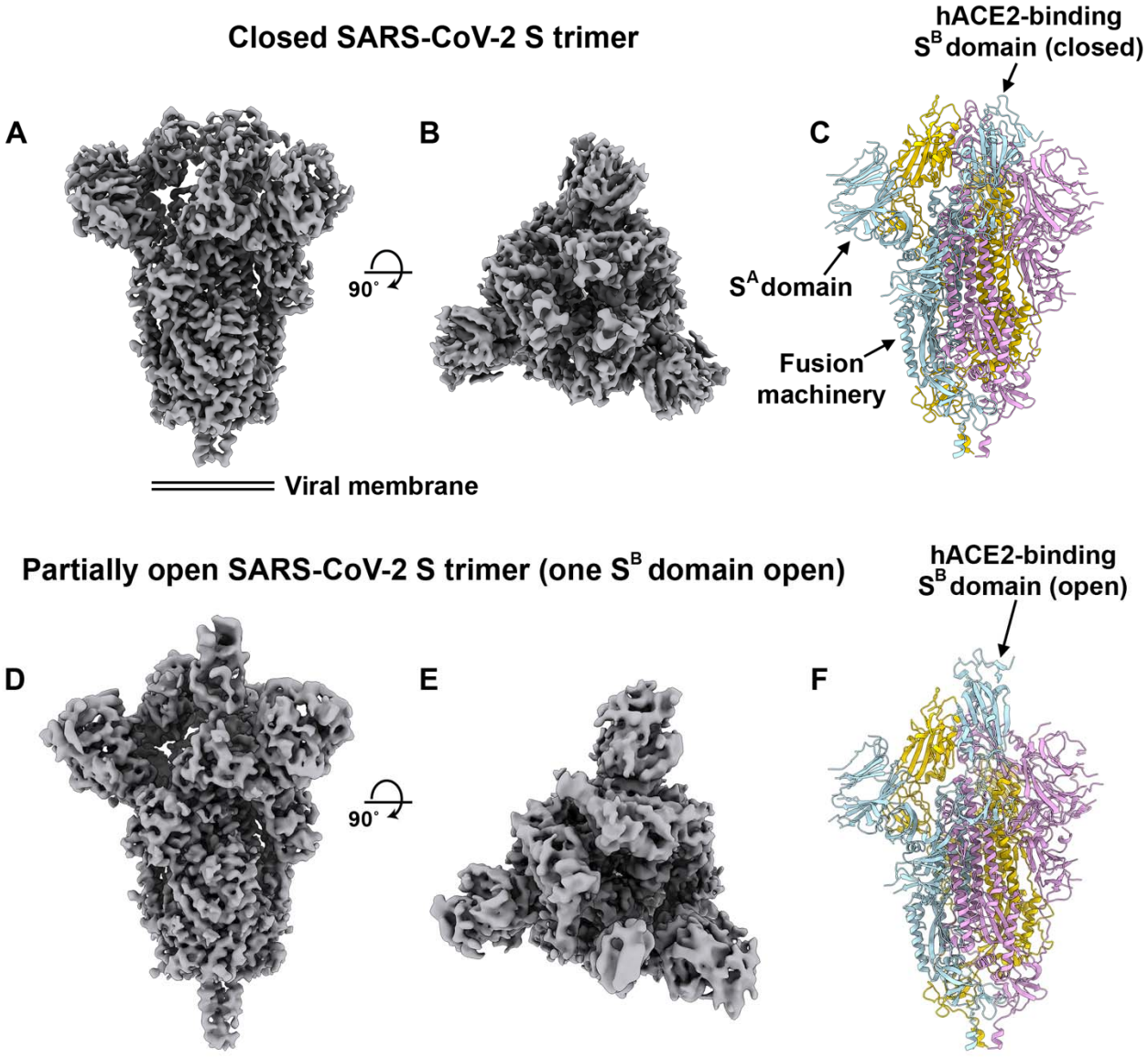
As is the case for other β-coronavirus S glycoproteins, including SARS-CoV S (Gui et al., 2017; Kirchdoerfer et al., 2018; Pallesen et al., 2017; Park et al., 2019; Song et al., 2018; Tortorici et al., 2019; Walls et al., 2016a; Walls et al., 2016b; Walls et al., 2019; Yuan et al., 2017), the S_1 subunit has a V-shaped architecture **(Fig S3 B-E)**. In the closed S trimer, the three hACE2-recognition motifs, whose location was inferred based on the crystal structure of SARS-CoV S^B in complex with hACE2 (Li et al., 2005a), are buried at the interface between protomers. As a result, SARS-CoV-2 S^B opening is expected to be necessary for interacting with ACE2 at the host cell surface and initiating the conformational changes leading to cleavage of the S_2' site, membrane fusion and viral entry (Gui et al., 2017; Kirchdoerfer et al., 2018; Millet and Whittaker, 2014; Pallesen et al., 2017; Park et al., 2016; Song et al., 2018; Walls et al., 2019; Yuan et al., 2017).

As the SARS-CoV-2 and SARS-CoV S_2 subunits share 88% sequence identity, they are structurally conserved and can be superimposed with 1.2 Å rmsd over 417 aligned Cα positions **(Fig S3 E-F)**. Only the most N-terminal part of the fusion peptide is resolved in the map, as was the case for previously determined SARS-CoV S structures. The sequence and conformational conservation of the fusion peptide region observed across SARS-CoV-2 S and SARS-CoV S suggests that antibodies targeting this functionally important motif might cross-react and neutralize the two viruses as well as related coronaviruses **(Fig S3 G-H)**

We previously showed that coronavirus S glycoproteins are densely decorated by heterogeneous N-linked glycans protruding from the trimer surface (Walls et al., 2016b; Walls et al., 2019; Xiong et al., 2017). These oligosaccharides participate in S folding (Rossen et al., 1998), they affect priming by host proteases (Yang et al., 2015b) and modulate antibody recognition (Pallesen et al., 2017; Walls et al., 2019). SARS-

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

CoV-2 S comprise 22 N-linked glycosylation sequons per protomer and 16 of them are resolved in the cryoEM map **(Fig. 4)**. By comparison, SARS-CoV S possesses 23 N-linked glycosylation sequons per protomer and we previously confirmed experimentally that at least 19 of them are glycosylated (Walls et al., 2019). 20 out of 22 SARS-CoV-2 S N-linked glycosylation sequons are conserved in SARS-CoV S **(Table 3)**. Specifically, 9 out of 13 glycans in the S_1 subunit and all 9 glycans in the S_2 subunit are conserved among SARS-CoV-2 S and SARS-CoV S. Furthermore, S_2 oligosaccharides are mostly conserved across SARSr-CoV S glycoproteins, suggesting that accessibility of the fusion machinery to antibodies will be comparable among these viruses.



bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Figure 3. CryoEM structures of the SARS-CoV-2 S glycoprotein. A-B. Two orthogonal views of the closed SARS-CoV-2 S trimer cryoEM map. **C.** Atomic model of the closed SARS-CoV-2 S trimer in the same orientation as in panel A. **D-E.** Two orthogonal views of the partially open SARS-CoV-2 S trimer cryoEM map (one S^B domain is open). **F.** Atomic model of the closed SARS-CoV-2 S trimer in the same orientation as in panel D. The glycans were omitted for clarity.

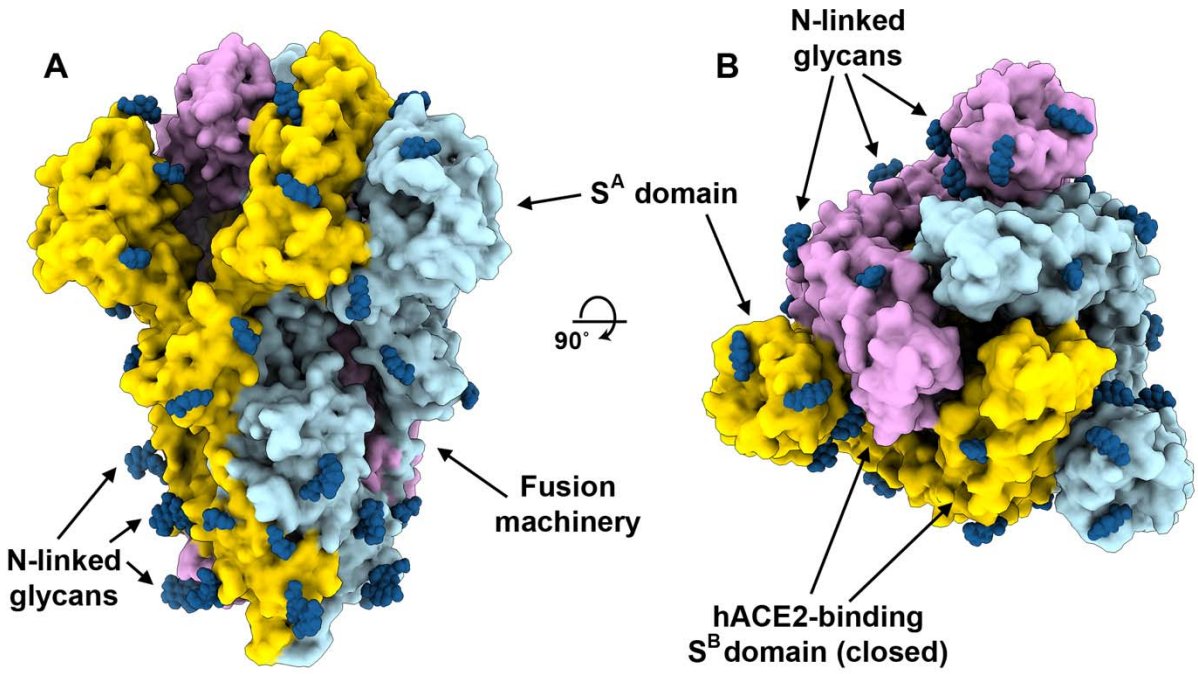


Figure 4. Organization of the SARS-CoV-2 S glycan shield. A-B. Two orthogonal views of the SARS-CoV-2 S structure rendered as a surface with glycans rendered as dark blue spheres.

SARS-CoV S elicits neutralizing antibodies against SARS-CoV-2 S

Mapping of the sequence conservation across multiple sarbecovirus S sequences underscores that the S_2 fusion machinery is more conserved than the S_1 subunit with the highest divergence found within S^A and S^B **(Fig. 5A-B)**. These

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

observations are in line with (i) the fact that some, but not all of these viruses use ACE2 as entry receptor (Ge et al., 2013; Ren et al., 2008; Yang et al., 2015a); and (ii) that the S_1 subunit is more exposed at the viral surface than the fusion machinery, and is likely to be subject to a more stringent selection pressure from the immune system.

Based on these observations, we hypothesized that exposure to one of the two viruses could elicit cross-reactive and potentially neutralizing antibodies against the other virus. We therefore investigated the ability of plasma from four mice immunized with a stabilized SARS-CoV S to inhibit SARS-CoV-2 S- and SARS-CoV S-mediated entry into target cells. All sera tested completely inhibited transduction of SARS-CoV S-MLV and reduced SARS-CoV-2 S-MLV transduction to ~10% of control in VeroE6 cells **(Fig. 5C)**. The elicitation of a heterotypic response blocking SARS-CoV-2 S-mediated entry into host cells concurs with the sequence and structural conservation of SARS-CoV-2 S and SARS-CoV S along with their comparable glycans shields, and suggests that immunity against one virus of the sarbecovirus subgenus can potentially provide protection against related viruses.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

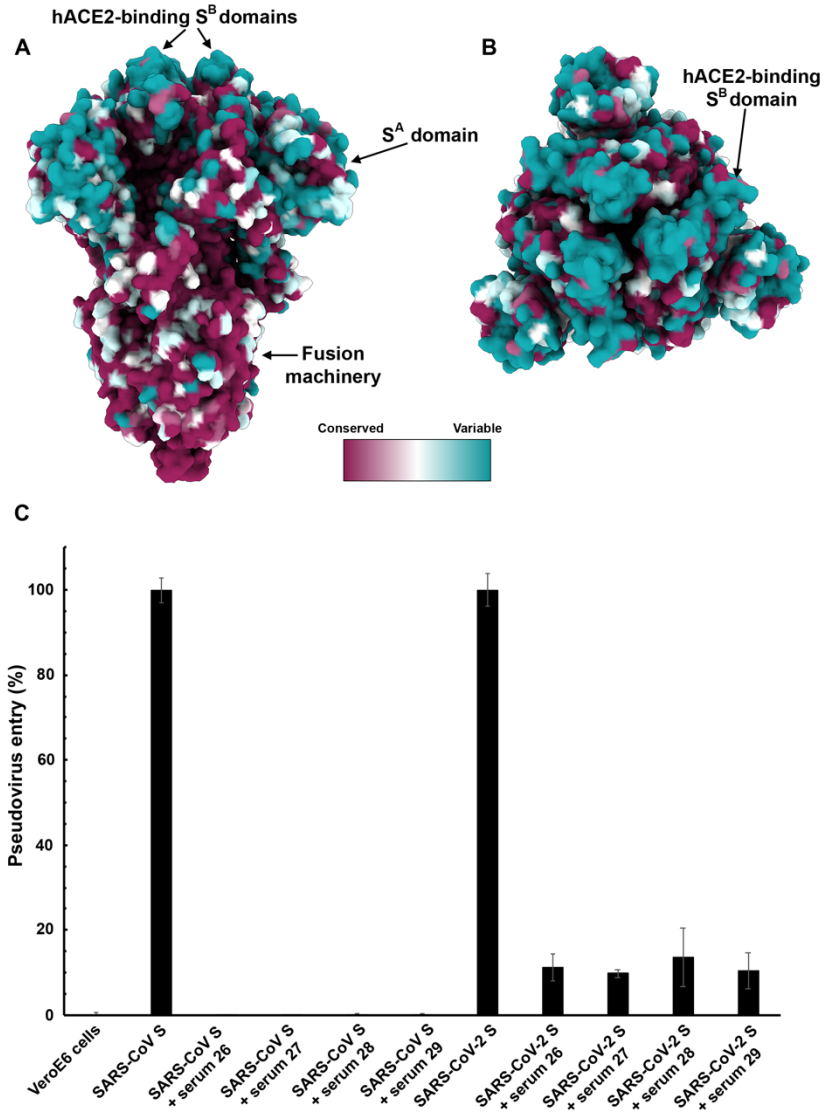


Figure 5. SARS-CoV S elicits antibodies neutralizing SARS-CoV-2 S mediated entry into host cells. A-B. Sequence conservation of sarbecovirus S glycoproteins plotted on the SARS-CoV-2 S structure. **C.** Entry of SARS-CoV-2 and SARS-CoV S pseudotyped viruses is potently inhibited by four SARS-CoV S mouse polyclonal immune plasma.

DISCUSSION

Receptor recognition is the first step of viral infection and is a key determinant of host cell and tissue tropism. Enhanced binding affinity between SARS-CoV S and hACE2 was proposed to correlate with increased virus transmissibility and disease

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

severity in humans (Li et al., 2005c). Indeed, SARS-CoV isolates from the three phases of the 2002-2003 epidemic were more efficiently transmitted among humans and more pathogenic than the isolates associated with the 2003-2004 re-emergence that caused only a few cases, in line with their binding affinities for hACE2 (Consortium, 2004; Kan et al., 2005; Li et al., 2005c). Moreover, the ability to engage ACE2 from different animal species appears to reflect host susceptibility to SARS-CoV infection and facilitated the jump of the virus from animals to humans (Li, 2008; Li et al., 2004). We report here that SARS-CoV-2 uses hACE2 as entry receptor to which it binds with similar affinity than the 2002-2003 SARS-CoV isolates which suggests it can spread efficiently in humans, in agreement with the numerous SARS-CoV-2 human-to-human transmission events reported to date.

Besides binding to host cell receptors, priming of the S glycoprotein by host proteases through cleavage at the S_1/S_2 and the S_2' sites is another crucial factor modulating tropism and pathogenicity (Millet and Whittaker, 2015). For instance, entry of the MERS-CoV-related bat coronavirus HKU4 into human cells required addition of exogenous trypsin, indicating that S proteolytic activation of this bat virus did not occur in human cells (despite its ability to recognize human DPP4)(Wang et al., 2014; Yang et al., 2014). Subsequent work suggested that a glycan present near the S_1/S_2 boundary accounted for the lack of proteolytic priming of HKU4 S and that its removal enhanced pseudovirus entry in human cells (Yang et al., 2015b). The presence of a polybasic cleavage site, that can be processed by furin-like proteases, is a signature of several highly pathogenic avian influenza viruses and pathogenic Newcastle disease virus (Klenk and Garten, 1994; Steinhauer, 1999). Strikingly, SARS-CoV-2 S harbors a furin cleavage site at the S_1/S_2 boundary, which is processed during biosynthesis. The presence of a furin cleavage site sets SARS-CoV-2 S apart from SARS-CoV S (and SARSr-CoV S) that possesses a monobasic S_1/S_2 cleavage site processed upon entry of target cells (Belouzard et al., 2009; Bosch et al., 2008; Glowacka et al., 2011; Matsuyama et al., 2010; Millet and Whittaker, 2015; Shulla et al., 2011). We speculate that the almost ubiquitous expression of furin-like proteases could participate in

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

expanding SARS-CoV-2 cell and tissue tropism, relative to SARS-CoV, as well as increase its transmissibility and/or alter its pathogenicity.

We previously suggested coronaviruses use conformational masking and glycan shielding to limit recognition by the immune response of infected hosts (Walls et al., 2016b; Walls et al., 2019; Xiong et al., 2017). Similarly to SARS-CoV S and MERS-CoV S (Gui et al., 2017; Kirchdoerfer et al., 2018; Pallesen et al., 2017; Song et al., 2018; Walls et al., 2019; Yuan et al., 2017), we found that the SARS-CoV-2 S trimer exist in multiple, distinct conformational states resulting from S^B opening at the trimer apex. These structural changes are necessary for receptor engagement of these three viruses and lead to initiation of fusogenic conformational changes (Song et al., 2018; Walls et al., 2017b; Walls et al., 2019). In contrast, only closed S trimers have been detected for HCoV-NL63 (Walls et al., 2016b), HCoV-OC43 (Tortorici et al., 2019), HCoV-HKU1 (Kirchdoerfer et al., 2016) and HCoV-229E (Li et al., 2019). As HCoV-NL63 and HCoV-229E are known to engage protein receptors through S^B (Wong et al., 2017; Wu et al., 2009), trimer opening is also expected to occur to expose their receptor-recognition motifs that are otherwise buried at the interface between protomers in the closed S trimer (Li et al., 2019; Walls et al., 2016b). Regardless of the nature of the receptor and of the location of the receptor-binding domains, removal of the trimeric S_1 crown is expected to be necessary for all coronaviruses to allow the large-scale S_2 conformational changes leading to fusion of the viral and host membranes (Walls et al., 2017b). Collectively, these data underscores that S glycoprotein trimers found in highly pathogenic human coronaviruses appear to exist in partially opened states whereas they remain largely closed in human coronaviruses associated with common colds. Based on the aforementioned data correlating the binding affinity of SARS-CoV for hACE2 with the rate of transmissibility, viral replication in distinct species and disease severity (Guan et al., 2003; Li et al., 2004; Li et al., 2005c; Wan et al., 2020), we hypothesize that the most pathogenic coronaviruses will exhibit S glycoprotein trimers spontaneously sampling closed and open conformations, as is the case for SARS-CoV-2, SARS-CoV and MERS-CoV.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

The striking structural similarity and sequence conservation among the SARS-CoV-2 S and SARS-CoV S glycoproteins emphasize the close relationship between these two viruses that recognize hACE2 to enter target cells. This resemblance is further strengthened by our finding that SARS-CoV S elicited polyclonal Ab responses potently neutralizing SARS-CoV-2 S-mediated entry into cells. We surmise most of these Abs target the highly conserved S_2 subunit (including the fusion peptide region) based on its structural similarity across SARS-CoV-2 and SARS-CoV and previous reports showing that sera from SARS-CoV-infected individuals target this region (Zhang et al., 2004). We note that most SARS-CoV neutralizing Abs isolated to date target the S^B domain and several of them recognize the RBM and prevent receptor engagement (Hwang et al., 2006; Rockx et al., 2008; Rockx et al., 2010; Traggiai et al., 2004; Walls et al., 2019). As the SARS-CoV-2 and SARS-CoV S^B domains share 75% amino acid sequence identity, future work will be necessary to evaluate how many of these antibodies neutralize the newly emerged coronavirus. These findings also indicate that it might be difficult to distinguish exposure to SARS-CoV-2 from other SARSr-CoVs in serological studies using S ectodomain trimers and that specific assays will need to be designed. Our results provide a structural framework to identify conserved and accessible epitopes across S glycoproteins that will support ongoing vaccine design efforts. Finally, elicitation of diverse, polyclonal Ab responses might prove key in light of the diversity of viruses circulating in animal reservoirs and to prevent the possible emergence of viral neutralization escape mutants.

ACKNOWLEDGEMENTS

This study was supported by the National Institute of General Medical Sciences (R01GM120553, D.V.), the National Institute of Allergy and Infectious Diseases (HHSN272201700059C, DV), a Pew Biomedical Scholars Award (D.V.), an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund (D.V.), the Open Philantropic Foundation (D.V.) and the Pasteur Institute (M.A.T.). We are grateful to Lynda Stuart for sharing the full-length hACE2 plasmid, to Gary Whittaker for the MLV pseudotyping system and to Ning Zheng for providing access to the Octet device.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

DECLARATION OF INTERESTS

The other authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

A.C.W., Y.J.P, A.T.M, and D.V. designed the experiments. A.C.W. and M.A.T. expressed and purified the proteins. A.C.W. carried out binding assays and pseudovirus entry assays. Y.J.P. prepared samples for cryoEM and collected the data. Y.J.P. and D.V. processed the data, built and refined the atomic models. A.W. and A.T.M. immunized mice. A.C.W., Y.J.P and D.V. analyzed the data and prepared the manuscript with input from all authors.

Table 1. Kinetic analysis of hACE2 binding to SARS-CoV-2 S^B and SARS-CoV S^B by biolayer interferometry. Values reported represent the global fit to the data shown in Fig. 2 A-B and the averages obtained from 5 (SARS-CoV-2) or 4 (SARS-CoV) replicates carried out with different protein preparations.

	SARS-CoV-2 S^B	SARS-CoV-2 S^B Average	SARS-CoV S^B	SARS-CoV S^B Average
K_D (nM)	1.2 +/- 0.1	2.9 +/- 3.6	4.9 +/- 0.1	7.7 +/- 8.8
k_{on}	1.4 x 10^5	2.3 +/- 1.4 x 10^5	1.4 x 10^5	1.7 +/- 7.1 x 10^5

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

($M^{-1}.s^{-1}$)				
k_{off} (s^{-1})	1.6 x 10^{-4}	1.7 +/- 0.84 x 10^{-4}	7.61x 10^{-4}	8.7 +/- 5.1 x 10^{-4}

Table 2. CryoEM data collection and refinement statistics

	SARS-CoV-2 S (one S^{B} open)	SARS-CoV-2 S (pseudo-closed)
Data collection and processing		
Magnification	130,000	130,000
Voltage (kV)	300	300
Electron exposure (e^{-}/Å^{2})	70	70
Defocus range (μm)	0.5-3.0	0.5-3.0
Pixel size (Å)	0.525	0.525
Symmetry imposed	C1	C1
Final particle images (no.)	115,591	99,530
Map resolution (Å)	3.3	3.8
FSC threshold	0.143	0.143
Map sharpening *B* factor (Å^{2})	-136.8	-136.8
Validation		
MolProbity score	1.23	
Clashscore	1.66	
Poor rotamers (%)	0.24	
Ramachandran plot		
Favored (%)	97.57	
Allowed (%)	99.76	
Disallowed (%)	0.24	
EMRinger Score	2.75	

Table 3. Conservation of N-linked glycosylation sequons in SARS-CoV-2 S and SARS-CoV S. Red indicates the absence of a glycosylation sequon and deletions are indicated with periods. Glycans observed in the SARS-CoV-2 S cryoEM map are underlined.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

SARS-CoV-2 S	SARS-CoV S
N_{26}LT	T_{21}FD
P34PA	N_{29}YT
N_{70}VT	N_{65}VT
H_{78}VS	N_{73}HT
N_{83}GT	. . .
S_{121}KT	N_{109}KS
N_{130}NA	N_{118}NS
N_{131}AT	N_{119}ST
N_{158}KS	T_{146}QT
N_{174}CT	N_{158}CT
N_{243}IT	N_{227}IT
N_{291}GT	N_{269}GT
N_{340}IT	N_{318}IT
N_{352}AT	N_{330}AT
N_{379}SA	N_{357}ST
N_{612}TS	N_{589}AS
N_{625}CT	N_{602}CT
N_{666}NS	D_{643}TS
N_{718}NS	N_{691}NT
N_{726}FT	N_{699}FS
N_{810}FS	N_{783}FS
N_{1083}FT	N_{1056}FT
N_{1107}GT	N_{1080}GT
N_{1143}NT	N_{1116}NT
N_{1167}HT	N_{1140}HT
N_{1182}AS	N_{1155}AS
N_{1203}ES	N_{1176}ES

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

MATERIALS AND METHODS

Transient expression of SARS-CoV-2 and SARS-CoV S^B

The SARS-CoV S^B construct was cloned from a SARS-CoV S ectodomain (Walls et al., 2019) synthesized by GeneArt (ThermoFisher Scientific) into a modified pOPING vector with an N-terminal mu-phosphatase signal peptide and a C-terminal hexa-histidine tag (G-HHHHHH). The boundaries of the construct are N-terminal $_{306}$RVVPSG$_{311}$ and C-terminal $_{571}$LDISP$_{5975}$ The SARS-CoV-2 S^B construct was synthesized by GenScript into pcDNA3.1- with an N-terminal mu-phosphatase signal peptide and a C-terminal hexa-histidine tag (G-HHHHHH). The boundaries of the construct are N-terminal $_{337}$RFPN$_{340}$ and C-terminal $_{539}$STNL$_{542}$. Both constructs were produced in 500mL HEK293F cells grown in suspension using FreeStyle 293 expression medium (Life technologies) at 37˚C in a humidified 8% CO_2 incubator rotating at 130 rpm. The cultures were transfected using 293-Free transfection reagent (Millipore) with cells grown to a density of 1 million cells per mL and cultivated for 3 days. The supernatants were harvested and cells resuspended for another 3 days, yielding two harvests. Proteins were purified from clarified supernatants using a 5mL Cobalt affinity column (Takara), concentrated and flash frozen in a buffer containing 20 mM Tris pH 8.0 and 300 mM NaCl prior to analysis. SDS-PAGE was run to check purity.

Transient expression of hACE2

Expression and purification of human angiotensin-converting enzyme ectodomain (ACE2, residues 1–614) fused to the Fc region of human IgG (hFc) was performed as previously described (Walls et al., 2019) and cleavage of the Fc fragment was carried out with thrombin. Briefly, hACE2 was produced in 500mL HEK293F cells grown in suspension using FreeStyle 293 expression medium (Life technologies) at 37˚C in a humidified 8% CO_2 incubator rotating at 130 rpm. The cultures were transfected using 293-Free transfection reagent (Millipore) with cells grown to a density of 1 million cells per mL and cultivated for 4 days. The supernatants were harvested and cells resuspended for another four days, yielding two harvests. Clarified supernatants were concentrated using Vivaflow tangential filtration cassettes (Sartorius, 10-kDa cut-off) before affinity purification using a Protein A column (GE LifeSciences) followed by gel

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

filtration chromatography using a Superdex 200 10/300 GL column (GE Life Sciences) equilibrated in 20 mM Tris-HCl, pH 8, 100 mM NaCl. The Fc tag was removed by thrombin cleavage in a reaction mixture containing 3 mg of recombinant ACE2-FC ectodomain and 10 µg of thrombin in 10 mMTris-HCl pH8.0 and 20mM $CaCl_2$.The reaction mixture was incubated at 25 °C overnight and re-loaded in a Protein A column to remove uncleaved protein and the Fc tag. The cleaved protein was further purified by gel filtration using a Superdex 200 column 10/300 GL (GE Life Sciences) equilibrated in a buffer containing 20 mM Tris pH 8.0 and 100 mM NaCl. The purified protein was quantified using absorption at 280 nm and concentrated to approximately 1 mg/ml·

Pseudovirus production

MLV-based SARS-CoV S, SARS-CoV-2 S, and SARS-CoV-2 $S_{fur/mut}$ pseudotypes were prepared as previously described (Millet and Whittaker, 2016). HEK293T cells were co-transfected using Lipofectamine 2000 (Life Technologies) with an S encoding-plasmid, an MLV Gag-Pol packaging construct and the MLV transfer vector encoding a luciferase reporter, according to the manufacturer's instructions. Cells were incubated for 5 hours at 37˚C with transfection medium. Cells were then washed with DMEM 2X and DMEM containing 10% FBS was added for 60 hours. The supernatants were then harvested and filtered through 0.45-mm membranes and concentrated with a 30kDa membrane for 10 minutes at 3,000 rpm and then frozen at -80˚C.

Western Blotting

SARS-CoV S, SARS-CoV-2 S, and SARS-CoV-2 $S_{Fur/Mut}$ pseudovirions were thawed and 4X SDS loading buffer was added prior to boiling for 10 min at 95˚C. Samples were run on a 4%–20% gradient Tris-Glycine Gel (BioRad) and transferred to a PVDF membrane. An anti-S_2 SARS-CoV S monoclonal primary antibody (1:150 dilution) or an MLV p30 monoclonal antibody (1:1,000 dilution, AbCam) and an Alexa Fluor 680-conjugated goat anti-human or mouse secondary antibody (1:30,000 and 1:15,000 dilution respectively, Jackson Laboratory) were used for Western-blotting. A LI-COR processor was used to develop images.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Pseudovirus entry assays

VeroE6 and Baby Hamster Kidney (BHK) cells were cultured in 10% FBS, 1% PenStrep DMEM. BHK or VeroE6 cells were plated into 12 well plates at a density of 0.3 x 10^6 for 16 hours. BHK cells were either not transfected or transfected with 0.8 µg ACE2 per well using standard lipofectamine2000 (Life Technologies) protocols and incubated for another 16 hours. Concentrated pseudovirus was added to the wells after washing 3X with DMEM. After 2-3 hours 20% FBS and 2% PenStrep containing DMEM was added to the cells for 48 hours. Following the 48 hr infection, One-Glo-EX (Promega) was added to the cells in equivalent culturing volume and incubated in the dark for 10 minutes prior to reading on a Varioskan LUX plate reader (ThermoFisher). Measurements were done in triplicate and relative luciferase units (RLU) were plotted.

Biolayer interferometry

Assays were performed on an Octet Red (ForteBio) instrument at 30˚C with shaking at 1,000 RPM. Anti-penta His biosensors were hydrated in water for 30 minutes prior to a 1 min incubation in 10X kinetics buffer (undiluted). Either SARS-CoV-2 S^B or SARS-CoV S^B were loaded at 10-20 µg/mL in 10X Kinetics Buffer for 200-900 seconds prior to baseline equilibration for 180 seconds in 10X kinetics buffer. Association in 10X kinetics buffer diluted ACE2 at various concentrations in a two-fold dilution series from 20nM to 2.5nM was carried out for 1,000 seconds prior to dissociation at 10X kinetics buffer for 1,000 seconds. The data were baseline subtracted prior to fitting performed using a 1:1 binding model and the ForteBio data analysis software. Mean k_{on}, k_{off} values were determined with a global fit applied to all data. The experiments were done with two separate purification batches of SARS-CoV S^B, SARS-CoV-2-S^B, and ACE2 in triplicate and the average is reported.

Protein expression and purification.

The SARS-CoV-2 2P S and SARS-CoV 2P S ectodomains were produced in 500mL HEK293F cells grown in suspension using FreeStyle 293 expression medium (Life technologies) at 37°C in a humidified 8% CO2 incubator rotating at 130 r.p.m. The

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

cultures were transfected using 293fectin (ThermoFisher Scientific) with cells grown to a density of 1 million cells per mL and cultivated for three days. The supernatants were harvested and cells resuspended for another three days, yielding two harvests. Clarified supernatants were purified using a 5mL Cobalt affinity column (Takara). Purified protein was filtered or concentrated and flash frozen in Tris-saline (50 mM Tris pH 8.0, 150 mM NaCl) prior to negative staining and cryo-EM analysis.

Cryo-EM sample preparation and data collection.

3 µL of 0.16 mg/mL SARS-CoV-2 S was loaded onto a freshly glow discharged (30 s at 20 mA) lacey carbon grid with a thin layer of evaporated continuous carbon prior to plunge freezing using a vitrobot MarkIV (ThermoFisher Scientific) using a blot force of -1 and 2.5 second blot time at 100% humidity and 25°C. Data were acquired using the Leginon software (Suloway et al., 2005) to control an FEI Titan Krios transmission electron microscope operated at 300 kV and equipped with a Gatan K2 Summit direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon at a nominal magnification of 130,000x with a pixel size of 0.525 Å. The dose rate was adjusted to 8 counts/pixel/s, and each movie was acquired in super-resolution mode fractionated in 50 frames of 200 ms. 3,000 micrographs were collected in a single session with a defocus range comprised between 1.0 and 2.5 µm.

CryoEM data processing

Movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking and extraction were carried out using Warp (Tegunov and Cramer, 2019). Particle images were extracted with a box size of 800 binned to 400 yielding a pixel size of 1.05 Å. For each data set two rounds of reference-free 2D classification were performed using cryoSPARC (Punjani et al., 2017) to select well-defined particle images. Subsequently, two rounds of 3D classification with 50 iterations each (angular sampling 7.5˚ for 25 iterations and 1.8˚ with local search for 25 iterations) were carried out using Relion (Zivanov et al., 2018) without imposing symmetry to separate distinct SARS-CoV-2 S conformations. 3D refinements were carried out using

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

non-uniform refinement in cryoSPARC (Punjani et al., 2017) using an ab initio model generated using cryoSPARC. Particle images were subjected to Bayesian polishing (Zivanov et al., 2019) before performing another round of on-uniform refinement in cryoSPARC (Punjani et al., 2017) followed by per-particle defocus refinement and again non-uniform refinement. Local resolution estimation, filtering and sharpening was carried out using CryoSPARC. Reported resolutions are based on the gold-standard Fourier shell correlation (FSC) of 0.143 criterion and Fourier shell correlation curves were corrected for the effects of soft masking by high-resolution noise substitution (Scheres and Chen, 2012).

CryoEM model building and analysis.

UCSF Chimera (Goddard et al., 2007) and Coot were used to fit atomic models (PDB 6NB6) into the cryoEM map. The model was subsequently manually rebuilt using Coot and the C3-symmetrized map (Brown et al., 2015; Emsley et al., 2010). This model was then used to build features specific of the C1-map. N-linked glycans were hand-built into the density where visible and the models were refined and relaxed using Rosetta (Wang et al., 2016). Glycan refinement relied on a dedicated Rosetta protocol, which uses physically realistic geometries based on prior knowledge of saccharide chemical properties (Frenz, 2019), and was aided by using both sharpened and unsharpened maps. Models were analyzed using MolProbity (Chen et al., 2010) and privateer (Agirre et al., 2015). Figures were generated using UCSF ChimeraX (Goddard et al., 2018).

Immunizations with SARS-CoV S

57BL/6J mice were obtained from the Jackson Laboratory and maintained at the Comparative Medicine Facility at the Fred Hutchinson Cancer Research Center. After collecting a pre-bleed, mice were immunized with 5µg of SARS-CoV 2P S ectodomain trimer formulated in 50µl of PBS and 50µl of Sigma Adjuvant System subcutaneously in the base of the tail at weeks 0, 4, and 12 and bleeds were collected 2 weeks after each immunization. Plasma from immunized animals was heat inactivated at 56ºC for 1 hour and then stored at 4ºC until use. All experiments were conducted at the Fred

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Hutchinson Cancer Research Center according to approved Institutional Animal Care and Use Committee protocols.

References

Agirre, J., Iglesias-Fernandez, J., Rovira, C., Davies, G.J., Wilson, K.S., and Cowtan, K.D. (2015). Privateer: software for the conformational validation of carbohydrate structures. Nature structural & molecular biology *22*, 833-834.

Anthony, S.J., Gilardi, K., Menachery, V.D., Goldstein, T., Ssebide, B., Mbabazi, R., Navarrete-Macias, I., Liang, E., Wells, H., Hicks, A., *et al.* (2017). Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus. MBio *8*.

Belouzard, S., Chu, V.C., and Whittaker, G.R. (2009). Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proc Natl Acad Sci U S A *106*, 5871-5876.

Bosch, B.J., Bartelink, W., and Rottier, P.J. (2008). Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide. J Virol *82*, 8887-8890.

Bosch, B.J., van der Zee, R., de Haan, C.A., and Rottier, P.J. (2003). The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J Virol *77*, 8801-8811.

Brown, A., Long, F., Nicholls, R.A., Toots, J., Emsley, P., and Murshudov, G. (2015). Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. Acta Crystallogr D Biol Crystallogr *71*, 136-153.

Burkard, C., Verheije, M.H., Wicht, O., van Kasteren, S.I., van Kuppeveld, F.J., Haagmans, B.L., Pelkmans, L., Rottier, P.J., Bosch, B.J., and de Haan, C.A. (2014). Coronavirus cell entry occurs through the endo-/lysosomal pathway in a proteolysis-dependent manner. PLoS Pathog *10*, e1004502.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr *66*, 12-21.

Consortium, C.S.M.E. (2004). Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science *303*, 1666-1669.

Corti, D., Zhao, J., Pedotti, M., Simonelli, L., Agnihothram, S., Fett, C., Fernandez-Rodriguez, B., Foglierini, M., Agatic, G., Vanzetta, F., *et al.* (2015). Prophylactic and postexposure efficacy of a potent human monoclonal antibody against MERS coronavirus. Proc Natl Acad Sci U S A *112*, 10473-10478.

Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H.R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R.A., *et al.* (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med *348*, 1967-1976.

Elbe, S., and Buckland-Merrett, G. (2017). Data, disease and diplomacy: GISAID's innovative contribution to global health. Glob Chall *1*, 33-46.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Frenz, B.D.R., S.; Borst A.J.; Walls, A.C.; Adolf-Bryfogle, J.; Schief, W.R.; Veesler, D.; DiMaio, F. (2019). Automatically fixing errors in glycoprotein structures with Rosetta. Structure *27*, 1-6.

Ge, X.Y., Li, J.L., Yang, X.L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Hu, B., Zhang, W., Peng, C., *et al.* (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature *503*, 535-538.

Glowacka, I., Bertram, S., Müller, M.A., Allen, P., Soilleux, E., Pfefferle, S., Steffen, I., Tsegaye, T.S., He, Y., Gnirss, K., *et al.* (2011). Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. J Virol *85*, 4122-4134.

Goddard, T.D., Huang, C.C., and Ferrin, T.E. (2007). Visualizing density maps with UCSF Chimera. Journal of structural biology *157*, 281-287.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci *27*, 14-25.

Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L., Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., *et al.* (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science *302*, 276-278.

Gui, M., Song, W., Zhou, H., Xu, J., Chen, S., Xiang, Y., and Wang, X. (2017). Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. Cell Res *27*, 119-129.

Haagmans, B.L., Al Dhahiry, S.H., Reusken, C.B., Raj, V.S., Galiano, M., Myers, R., Godeke, G.J., Jonges, M., Farag, E., Diab, A., *et al.* (2014). Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. Lancet Infect Dis *14*, 140-145.

Heald-Sargent, T., and Gallagher, T. (2012). Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. Viruses *4*, 557-580.

Hoffmann, M., Kleine-Weber, H., Krüger, N., Müller, M., Drosten, C., and Pöhlmann, S. (2020). The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells. bioRxiv, 2020.2001.2031.929042.

Hu, B., Zeng, L.P., Yang, X.L., Ge, X.Y., Zhang, W., Li, B., Xie, J.Z., Shen, X.R., Zhang, Y.Z., Wang, N., *et al.* (2017). Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog *13*, e1006698.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., *et al.* (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet.

Hulswit, R.J.G., Lang, Y., Bakkers, M.J.G., Li, W., Li, Z., Schouten, A., Ophorst, B., van Kuppeveld, F.J.M., Boons, G.J., Bosch, B.J., *et al.* (2019). 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.

Hwang, W.C., Lin, Y., Santelli, E., Sui, J., Jaroszewski, L., Stec, B., Farzan, M., Marasco, W.A., and Liddington, R.C. (2006). Structural basis of neutralization by a human anti-severe acute respiratory syndrome spike protein antibody, 80R. J Biol Chem *281*, 34610-34616.

Kan, B., Wang, M., Jing, H., Xu, H., Jiang, X., Yan, M., Liang, W., Zheng, H., Wan, K., Liu, Q., *et al.* (2005). Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. J Virol *79*, 11892-11900.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Kirchdoerfer, R.N., Cottrell, C.A., Wang, N., Pallesen, J., Yassine, H.M., Turner, H.L., Corbett, K.S., Graham, B.S., McLellan, J.S., and Ward, A.B. (2016). Pre-fusion structure of a human coronavirus spike protein. Nature *531*, 118-121.

Kirchdoerfer, R.N., Wang, N., Pallesen, J., Wrapp, D., Turner, H.L., Cottrell, C.A., Corbett, K.S., Graham, B.S., McLellan, J.S., and Ward, A.B. (2018). Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. Sci Rep *8*, 15701.

Klenk, H.D., and Garten, W. (1994). Host cell proteases controlling virus pathogenicity. Trends Microbiol *2*, 39-43.

Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W.*, et al.* (2003). A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med *348*, 1953-1966.

Letko, M., and Munster, V. (2020). Functional assessment of cell entry and receptor usage for lineage B β-coronaviruses, including 2019-nCoV. bioRxiv, 2020.2001.2022.915660.

Li, F. (2008). Structural analysis of major species barriers between humans and palm civets for severe acute respiratory syndrome coronavirus infections. J Virol *82*, 6984-6991.

Li, F., Li, W., Farzan, M., and Harrison, S.C. (2005a). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science *309*, 1864-1868.

Li, W., Greenough, T.C., Moore, M.J., Vasilieva, N., Somasundaran, M., Sullivan, J.L., Farzan, M., and Choe, H. (2004). Efficient replication of severe acute respiratory syndrome coronavirus in mouse cells is limited by murine angiotensin-converting enzyme 2. J Virol *78*, 11429-11433.

Li, W., Hulswit, R.J.G., Widjaja, I., Raj, V.S., McBride, R., Peng, W., Widagdo, W., Tortorici, M.A., van Dieren, B., Lang, Y.*, et al.* (2017). Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. Proc Natl Acad Sci U S A *114*, E8508-E8517.

Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C.*, et al.* (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature *426*, 450-454.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H.*, et al.* (2005b). Bats are natural reservoirs of SARS-like coronaviruses. Science *310*, 676-679.

Li, W., Zhang, C., Sui, J., Kuhn, J.H., Moore, M.J., Luo, S., Wong, S.K., Huang, I.C., Xu, K., Vasilieva, N.*, et al.* (2005c). Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. EMBO J *24*, 1634-1643.

Li, Z., Tomlinson, A.C., Wong, A.H., Zhou, D., Desforges, M., Talbot, P.J., Benlekbir, S., Rubinstein, J.L., and Rini, J.M. (2019). The human coronavirus HCoV-229E S-protein structure and receptor binding. Elife *8*.

Lu, G., Hu, Y., Wang, Q., Qi, J., Gao, F., Li, Y., Zhang, Y., Zhang, W., Yuan, Y., Bao, J.*, et al.* (2013). Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature *500*, 227-231.

Madu, I.G., Roth, S.L., Belouzard, S., and Whittaker, G.R. (2009). Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide. J Virol *83*, 7411-7421.

Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., and Taguchi, F. (2010). Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. J Virol *84*, 12658-12664.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Memish, Z.A., Mishra, N., Olival, K.J., Fagbo, S.F., Kapoor, V., Epstein, J.H., Alhakeem, R., Durosinloun, A., Al Asmari, M., Islam, A., *et al.* (2013). Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. Emerg Infect Dis *19*, 1819-1823.

Menachery, V.D., Yount, B.L., Jr., Debbink, K., Agnihothram, S., Gralinski, L.E., Plante, J.A., Graham, R.L., Scobey, T., Ge, X.Y., Donaldson, E.F., *et al.* (2015). A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. Nat Med *21*, 1508-1513.

Menachery, V.D., Yount, B.L., Jr., Sims, A.C., Debbink, K., Agnihothram, S.S., Gralinski, L.E., Graham, R.L., Scobey, T., Plante, J.A., Royal, S.R., *et al.* (2016). SARS-like WIV1-CoV poised for human emergence. Proc Natl Acad Sci U S A *113*, 3048-3053.

Millet, J.K., and Whittaker, G.R. (2014). Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. Proceedings of the National Academy of Sciences of the United States of America *111*, 15214-15219.

Millet, J.K., and Whittaker, G.R. (2015). Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. Virus Res *202*, 120-134.

Millet, J.K., and Whittaker, G.R. (2016). Murine Leukemia Virus (MLV)-based Coronavirus Spike-pseudotyped Particle Production and Infection. Bio Protoc *6*.

Miroshnikov, K.A., Marusich, E.I., Cerritelli, M.E., Cheng, N., Hyde, C.C., Steven, A.C., and Mesyanzhinov, V.V. (1998). Engineering trimeric fibrous proteins based on bacteriophage T4 adhesins. Protein Eng *11*, 329-332.

Pallesen, J., Wang, N., Corbett, K.S., Wrapp, D., Kirchdoerfer, R.N., Turner, H.L., Cottrell, C.A., Becker, M.M., Wang, L., Shi, W., *et al.* (2017). Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. Proc Natl Acad Sci U S A *114*, E7348-E7357.

Park, J.E., Li, K., Barlan, A., Fehr, A.R., Perlman, S., McCray, P.B., Jr., and Gallagher, T. (2016). Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism. Proc Natl Acad Sci U S A *113*, 12262-12267.

Park, Y.J., Walls, A.C., Wang, Z., Sauer, M.M., Li, W., Tortorici, M.A., Bosch, B.J., DiMaio, F., and Veesler, D. (2019). Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. Nat Struct Mol Biol *26*, 1151-1157.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature methods *14*, 290-296.

Raj, V.S., Mou, H., Smits, S.L., Dekkers, D.H., Muller, M.A., Dijkman, R., Muth, D., Demmers, J.A., Zaki, A., Fouchier, R.A., *et al.* (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature *495*, 251-254.

Ren, W., Qu, X., Li, W., Han, Z., Yu, M., Zhou, P., Zhang, S.Y., Wang, L.F., Deng, H., and Shi, Z. (2008). Difference in receptor usage between severe acute respiratory syndrome (SARS) coronavirus and SARS-like coronavirus of bat origin. J Virol *82*, 1899-1907.

Rockx, B., Corti, D., Donaldson, E., Sheahan, T., Stadler, K., Lanzavecchia, A., and Baric, R. (2008). Structural basis for potent cross-neutralizing human monoclonal antibody protection against lethal human and zoonotic severe acute respiratory syndrome coronavirus challenge. J Virol *82*, 3220-3235.

Rockx, B., Donaldson, E., Frieman, M., Sheahan, T., Corti, D., Lanzavecchia, A., and Baric, R.S. (2010). Escape from human monoclonal antibody neutralization affects in vitro and in vivo fitness of severe acute respiratory syndrome coronavirus. J Infect Dis *201*, 946-955.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Rossen, J.W., de Beer, R., Godeke, G.J., Raamsman, M.J., Horzinek, M.C., Vennema, H., and Rottier, P.J. (1998). The viral spike protein is not involved in the polarized sorting of coronaviruses in epithelial cells. J Virol *72*, 497-503.

Scheres, S.H., and Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. Nature methods *9*, 853-854.

Shulla, A., Heald-Sargent, T., Subramanya, G., Zhao, J., Perlman, S., and Gallagher, T. (2011). A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. J Virol *85*, 873-882.

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

Steinhauer, D.A. (1999). Role of hemagglutinin cleavage for the pathogenicity of influenza virus. Virology *258*, 1-20.

Sui, J., Li, W., Murakami, A., Tamin, A., Matthews, L.J., Wong, S.K., Moore, M.J., Tallarico, A.S., Olurinde, M., Choe, H., *et al.* (2004). 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 *101*, 2536-2541.

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: the new Leginon system. Journal of structural biology *151*, 41-60.

Tegunov, D., and Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. Nature methods.

Tortorici, M.A., and Veesler, D. (2019). Structural insights into coronavirus entry. Adv Virus Res *105*, 93-116.

Tortorici, M.A., Walls, A.C., Lang, Y., Wang, C., Li, Z., Koerhuis, D., Boons, G.J., Bosch, B.J., Rey, F.A., de Groot, R.J., *et al.* (2019). Structural basis for human coronavirus attachment to sialic acid receptors. Nat Struct Mol Biol *26*, 481-489.

Traggiai, E., Becker, S., Subbarao, K., Kolesnikova, L., Uematsu, Y., Gismondo, M.R., Murphy, B.R., Rappuoli, R., and Lanzavecchia, A. (2004). An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. Nat Med *10*, 871-875.

Vlasak, R., Luytjes, W., Spaan, W., and Palese, P. (1988). Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. Proc Natl Acad Sci U S A *85*, 4526-4529.

Walls, A., Tortorici, M.A., Bosch, B.J., Frenz, B., Rottier, P.J., DiMaio, F., Rey, F.A., and Veesler, D. (2017a). Crucial steps in the structure determination of a coronavirus spike glycoprotein using cryo-electron microscopy. Protein Sci *26*, 113-121.

Walls, A.C., Tortorici, M.A., Bosch, B.J., Frenz, B., Rottier, P.J., DiMaio, F., Rey, F.A., and Veesler, D. (2016a). Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. Nature *531*, 114-117.

Walls, A.C., Tortorici, M.A., Frenz, B., Snijder, J., Li, W., Rey, F.A., DiMaio, F., Bosch, B.J., and Veesler, D. (2016b). Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. Nat Struct Mol Biol *23*, 899-905.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Walls, A.C., Tortorici, M.A., Snijder, J., Xiong, X., Bosch, B.-J., Rey, F.A., and Veesler, D. (2017b). Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. Proceedings of the National Academy of Sciences *114*, 11157-11162.

Walls, A.C., Xiong, X., Park, Y.J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., *et al.* (2019). Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. Cell *176*, 1026-1039 e1015.

Wan, Y., Shang, J., Graham, R., Baric, R.S., and Li, F. (2020). Receptor recognition by novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS. J Virol.

Wang, M., Yan, M., Xu, H., Liang, W., Kan, B., Zheng, B., Chen, H., Zheng, H., Xu, Y., Zhang, E., *et al.* (2005). SARS-CoV infection in a restaurant from palm civet. Emerg Infect Dis *11*, 1860-1865.

Wang, Q., Qi, J., Yuan, Y., Xuan, Y., Han, P., Wan, Y., Ji, W., Li, Y., Wu, Y., Wang, J., *et al.* (2014). Bat origins of MERS-CoV supported by bat coronavirus HKU4 usage of human receptor CD26. Cell Host Microbe *16*, 328-337.

Wang, R.Y., Song, Y., Barad, B.A., Cheng, Y., Fraser, J.S., and DiMaio, F. (2016). Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. eLife *5*.

Wong, A.H.M., Tomlinson, A.C.A., Zhou, D., Satkunarajah, M., Chen, K., Sharon, C., Desforges, M., Talbot, P.J., and Rini, J.M. (2017). Receptor-binding loops in alphacoronavirus adaptation and evolution. Nat Commun *8*, 1735.

Wu, K., Li, W., Peng, G., and Li, F. (2009). Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. Proc Natl Acad Sci U S A *106*, 19970-19974.

Xiong, X., Tortorici, M.A., Snijder, J., Yoshioka, C., Walls, A.C., Li, W., McGuire, A.T., Rey, F.A., Bosch, B.J., and Veesler, D. (2017). Glycan shield and fusion activation of a deltacoronavirus spike glycoprotein fine-tuned for enteric infections. J Virol.

Yang, X.L., Hu, B., Wang, B., Wang, M.N., Zhang, Q., Zhang, W., Wu, L.J., Ge, X.Y., Zhang, Y.Z., Daszak, P., *et al.* (2015a). Isolation and Characterization of a Novel Bat Coronavirus Closely Related to the Direct Progenitor of Severe Acute Respiratory Syndrome Coronavirus. J Virol *90*, 3253-3256.

Yang, Y., Du, L., Liu, C., Wang, L., Ma, C., Tang, J., Baric, R.S., Jiang, S., and Li, F. (2014). Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. Proc Natl Acad Sci U S A *111*, 12516-12521.

Yang, Y., Liu, C., Du, L., Jiang, S., Shi, Z., Baric, R.S., and Li, F. (2015b). Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome Coronavirus. J Virol *89*, 9119-9123.

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

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D., and Fouchier, R.A. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med *367*, 1814-1820.

Zhang, H., Wang, G., Li, J., Nie, Y., Shi, X., Lian, G., Wang, W., Yin, X., Zhao, Y., Qu, X., *et al.* (2004). Identification of an antigenic determinant on the S2 domain of the severe acute respiratory syndrome coronavirus spike glycoprotein capable of inducing neutralizing antibodies. J Virol *78*, 6938-6945.

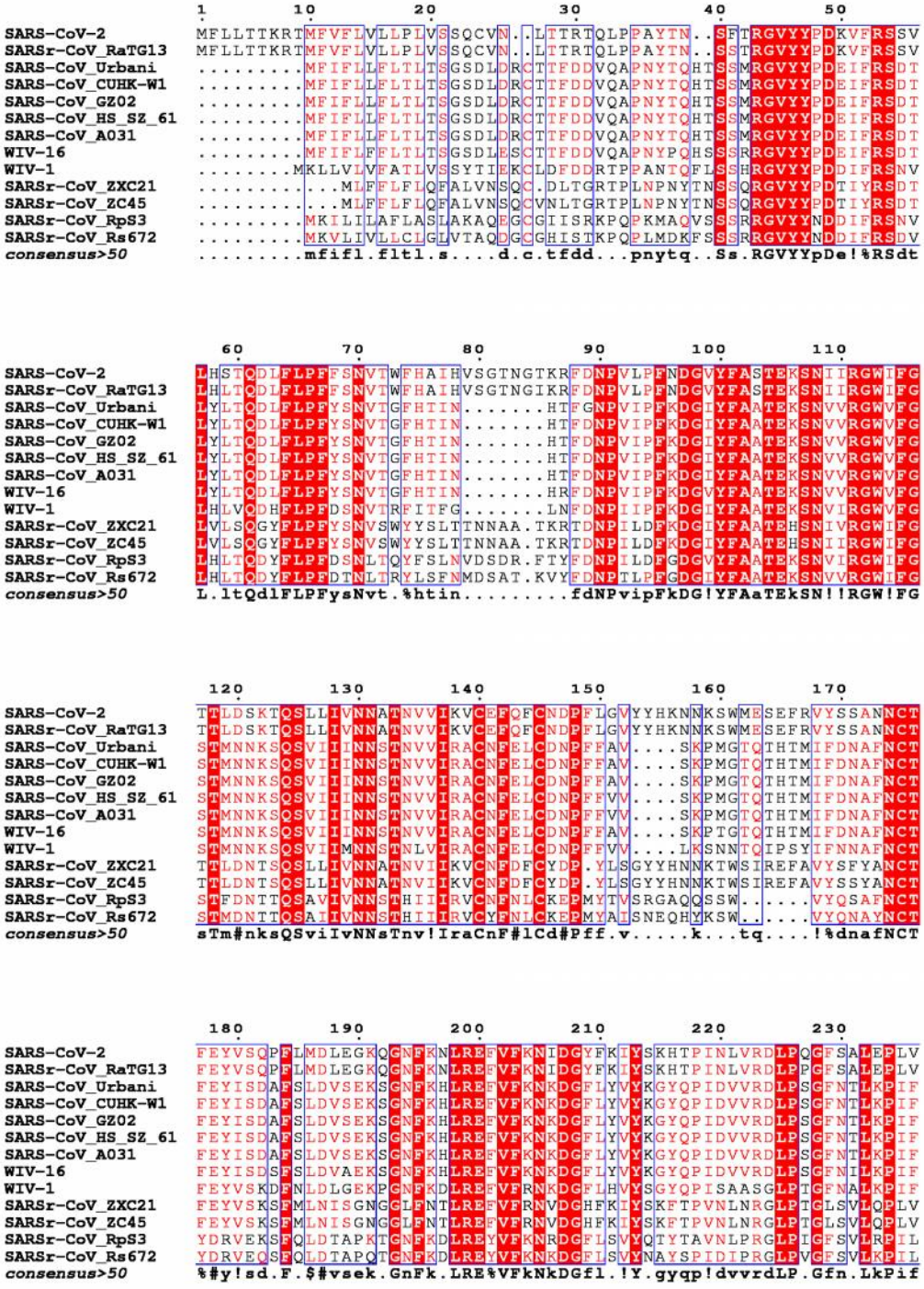
bioRxiv preprint doi: https://doi.org/10.1101/2020.02.19.956581. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

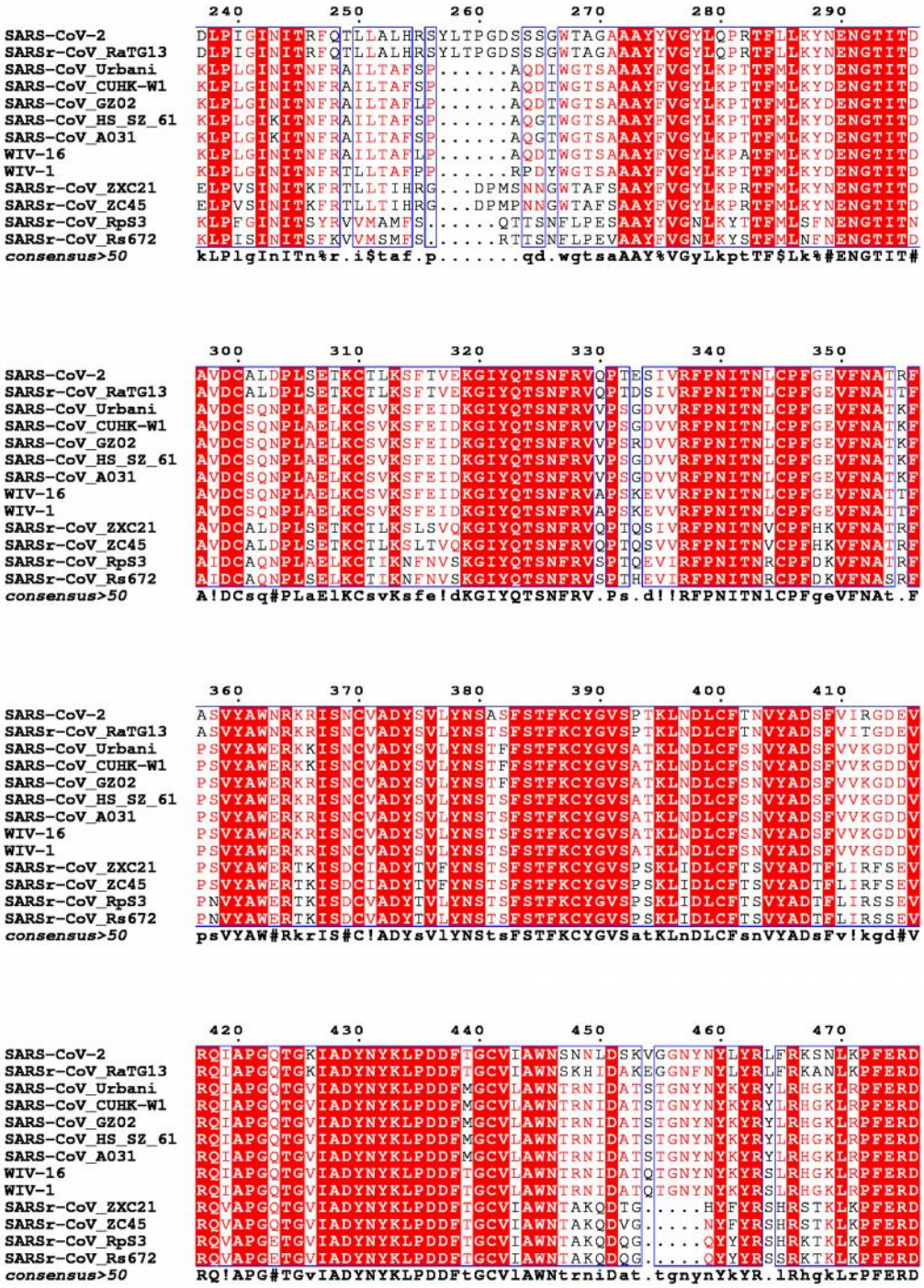
Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., *et al.* (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature.

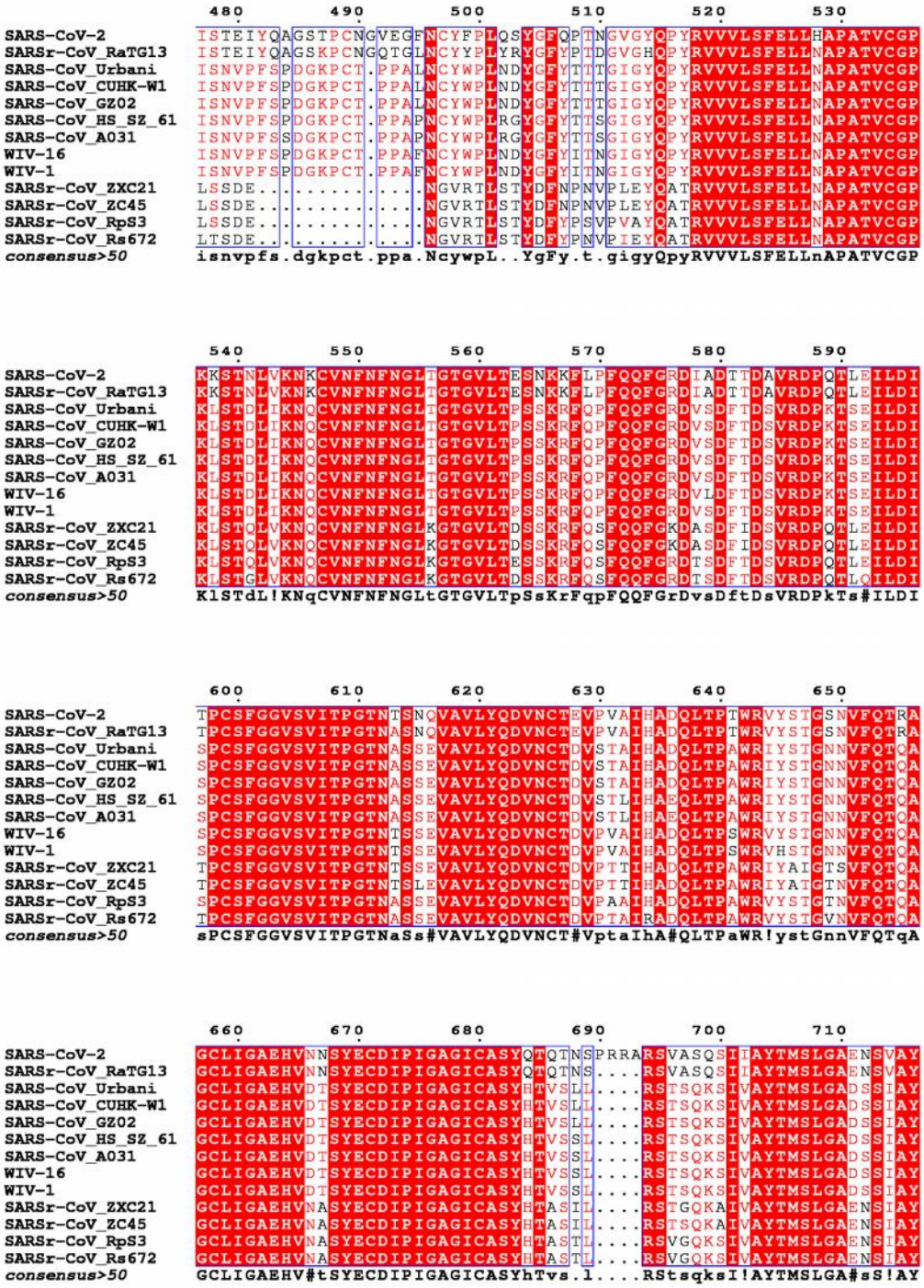
Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., *et al.* (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med.

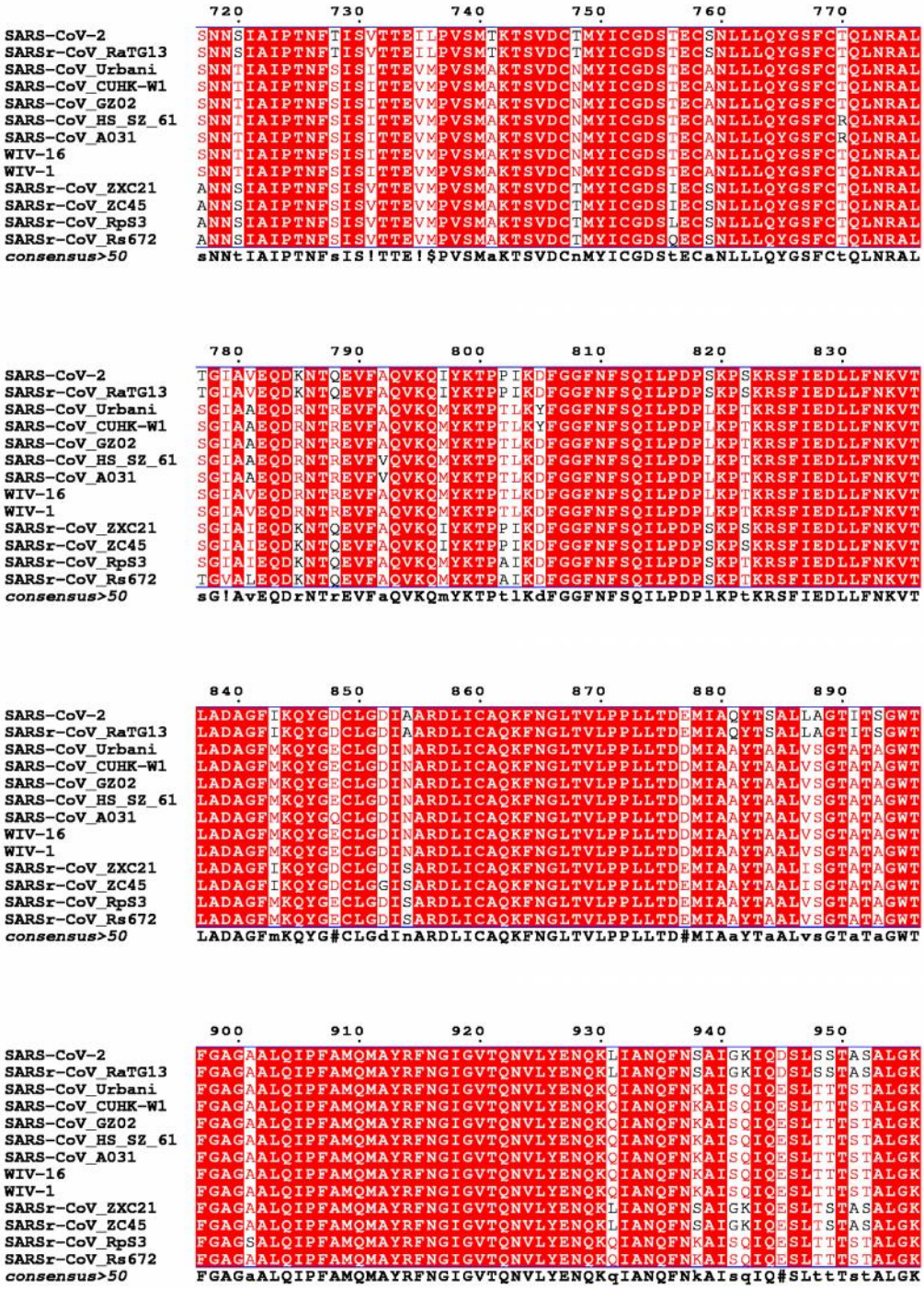
Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*.

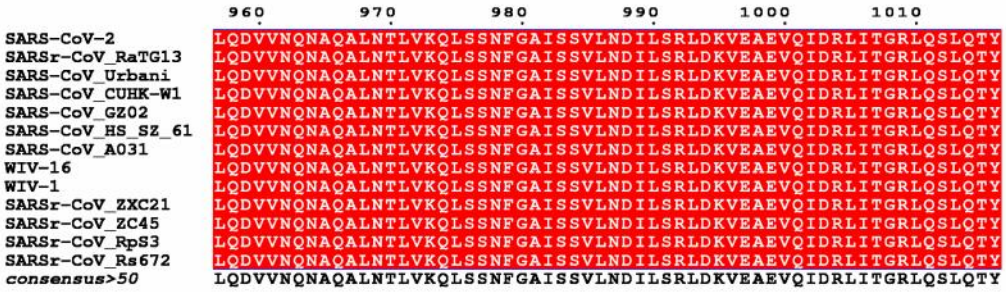
Zivanov, J., Nakane, T., and Scheres, S.H.W. (2019). A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. IUCrJ *6*.

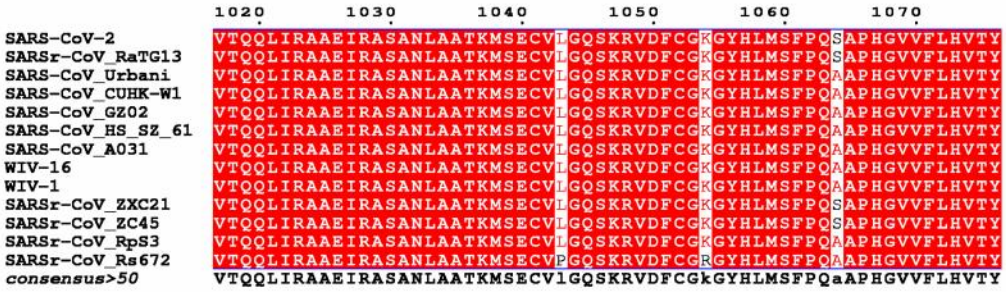

SARS-CoV-2
SARSr-CoV_RaTG13
SARS-CoV_Urbani
SARS-CoV_CUHK-W1
SARS-CoV_GZ02
SARS-CoV_HS_SZ_61
SARS-CoV_A031
WIV-16
WIV-1
SARSr-CoV_ZXC21
SARSr-CoV_ZC45
SARSr-CoV_RpS3
SARSr-CoV_Rs672
consensus>50

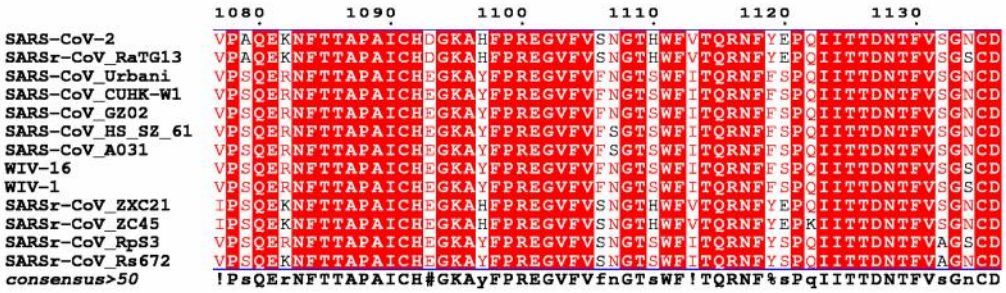

240 250 260 270 280 290
SARS-CoV-2 DLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
SARSr-CoV_RaTG13 DLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
SARS-CoV_Urbani KLPLGINITNFRAILTAFSP......AQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARS-CoV_CUHK-W1 KLPLGINITNFRAILTAFSP......AQDTWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARS-CoV_GZ02 KLPLGINITNFRAILTAFLP......AQDTWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARS-CoV_HS_SZ_61 KLPLGIKITNFRAILTAFSP......AQGTWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARS-CoV_A031 KLPLGIKITNFRAILTAFSP......AQGTWGTSAAAYFVGYLKPTTFMLKYDENGTITD
WIV-16 KLPLGINITNFRAILTAFLP......AQDTWGTSAAAYFVGYLKPATFMLKYDENGTITD
WIV-1 KLPLGINITNFRTLLTAFPP......RPDYWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARSr-CoV_ZXC21 ELPVSINITKFRTLLTIHRG...DPMSNNGWTAFSAAYFVGYLKPRTFMLKYNENGTITD
SARSr-CoV_ZC45 ELPVSINITKFRTLLTIHRG...DPMPNNGWTAFSAAYFVGYLKPRTFMLKYNENGTITD
SARSr-CoV_RpS3 KLPFGINITSYRVVMAMFS.....QTTSNFLPESAAYYVGNLKYTTFMLSFNENGTITN
SARSr-CoV_Rs672 KLPISINITSFKVVMSMFS.....RTTSNFLPEVAAYFVGNLKYSTFMLNFNENGTITD
consensus>50 kLPlgInITn%r.i$taf.p.......qd.wgtsaAAY%VGyLkptTF$Lk%#ENGTIT#
300 310 320 330 340 350
SARS-CoV-2 AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRF
SARSr-CoV_RaTG13 AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTDSIVRFPNITNLCPFGEVFNATTF
SARS-CoV_Urbani AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
SARS-CoV_CUHK-W1 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
SARS-CoV_GZ02 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSRDVVRFPNITNLCPFGEVFNATKF
SARS-CoV_HS_SZ_61 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
SARS-CoV_A031 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
WIV-16 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVAPSKEVVRFPNITNLCPFGEVFNATTF
WIV-1 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVAPSKEVVRFPNITNLCPFGEVFNATTF
SARSr-CoV_ZXC21 AVDCALDPLSETKCTLKSLSVQKGIYQTSNFRVQPTQSIVRFPNITNVCPFHKVFNATRF
SARSr-CoV_ZC45 AVDCALDPLSETKCTLKSLTVQKGIYQTSNFRVQPTQSVVRFPNITNVCPFHKVFNATRF
SARSr-CoV_RpS3 AIDCAQNPLAELKCTIKNFNVSKGIYQTSNFRVSPTQEVIRFPNITNRCPFDKVFNATRF
SARSr-CoV_Rs672 AIDCAQNPLSELKCTIKNFNVSKGIYQTSNFRVSPTHEVIRFPNITNRCPFDKVFNASRF
consensus>50 A!DCsq#PLaElKCsvKsfe!dKGIYQTSNFRV.Ps.d!!RFPNITNlCPFgeVFNAt.F
360 370 380 390 400 410
SARS-CoV-2 ASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEV
SARSr-CoV_RaTG13 ASVYAWNRKRISNCVADYSVLYNSTSFSTFKCYGVSPTKLNDLCFTNVYADSFVITGDEV
SARS-CoV_Urbani PSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARS-CoV_CUHK-W1 PSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARS-CoV_GZ02 PSVYAWERKRISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARS-CoV_HS_SZ_61 PSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARS-CoV_A031 PSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
WIV-16 PSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
WIV-1 PSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARSr-CoV_ZXC21 PSVYAWERTKISDCIADYTVFYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRFSEV
SARSr-CoV_ZC45 PSVYAWERTKISDCIADYTVFYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRFSEV
SARSr-CoV_RpS3 PNVYAWERTKISDCVADYTVLYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRSSEV
SARSr-CoV_Rs672 PNVYAWERTKISDCVADYTVLYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRSSEV
consensus>50 psVYAW#RkrIS#C!ADYsVlYNStsFSTFKCYGVSatKLnDLCFsnVYADsFv!kgd#V
420 430 440 450 460 470
SARS-CoV-2 RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERD
SARSr-CoV_RaTG13 RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSKHIDAKEGGNFNYLYRLFRKANLKPFERD
SARS-CoV_Urbani RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
SARS-CoV_CUHK-W1 RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
SARS-CoV_GZ02 RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
SARS-CoV_HS_SZ_61 RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
SARS-CoV_A031 RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
WIV-16 RQIAPGQTGVIADYNYKLPDDFTGCVLAWNTRNIDATQTGNYNYKYRSLRHGKLRPFERD
WIV-1 RQIAPGQTGVIADYNYKLPDDFTGCVLAWNTRNIDATQTGNYNYKYRSLRHGKLRPFERD
SARSr-CoV_ZXC21 RQVAPGQTGVIADYNYKLPDDFTGCVIAWNTAKQDTG.....HYFYRSHRSTKLKPFERD
SARSr-CoV_ZC45 RQVAPGQTGVIADYNYKLPDDFTGCVIAWNTAKQDVG.....NYFYRSHRSTKLKPFERD
SARSr-CoV_RpS3 RQVAPGETGVIADYNYKLPDDFTGCVIAWNTAKQDQG.....QYYYRSHRKTKLKPFERD
SARSr-CoV_Rs672 RQVAPGETGVIADYNYKLPDDFTGCVIAWNTAKQDQG.....QYYYRSSRKTKLKPFERD
consensus>50 RQ!APG#TGvIADYNYKLPDDFtGCVlAWNtrniDat.tgnynYkYR.lRhgkLrPFERD


480 490 500 510 520 530
SARS-CoV-2 ISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGP
SARSr-CoV_RaTG13 ISTEIYQAGSKPCNGQTGLNCYYPLYRYGFYPTDGVGHQPYRVVVLSFELLNAPATVCGP
SARS-CoV_Urbani ISNVPFSPDGKPCT.PPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGP
SARS-CoV_CUHK-W1 ISNVPFSPDGKPCT.PPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGP
SARS-CoV_GZ02 ISNVPFSPDGKPCT.PPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGP
SARS-CoV_HS_SZ_61 ISNVPFSPDGKPCT.PPAPNCYWPLRGYGFYTTSGIGYQPYRVVVLSFELLNAPATVCGP
SARS-CoV_A031 ISNVPFSSDGKPCT.PPAPNCYWPLRGYGFYTTSGIGYQPYRVVVLSFELLNAPATVCGP
WIV-16 ISNVPFSPDGKPCT.PPAFNCYWPLNDYGFYITNGIGYQPYRVVVLSFELLNAPATVCGP
WIV-1 ISNVPFSPDGKPCT.PPAFNCYWPLNDYGFYITNGIGYQPYRVVVLSFELLNAPATVCGP
SARSr-CoV_ZXC21 LSSDE..........NGVRTLSTYDFNPNVPLEYQATRVVVLSFELLNAPATVCGP
SARSr-CoV_ZC45 LSSDE..........NGVRTLSTYDFNPNVPLEYQATRVVVLSFELLNAPATVCGP
SARSr-CoV_RpS3 LSSDE..........NGVRTLSTYDFYPSVPVAYQATRVVVLSFELLNAPATVCGP
SARSr-CoV_Rs672 LTSDE..........NGVRTLSTYDFYPNVPIEYQATRVVVLSFELLNAPATVCGP
consensus>50 isnvpfs.dgkpct.ppa.NcywpL..YgFy.t.gigyQpyRVVVLSFELLnAPATVCGP
540 550 560 570 580 590
SARS-CoV-2 KKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDI
SARSr-CoV_RaTG13 KKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDI
SARS-CoV_Urbani KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARS-CoV_CUHK-W1 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARS-CoV_GZ02 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARS-CoV_HS_SZ_61 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARS-CoV_A031 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
WIV-16 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVLDFTDSVRDPKTSEILDI
WIV-1 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARSr-CoV_ZXC21 KLSTQLVKNQCVNFNFNGLKGTGVLTDSSKRFQSFQQFGKDASDFIDSVRDPQTLEILDI
SARSr-CoV_ZC45 KLSTQLVKNQCVNFNFNGLKGTGVLTDSSKRFQSFQQFGKDASDFIDSVRDPQTLEILDI
SARSr-CoV_RpS3 KLSTQLVKNQCVNFNFNGLKGTGVLTESSKRFQSFQQFGRDTSDFTDSVRDPQTLEILDI
SARSr-CoV_Rs672 KLSTGLVKNQCVNFNFNGLKGTGVLTDSSKRFQSFQQFGRDTSDFTDSVRDPQTLQILDI
consensus>50 KlSTdL!KNqCVNFNFNGLtGTGVLTpSsKrFqpFQQFGrDvsDftDsVRDPkTs#ILDI
600 610 620 630 640 650
SARS-CoV-2 TPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRA
SARSr-CoV_RaTG13 TPCSFGGVSVITPGTNASNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRA
SARS-CoV_Urbani SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQA
SARS-CoV_CUHK-W1 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQA
SARS-CoV_GZ02 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQA
SARS-CoV_HS_SZ_61 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTLIHAEQLTPAWRIYSTGNNVFQTQA
SARS-CoV_A031 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTLIHAEQLTPAWRIYSTGNNVFQTQA
WIV-16 SPCSFGGVSVITPGTNTSSEVAVLYQDVNCTDVPVAIHADQLTPSWRVYSTGNNVFQTQA
WIV-1 SPCSFGGVSVITPGTNTSSEVAVLYQDVNCTDVPVAIHADQLTPSWRVHSTGNNVFQTQA
SARSr-CoV_ZXC21 TPCSFGGVSVITPGTNTSSEVAVLYQDVNCTDVPTTIHADQLTPAWRIYAIGTSVFQTQA
SARSr-CoV_ZC45 TPCSFGGVSVITPGTNTSLEVAVLYQDVNCTDVPTTIHADQLTPAWRIYATGTNVFQTQA
SARSr-CoV_RpS3 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVPAAIHADQLTPAWRVYSTGTNVFQTQA
SARSr-CoV_Rs672 TPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVPTAIRADQLTPAWRVYSTGVNVFQTQA
consensus>50 sPCSFGGVSVITPGTNaSs#VAVLYQDVNCT#VptaIhA#QLTPaWR!ystGnnVFQTqA
660 670 680 690 700 710
SARS-CoV-2 GCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAY
SARSr-CoV_RaTG13 GCLIGAEHVNNSYECDIPIGAGICASYQTQTNS....RSVASQSIIAYTMSLGAENSVAY
SARS-CoV_Urbani GCLIGAEHVDTSYECDIPIGAGICASYHTVSLL....RSTSQKSIVAYTMSLGADSSIAY
SARS-CoV_CUHK-W1 GCLIGAEHVDTSYECDIPIGAGICASYHTVSLL....RSTSQKSIVAYTMSLGADSSIAY
SARS-CoV_GZ02 GCLIGAEHVDTSYECDIPIGAGICASYHTVSLL....RSTSQKSIVAYTMSLGADSSIAY
SARS-CoV_HS_SZ_61 GCLIGAEHVDTSYECDIPIGAGICASYHTVSSL....RSTSQKSIVAYTMSLGADSSIAY
SARS-CoV_A031 GCLIGAEHVDTSYECDIPIGAGICASYHTVSSL....RSTSQKSIVAYTMSLGADSSIAY
WIV-16 GCLIGAEHVDTSYECDIPIGAGICASYHTVSSL....RSTSQKSIVAYTMSLGADSSIAY
WIV-1 GCLIGAEHVDTSYECDIPIGAGICASYHTVSSL....RSTSQKSIVAYTMSLGADSSIAY
SARSr-CoV_ZXC21 GCLIGAEHVNASYECDIPIGAGICASYHTASIL....RSTGQKAIVAYTMSLGAENSIAY
SARSr-CoV_ZC45 GCLIGAEHVNASYECDIPIGAGICASYHTASIL....RSTSQKAIVAYTMSLGAENSIAY
SARSr-CoV_RpS3 GCLIGAEHVNASYECDIPIGAGICASYHTASTL....RSVGQKSIVAYTMSLGAENSIAY
SARSr-CoV_Rs672 GCLIGAEHVNASYECDIPIGAGICASYHTASTL....RSVGQKSIVAYTMSLGAENSIAY
consensus>50 GCLIGAEHV#tSYECDIPIGAGICASYhTvs.l....RStsqksI!AYTMSLGA#sS!AY

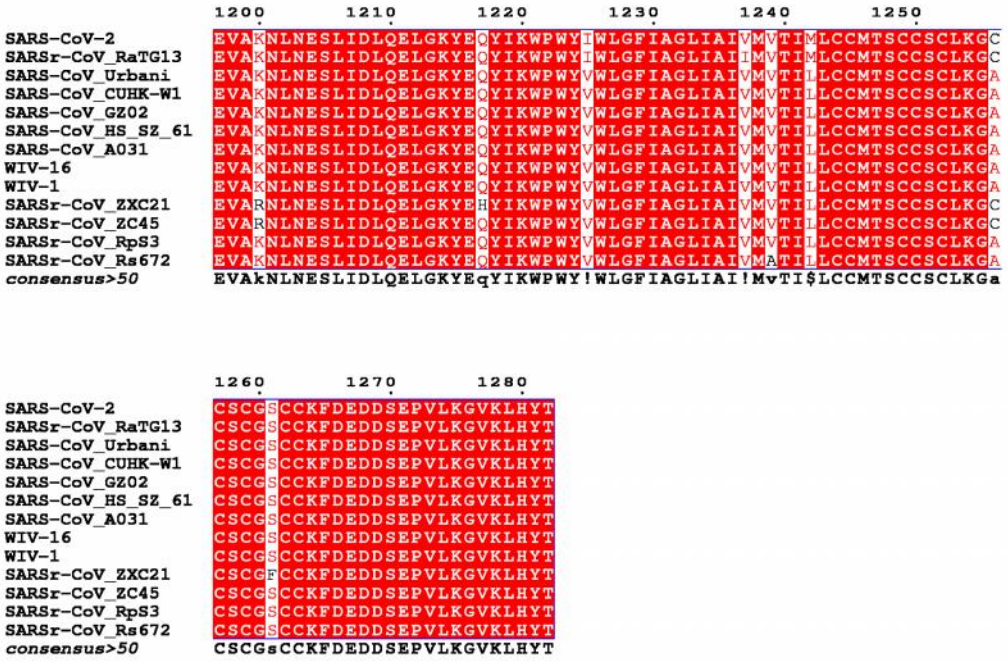

720 730 740 750 760 770
SARS-CoV-2 SNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRAL
SARSr-CoV_RaTG13 SNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRAL
SARS-CoV_Urbani SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SARS-CoV_CUHK-W1 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SARS-CoV_GZ02 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SARS-CoV_HS_SZ_61 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCRQLNRAL
SARS-CoV_A031 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCRQLNRAL
WIV-16 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
WIV-1 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SARSr-CoV_ZXC21 ANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSIECSNLLLQYGSFCTQLNRAL
SARSr-CoV_ZC45 ANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSIECSNLLLQYGSFCTQLNRAL
SARSr-CoV_RpS3 ANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSLECSNLLLQYGSFCTQLNRAL
SARSr-CoV_Rs672 ANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSQECSNLLLQYGSFCTQLNRAL
consensus>50 sNNtIAIPTNFsIS!TTE!$PVSMaKTSVDCnMYICGDStECaNLLLQYGSFCtQLNRAL
780 790 800 810 820 830
SARS-CoV-2 TGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
SARSr-CoV_RaTG13 TGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
SARS-CoV_Urbani SGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARS-CoV_CUHK-W1 SGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARS-CoV_GZ02 SGIAAEQDRNTREVFAQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARS-CoV_HS_SZ_61 SGIAAEQDRNTREVFVQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARS-CoV_A031 SGIAAEQDRNTREVFVQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
WIV-16 SGIAVEQDRNTREVFAQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
WIV-1 SGIAVEQDRNTREVFAQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARSr-CoV_ZXC21 SGIAIEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
SARSr-CoV_ZC45 SGIAIEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
SARSr-CoV_RpS3 SGIAIEQDKNTQEVFAQVKQMYKTPAIKDFGGFNFSQILPDPSKPTKRSFIEDLLFNKVT
SARSr-CoV_Rs672 TGVALEQDKNTQEVFAQVKQMYKTPAIKDFGGFNFSQILPDPSKPTKRSFIEDLLFNKVT
consensus>50 sG!AvEQDrNTrEVFaQVKQmYKTPtlKdFGGFNFSQILPDPlKPtKRSFIEDLLFNKVT
840 850 860 870 880 890
SARS-CoV-2 LADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWT
SARSr-CoV_RaTG13 LADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWT
SARS-CoV_Urbani LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARS-CoV_CUHK-W1 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARS-CoV_GZ02 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARS-CoV_HS_SZ_61 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARS-CoV_A031 LADAGFMKQYGQCLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
WIV-16 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
WIV-1 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARSr-CoV_ZXC21 LADAGFIKQYGDCLGDISARDLICAQKFNGLTVLPPLLTDEMIAAYTAALISGTATAGWT
SARSr-CoV_ZC45 LADAGFIKQYGDCLGGISARDLICAQKFNGLTVLPPLLTDEMIAAYTAALISGTATAGWT
SARSr-CoV_RpS3 LADAGFMKQYGECLGDISARDLICAQKFNGLTVLPPLLTDEMIAAYTAALVSGTATAGWT
SARSr-CoV_Rs672 LADAGFMKQYGECLGDISARDLICAQKFNGLTVLPPLLTDEMIAAYTAALVSGTATAGWT
consensus>50 LADAGFmKQYG#CLGdInARDLICAQKFNGLTVLPPLLTD#MIAaYTaALvsGTaTaGWT
900 910 920 930 940 950
SARS-CoV-2 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGK
SARSr-CoV_RaTG13 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGK
SARS-CoV_Urbani FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARS-CoV_CUHK-W1 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARS-CoV_GZ02 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARS-CoV_HS_SZ_61 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARS-CoV_A031 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
WIV-16 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
WIV-1 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARSr-CoV_ZXC21 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQESLTSTASALGK
SARSr-CoV_ZC45 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQESLTSTASALGK
SARSr-CoV_RpS3 FGAGSALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARSr-CoV_Rs672 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
consensus>50 FGAGaALQIPFAMQMAYRFNGIGVTQNVLYENQKqIANQFNkAIsqIQ#SLttTstALGK


960 970 980 990 1000 1010
SARS-CoV-2 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_RaTG13 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_Urbani LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_CUHK-W1 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_GZ02 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_HS_SZ_61 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_A031 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
WIV-16 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
WIV-1 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_ZXC21 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_ZC45 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_RpS3 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_Rs672 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
consensus>50 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY


1020 1030 1040 1050 1060 1070
SARS-CoV-2 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTY
SARSr-CoV_RaTG13 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTY
SARS-CoV_Urbani VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARS-CoV_CUHK-W1 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARS-CoV_GZ02 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARS-CoV_HS_SZ_61 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARS-CoV_A031 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
WIV-16 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
WIV-1 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARSr-CoV_ZXC21 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTY
SARSr-CoV_ZC45 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTY
SARSr-CoV_RpS3 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARSr-CoV_Rs672 VTQQLIRAAEIRASANLAATKMSECVPGQSKRVDFCGRGYHLMSFPQAAPHGVVFLHVTY
consensus>50 VTQQLIRAAEIRASANLAATKMSECVlGQSKRVDFCGkGYHLMSFPQaAPHGVVFLHVTY


1080 1090 1100 1110 1120 1130
SARS-CoV-2 VPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD
SARSr-CoV_RaTG13 VPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGSCD
SARS-CoV_Urbani VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCD
SARS-CoV_CUHK-W1 VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCD
SARS-CoV_GZ02 VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCD
SARS-CoV_HS_SZ_61 VPSQERNFTTAPAICHEGKAYFPREGVFVFSGTSWFITQRNFFSPQIITTDNTFVSGNCD
SARS-CoV_A031 VPSQERNFTTAPAICHEGKAYFPREGVFVFSGTSWFITQRNFFSPQIITTDNTFVSGNCD
WIV-16 VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGSCD
WIV-1 VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGSCD
SARSr-CoV_ZXC21 IPSQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD
SARSr-CoV_ZC45 IPSQEKNFTTAPAICHEGKAHFPREGVFVSNGTHWFVTQRNFYEPKIITTDNTFVSGNCD
SARSr-CoV_RpS3 VPSQERNFTTAPAICHEGKAYFPREGVFVSNGTSWFITQRNFYSPQIITTDNTFVAGSCD
SARSr-CoV_Rs672 VPSQEKNFTTAPAICHEGKAYFPREGVFVSNGTSWFITQRNFYSPQIITTDNTFVAGNCD
consensus>50 !PsQErNFTTAPAICH#GKAyFPREGVFVfnGTsWF!TQRNF%sPqIITTDNTFVsGnCD

1140 1150 1160 1170 1180 1190
SARS-CoV-2 VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_RaTG13 VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARS-CoV_Urbani VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARS-CoV_CUHK-W1 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARS-CoV_GZ02 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQEEIDRLN
SARS-CoV_HS_SZ_61 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQEEIDRLN
SARS-CoV_A031 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQEEIDRLN
WIV-16 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
WIV-1 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_ZXC21 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDIDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_ZC45 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDIDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_RpS3 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_Rs672 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
consensus>50 VVIGI!NNTVYDPLQPELDSFKEELDKYFKNHTSPD!DLGDISGINASVVNIQkEIDRLN


1200 1210 1220 1230 1240 1250
SARS-CoV-2 EVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGC
SARSr-CoV_RaTG13 EVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIIMVTIMLCCMTSCCSCLKGC
SARS-CoV_Urbani EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARS-CoV_CUHK-W1 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARS-CoV_GZ02 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARS-CoV_HS_SZ_61 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARS-CoV_A031 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
WIV-16 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
WIV-1 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARSr-CoV_ZXC21 EVARNLNESLIDLQELGKYEHYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGC
SARSr-CoV_ZC45 EVARNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGC
SARSr-CoV_RpS3 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARSr-CoV_Rs672 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMATILLCCMTSCCSCLKGA
consensus>50 EVAkNLNESLIDLQELGKYEqYIKWPWY!WLGFIAGLIAI!MvTI$LCCMTSCCSCLKGa

1260 1270 1280
SARS-CoV-2 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_RaTG13 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_Urbani CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_CUHK-W1 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_GZ02 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_HS_SZ_61 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_A031 CSCGSCCKFDEDDSEPVLKGVKLHYT
WIV-16 CSCGSCCKFDEDDSEPVLKGVKLHYT
WIV-1 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_ZXC21 CSCGFCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_ZC45 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_RpS3 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_Rs672 CSCGSCCKFDEDDSEPVLKGVKLHYT
consensus>50 CSCGsCCKFDEDDSEPVLKGVKLHYT

Figure S1. Amino acid sequence alignment of sarbecovirus S glycoproteins. The following sequences were used: SARS-CoV-2 (), SARSr-CoV RaTG13 (QHR63300.1), SARS-CoV Urbani (AAP13441.1), SARS-CoV CUHK-W1 (AAP13567.1), SARS-CoV GZ02 (AAS00003.1), SARS-CoV HS/SZ/61/03 (AY515512.1), SARS-CoV A031 (AAV97988.1), WIV-16 (ALK02457.1), WIV-1 (AGZ48828.1), SARSr-CoV ZXC21 (AVP78042.1), SARSr-CoV ZC45 (AVP78031.1), SARSr-CoV Rp3 (Q3I5J5.1), SARSr-CoV Rs672 (ACU31032.1).

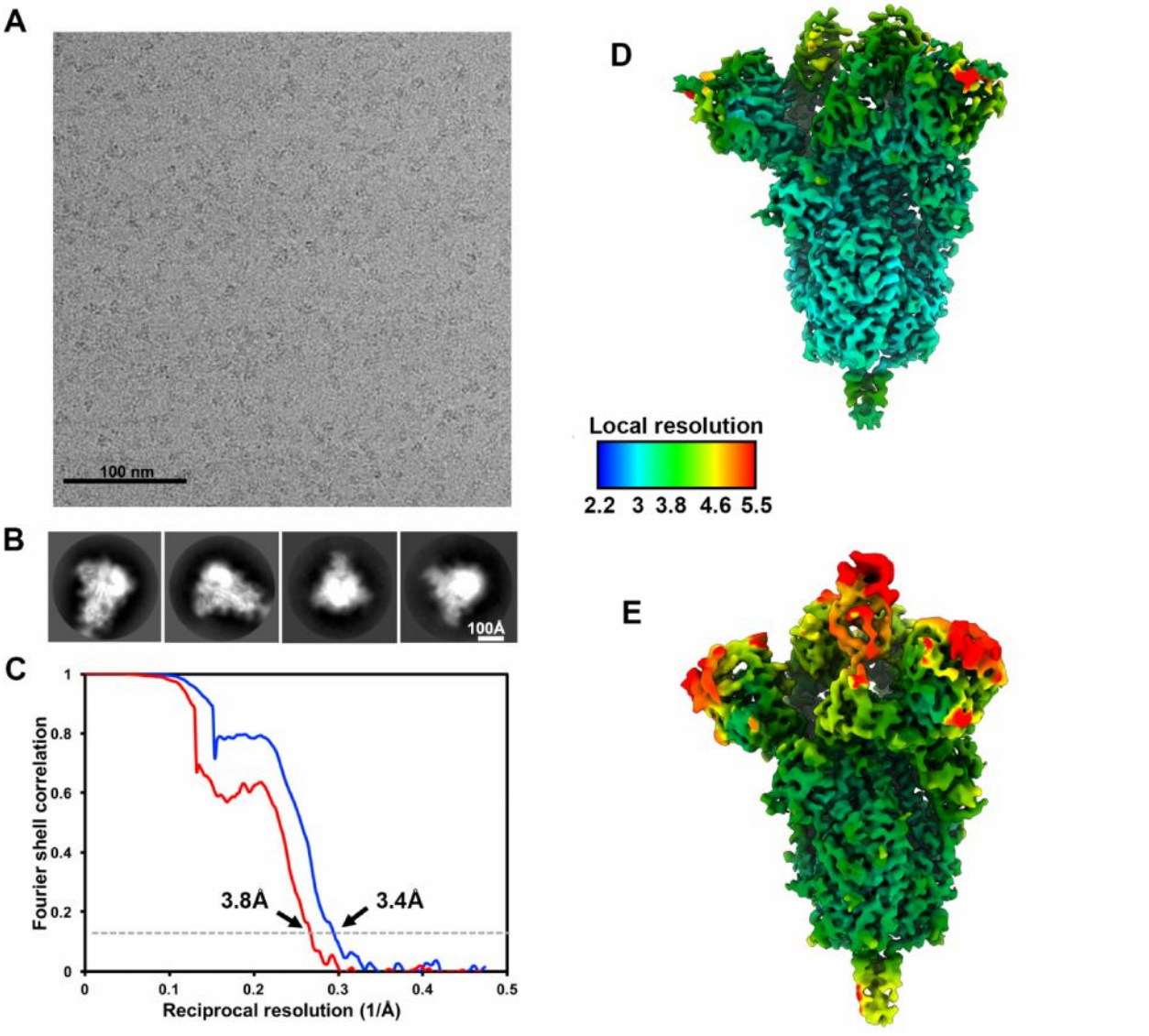


Figure S2. CryoEM data processing and validation. A-B. Representative electron micrograph (A) and class averages (B) of SARS-CoV-2 S embedded in vitreous ice. **C.** Gold-standard Fourier shell correlation curves for the closed (blue) and partially open trimers (red). The 0.143 cutoff is indicated by horizontal dashed lines. **D-E.** Local resolution map calculated using cryoSPARC for the closed (D) and partially open (E) reconstructions.

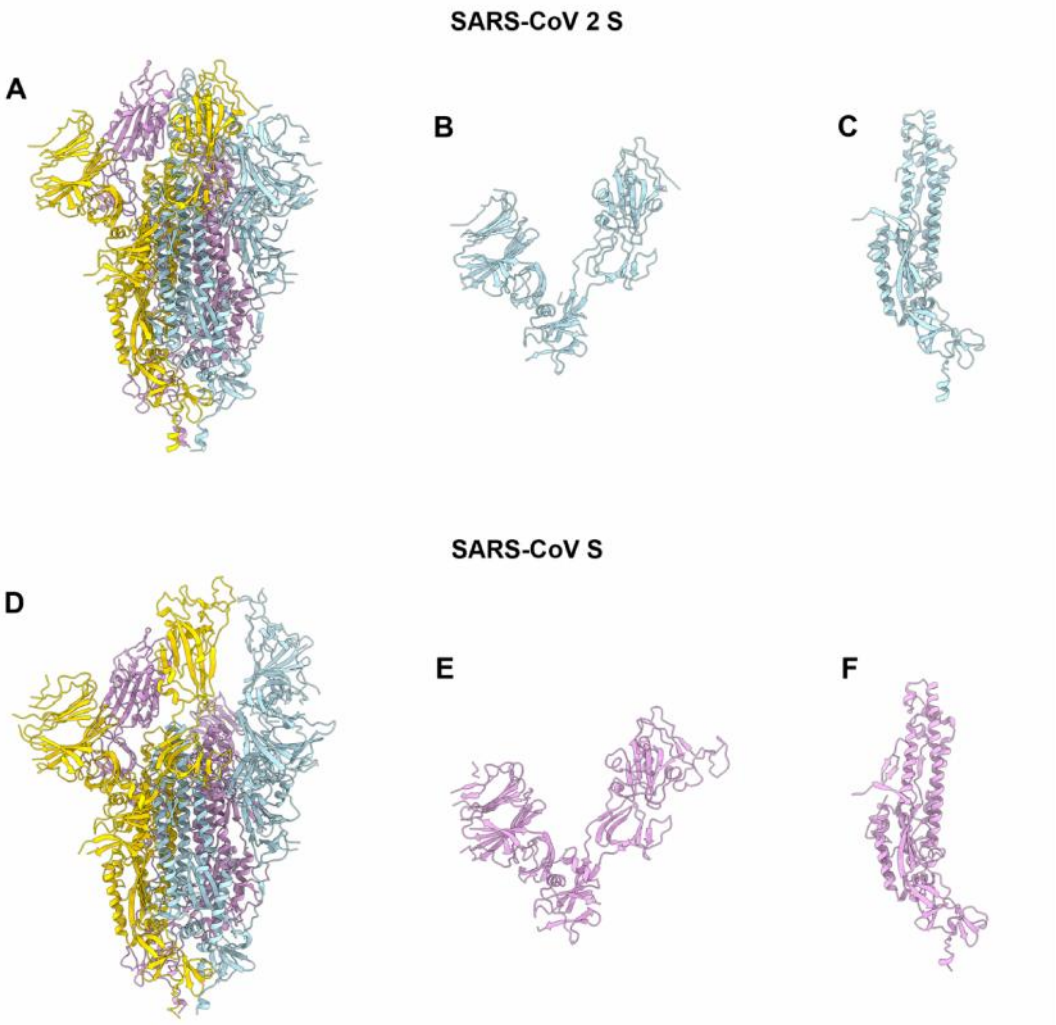
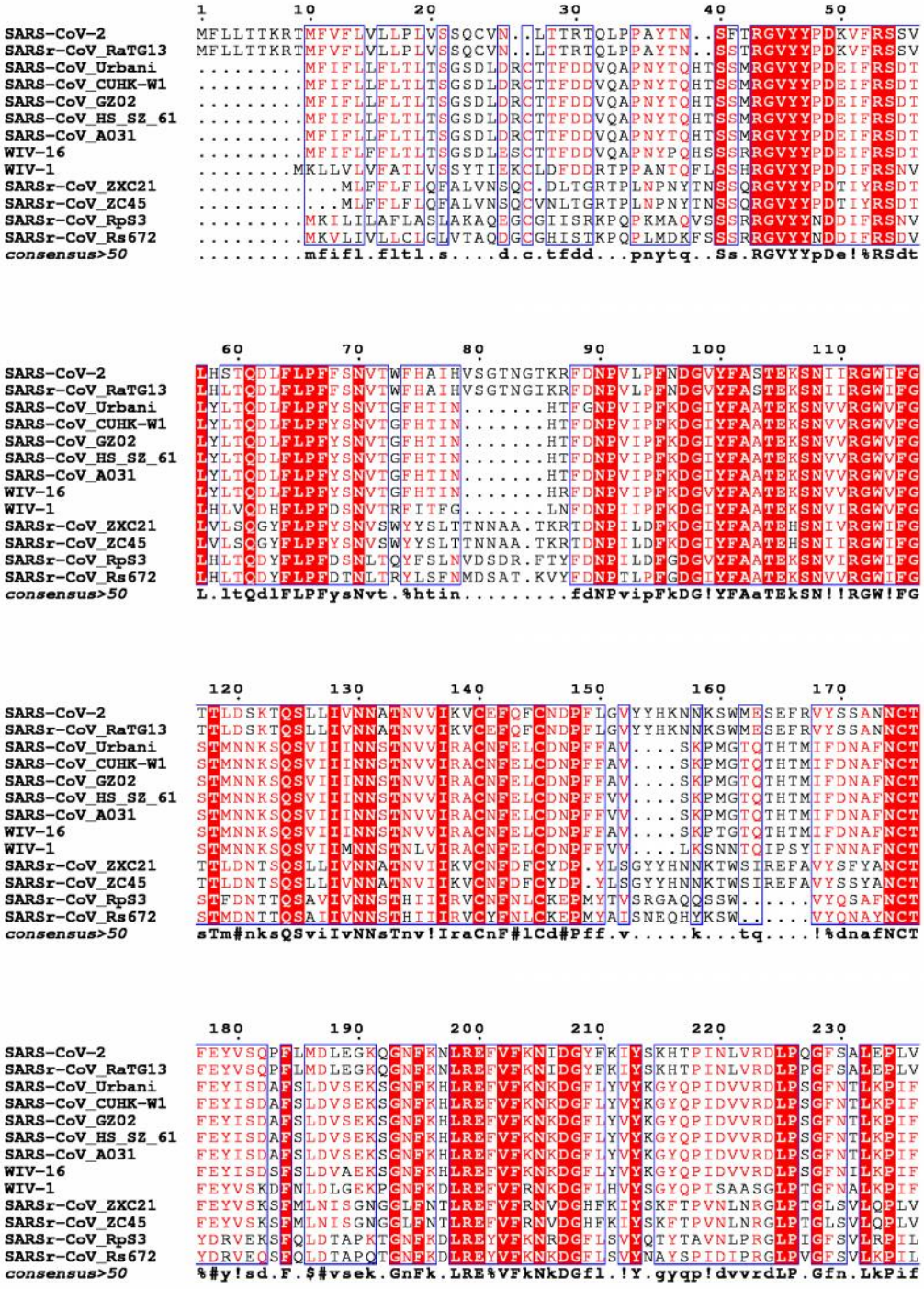
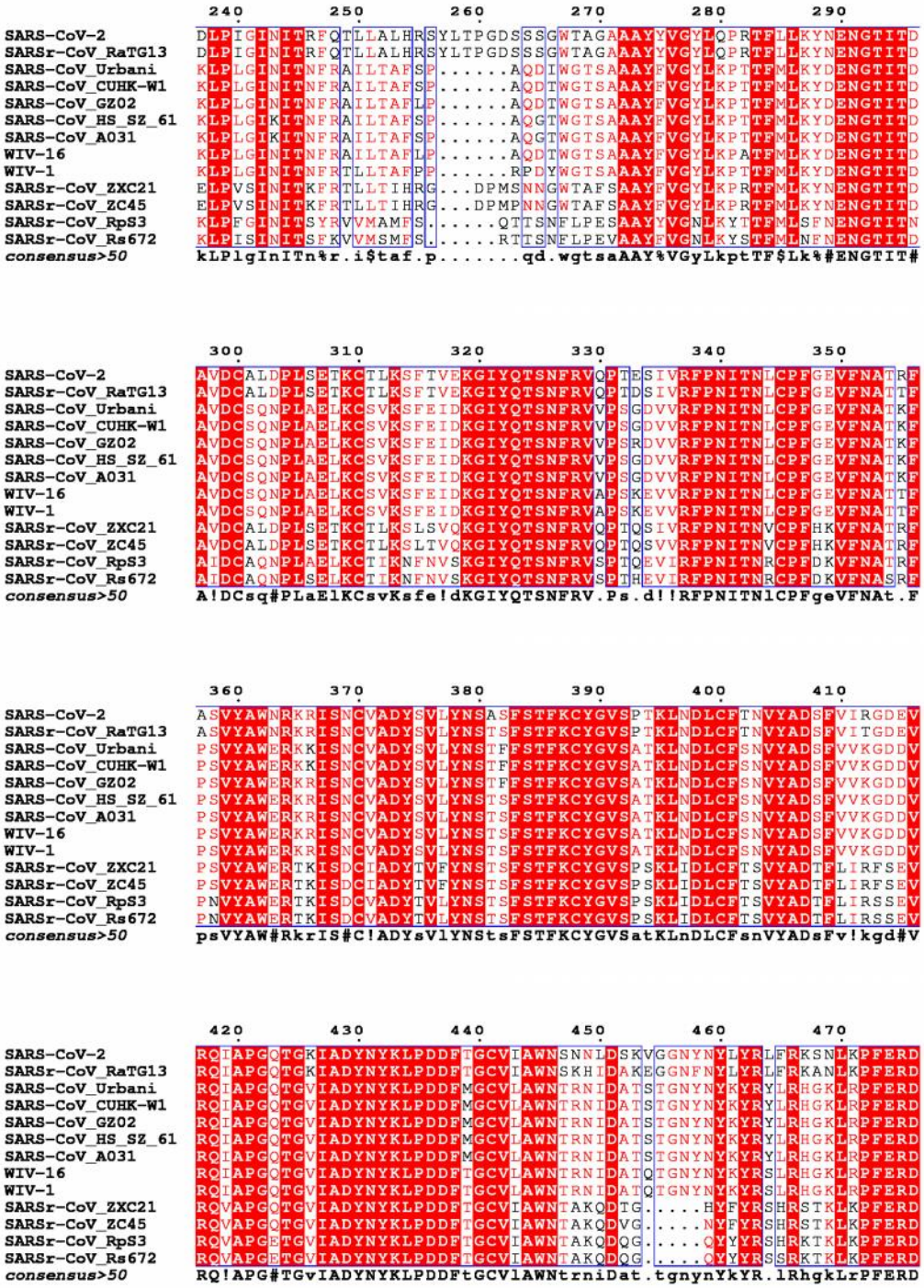
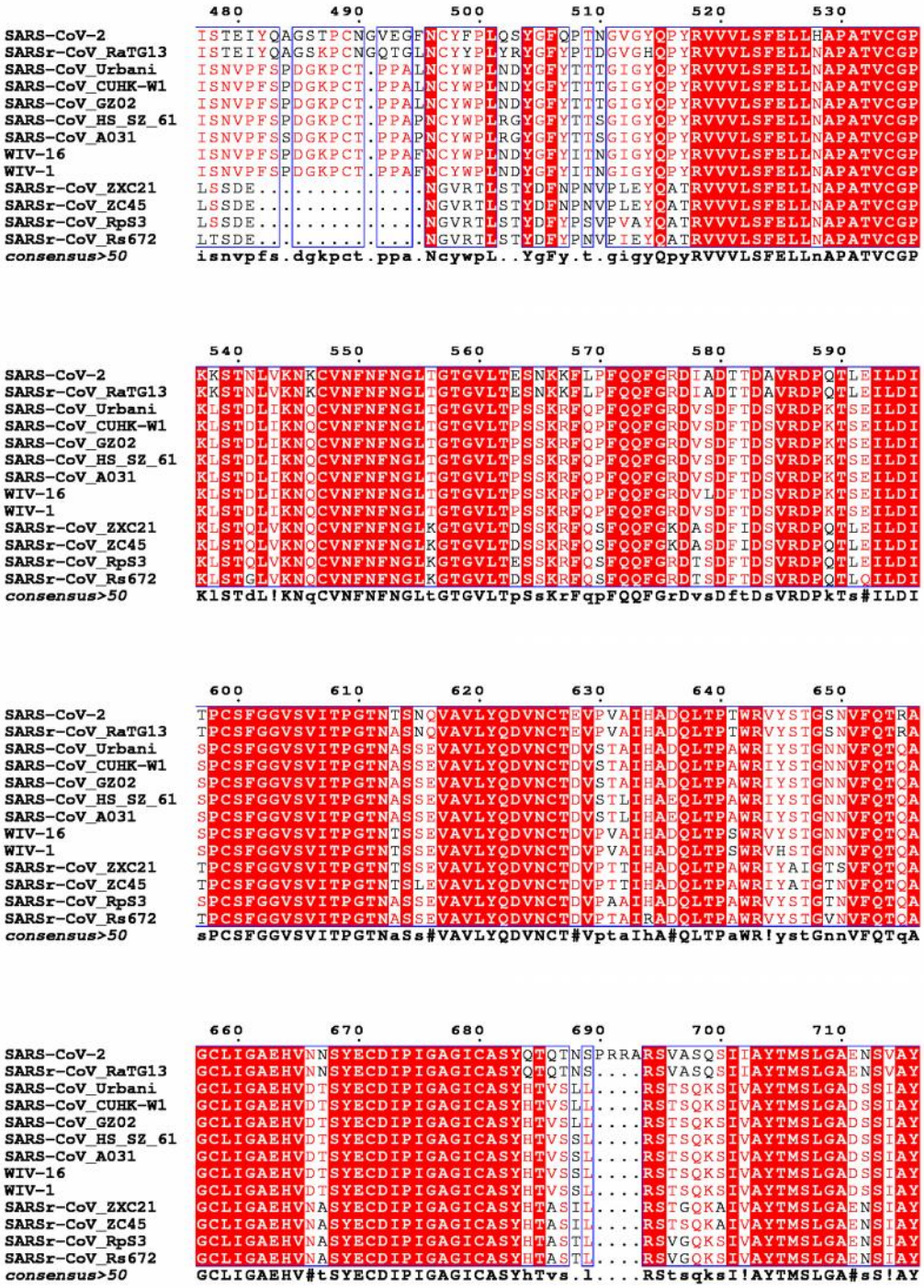
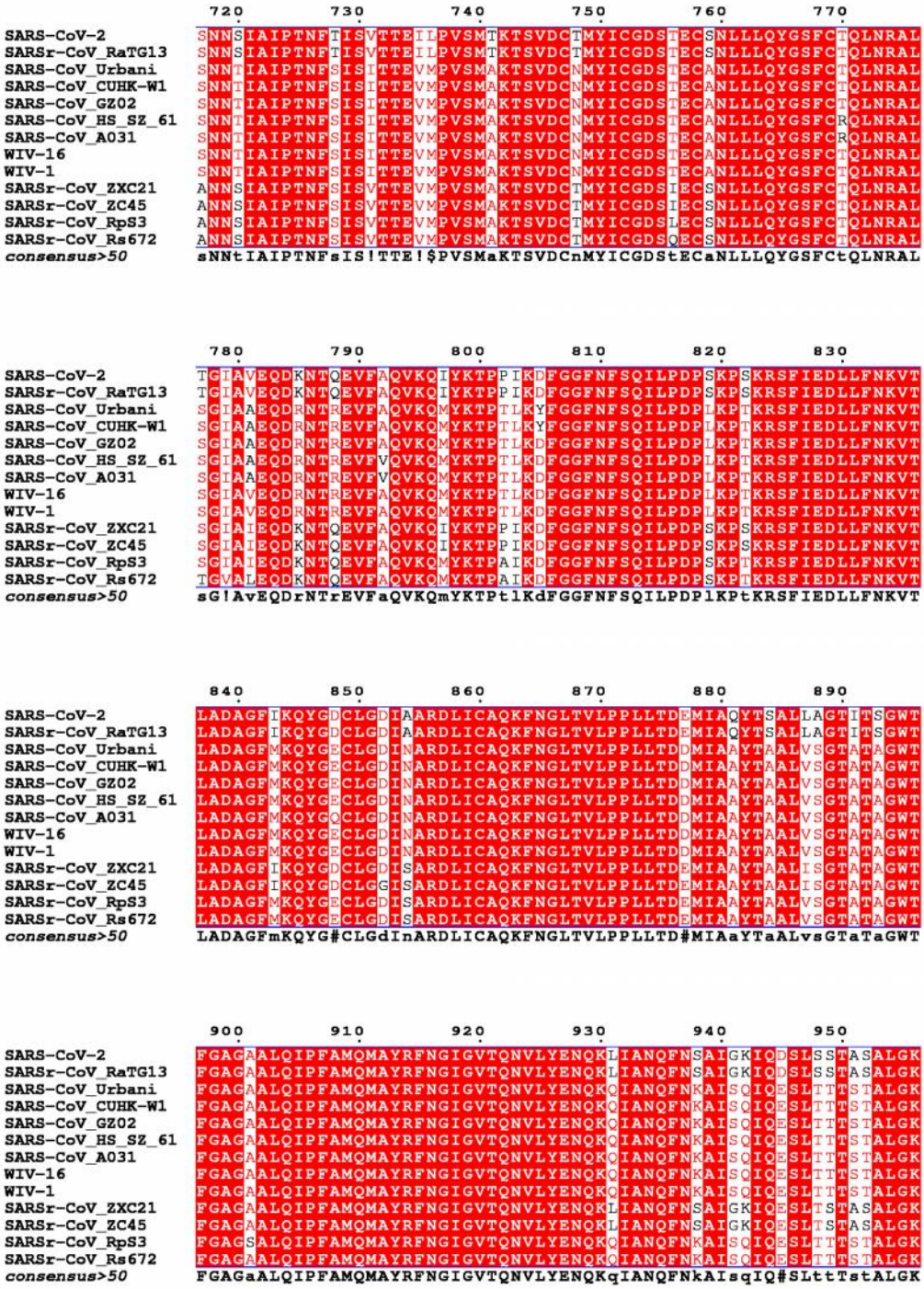


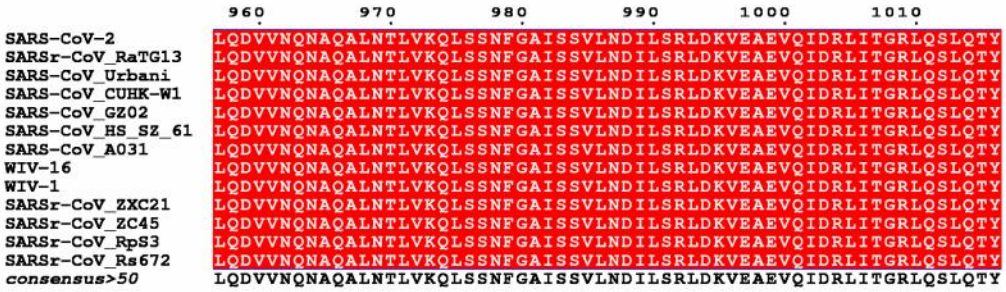
Figure S3. Comparison of the SARS-CoV-2 and SARS-CoV S structures. A-D. Ribbon diagrams of the SARS-CoV-2 S (A) and SARS-CoV S (PDB 6NB6, D) ectodomain cryoEM structures. **B-E,** The SARS-CoV-2 (C) and SARS-CoV (E) S_1 subunits. **C-F,** The SARS-CoV-2 (C) and SARS-CoV (F) S_2 subunits.

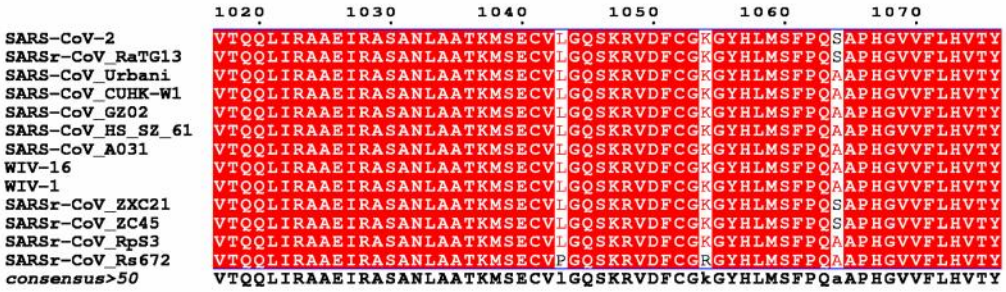

SARS-CoV-2
SARSr-CoV_RaTG13
SARS-CoV_Urbani
SARS-CoV_CUHK-W1
SARS-CoV_GZ02
SARS-CoV_HS_SZ_61
SARS-CoV_A031
WIV-16
WIV-1
SARSr-CoV_ZXC21
SARSr-CoV_ZC45
SARSr-CoV_RpS3
SARSr-CoV_Rs672
consensus>50

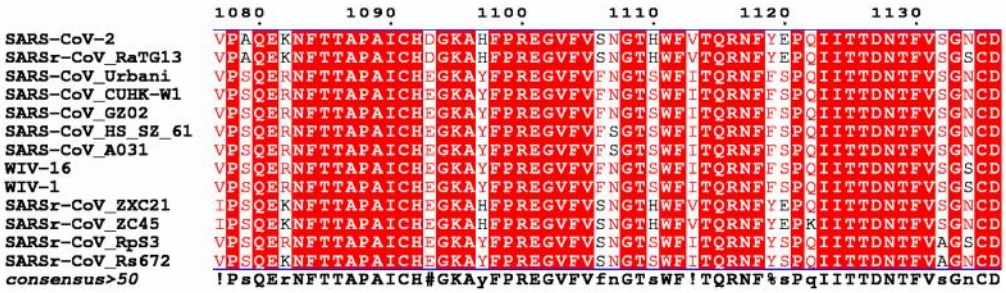

240 250 260 270 280 290
SARS-CoV-2 DLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
SARSr-CoV_RaTG13 DLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
SARS-CoV_Urbani KLPLGINITNFRAILTAFSP......AQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARS-CoV_CUHK-W1 KLPLGINITNFRAILTAFSP......AQDTWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARS-CoV_GZ02 KLPLGINITNFRAILTAFLP......AQDTWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARS-CoV_HS_SZ_61 KLPLGIKITNFRAILTAFSP......AQGTWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARS-CoV_A031 KLPLGIKITNFRAILTAFSP......AQGTWGTSAAAYFVGYLKPTTFMLKYDENGTITD
WIV-16 KLPLGINITNFRAILTAFLP......AQDTWGTSAAAYFVGYLKPATFMLKYDENGTITD
WIV-1 KLPLGINITNFRTLLTAFPP......RPDYWGTSAAAYFVGYLKPTTFMLKYDENGTITD
SARSr-CoV_ZXC21 ELPVSINITKFRTLLTIHRG...DPMSNNGWTAFSAAYFVGYLKPRTFMLKYNENGTITD
SARSr-CoV_ZC45 ELPVSINITKFRTLLTIHRG...DPMPNNGWTAFSAAYFVGYLKPRTFMLKYNENGTITD
SARSr-CoV_RpS3 KLPFGINITSYRVVMAMFS.....QTTSNFLPESAAYYVGNLKYTTFMLSFNENGTITN
SARSr-CoV_Rs672 KLPISINITSFKVVMSMFS.....RTTSNFLPEVAAYFVGNLKYSTFMLNFNENGTITD
consensus>50 kLPlgInITn%r.i$taf.p.......qd.wgtsaAAY%VGyLkptTF$Lk%#ENGTIT#
300 310 320 330 340 350
SARS-CoV-2 AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRF
SARSr-CoV_RaTG13 AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTDSIVRFPNITNLCPFGEVFNATTF
SARS-CoV_Urbani AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
SARS-CoV_CUHK-W1 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
SARS-CoV_GZ02 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSRDVVRFPNITNLCPFGEVFNATKF
SARS-CoV_HS_SZ_61 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
SARS-CoV_A031 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
WIV-16 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVAPSKEVVRFPNITNLCPFGEVFNATTF
WIV-1 AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVAPSKEVVRFPNITNLCPFGEVFNATTF
SARSr-CoV_ZXC21 AVDCALDPLSETKCTLKSLSVQKGIYQTSNFRVQPTQSIVRFPNITNVCPFHKVFNATRF
SARSr-CoV_ZC45 AVDCALDPLSETKCTLKSLTVQKGIYQTSNFRVQPTQSVVRFPNITNVCPFHKVFNATRF
SARSr-CoV_RpS3 AIDCAQNPLAELKCTIKNFNVSKGIYQTSNFRVSPTQEVIRFPNITNRCPFDKVFNATRF
SARSr-CoV_Rs672 AIDCAQNPLSELKCTIKNFNVSKGIYQTSNFRVSPTHEVIRFPNITNRCPFDKVFNASRF
consensus>50 A!DCsq#PLaElKCsvKsfe!dKGIYQTSNFRV.Ps.d!!RFPNITNlCPFgeVFNAt.F
360 370 380 390 400 410
SARS-CoV-2 ASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEV
SARSr-CoV_RaTG13 ASVYAWNRKRISNCVADYSVLYNSTSFSTFKCYGVSPTKLNDLCFTNVYADSFVITGDEV
SARS-CoV_Urbani PSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARS-CoV_CUHK-W1 PSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARS-CoV_GZ02 PSVYAWERKRISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARS-CoV_HS_SZ_61 PSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARS-CoV_A031 PSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
WIV-16 PSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
WIV-1 PSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDV
SARSr-CoV_ZXC21 PSVYAWERTKISDCIADYTVFYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRFSEV
SARSr-CoV_ZC45 PSVYAWERTKISDCIADYTVFYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRFSEV
SARSr-CoV_RpS3 PNVYAWERTKISDCVADYTVLYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRSSEV
SARSr-CoV_Rs672 PNVYAWERTKISDCVADYTVLYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRSSEV
consensus>50 psVYAW#RkrIS#C!ADYsVlYNStsFSTFKCYGVSatKLnDLCFsnVYADsFv!kgd#V
420 430 440 450 460 470
SARS-CoV-2 RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERD
SARSr-CoV_RaTG13 RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSKHIDAKEGGNFNYLYRLFRKANLKPFERD
SARS-CoV_Urbani RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
SARS-CoV_CUHK-W1 RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
SARS-CoV_GZ02 RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
SARS-CoV_HS_SZ_61 RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
SARS-CoV_A031 RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERD
WIV-16 RQIAPGQTGVIADYNYKLPDDFTGCVLAWNTRNIDATQTGNYNYKYRSLRHGKLRPFERD
WIV-1 RQIAPGQTGVIADYNYKLPDDFTGCVLAWNTRNIDATQTGNYNYKYRSLRHGKLRPFERD
SARSr-CoV_ZXC21 RQVAPGQTGVIADYNYKLPDDFTGCVIAWNTAKQDTG.....HYFYRSHRSTKLKPFERD
SARSr-CoV_ZC45 RQVAPGQTGVIADYNYKLPDDFTGCVIAWNTAKQDVG.....NYFYRSHRSTKLKPFERD
SARSr-CoV_RpS3 RQVAPGETGVIADYNYKLPDDFTGCVIAWNTAKQDQG.....QYYYRSHRKTKLKPFERD
SARSr-CoV_Rs672 RQVAPGETGVIADYNYKLPDDFTGCVIAWNTAKQDQG.....QYYYRSSRKTKLKPFERD
consensus>50 RQ!APG#TGvIADYNYKLPDDFtGCVlAWNtrniDat.tgnynYkYR.lRhgkLrPFERD


480 490 500 510 520 530
SARS-CoV-2 ISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGP
SARSr-CoV_RaTG13 ISTEIYQAGSKPCNGQTGLNCYYPLYRYGFYPTDGVGHQPYRVVVLSFELLNAPATVCGP
SARS-CoV_Urbani ISNVPFSPDGKPCT.PPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGP
SARS-CoV_CUHK-W1 ISNVPFSPDGKPCT.PPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGP
SARS-CoV_GZ02 ISNVPFSPDGKPCT.PPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGP
SARS-CoV_HS_SZ_61 ISNVPFSPDGKPCT.PPAPNCYWPLRGYGFYTTSGIGYQPYRVVVLSFELLNAPATVCGP
SARS-CoV_A031 ISNVPFSSDGKPCT.PPAPNCYWPLRGYGFYTTSGIGYQPYRVVVLSFELLNAPATVCGP
WIV-16 ISNVPFSPDGKPCT.PPAFNCYWPLNDYGFYITNGIGYQPYRVVVLSFELLNAPATVCGP
WIV-1 ISNVPFSPDGKPCT.PPAFNCYWPLNDYGFYITNGIGYQPYRVVVLSFELLNAPATVCGP
SARSr-CoV_ZXC21 LSSDE..........NGVRTLSTYDFNPNVPLEYQATRVVVLSFELLNAPATVCGP
SARSr-CoV_ZC45 LSSDE..........NGVRTLSTYDFNPNVPLEYQATRVVVLSFELLNAPATVCGP
SARSr-CoV_RpS3 LSSDE..........NGVRTLSTYDFYPSVPVAYQATRVVVLSFELLNAPATVCGP
SARSr-CoV_Rs672 LTSDE..........NGVRTLSTYDFYPNVPIEYQATRVVVLSFELLNAPATVCGP
consensus>50 isnvpfs.dgkpct.ppa.NcywpL..YgFy.t.gigyQpyRVVVLSFELLnAPATVCGP
540 550 560 570 580 590
SARS-CoV-2 KKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDI
SARSr-CoV_RaTG13 KKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDI
SARS-CoV_Urbani KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARS-CoV_CUHK-W1 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARS-CoV_GZ02 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARS-CoV_HS_SZ_61 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARS-CoV_A031 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
WIV-16 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVLDFTDSVRDPKTSEILDI
WIV-1 KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDI
SARSr-CoV_ZXC21 KLSTQLVKNQCVNFNFNGLKGTGVLTDSSKRFQSFQQFGKDASDFIDSVRDPQTLEILDI
SARSr-CoV_ZC45 KLSTQLVKNQCVNFNFNGLKGTGVLTDSSKRFQSFQQFGKDASDFIDSVRDPQTLEILDI
SARSr-CoV_RpS3 KLSTQLVKNQCVNFNFNGLKGTGVLTESSKRFQSFQQFGRDTSDFTDSVRDPQTLEILDI
SARSr-CoV_Rs672 KLSTGLVKNQCVNFNFNGLKGTGVLTDSSKRFQSFQQFGRDTSDFTDSVRDPQTLQILDI
consensus>50 KlSTdL!KNqCVNFNFNGLtGTGVLTpSsKrFqpFQQFGrDvsDftDsVRDPkTs#ILDI
600 610 620 630 640 650
SARS-CoV-2 TPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRA
SARSr-CoV_RaTG13 TPCSFGGVSVITPGTNASNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRA
SARS-CoV_Urbani SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQA
SARS-CoV_CUHK-W1 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQA
SARS-CoV_GZ02 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQA
SARS-CoV_HS_SZ_61 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTLIHAEQLTPAWRIYSTGNNVFQTQA
SARS-CoV_A031 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTLIHAEQLTPAWRIYSTGNNVFQTQA
WIV-16 SPCSFGGVSVITPGTNTSSEVAVLYQDVNCTDVPVAIHADQLTPSWRVYSTGNNVFQTQA
WIV-1 SPCSFGGVSVITPGTNTSSEVAVLYQDVNCTDVPVAIHADQLTPSWRVHSTGNNVFQTQA
SARSr-CoV_ZXC21 TPCSFGGVSVITPGTNTSSEVAVLYQDVNCTDVPTTIHADQLTPAWRIYAIGTSVFQTQA
SARSr-CoV_ZC45 TPCSFGGVSVITPGTNTSLEVAVLYQDVNCTDVPTTIHADQLTPAWRIYATGTNVFQTQA
SARSr-CoV_RpS3 SPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVPAAIHADQLTPAWRVYSTGTNVFQTQA
SARSr-CoV_Rs672 TPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVPTAIRADQLTPAWRVYSTGVNVFQTQA
consensus>50 sPCSFGGVSVITPGTNaSs#VAVLYQDVNCT#VptaIhA#QLTPaWR!ystGnnVFQTqA
660 670 680 690 700 710
SARS-CoV-2 GCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAY
SARSr-CoV_RaTG13 GCLIGAEHVNNSYECDIPIGAGICASYQTQTNS....RSVASQSIIAYTMSLGAENSVAY
SARS-CoV_Urbani GCLIGAEHVDTSYECDIPIGAGICASYHTVSLL....RSTSQKSIVAYTMSLGADSSIAY
SARS-CoV_CUHK-W1 GCLIGAEHVDTSYECDIPIGAGICASYHTVSLL....RSTSQKSIVAYTMSLGADSSIAY
SARS-CoV_GZ02 GCLIGAEHVDTSYECDIPIGAGICASYHTVSLL....RSTSQKSIVAYTMSLGADSSIAY
SARS-CoV_HS_SZ_61 GCLIGAEHVDTSYECDIPIGAGICASYHTVSSL....RSTSQKSIVAYTMSLGADSSIAY
SARS-CoV_A031 GCLIGAEHVDTSYECDIPIGAGICASYHTVSSL....RSTSQKSIVAYTMSLGADSSIAY
WIV-16 GCLIGAEHVDTSYECDIPIGAGICASYHTVSSL....RSTSQKSIVAYTMSLGADSSIAY
WIV-1 GCLIGAEHVDTSYECDIPIGAGICASYHTVSSL....RSTSQKSIVAYTMSLGADSSIAY
SARSr-CoV_ZXC21 GCLIGAEHVNASYECDIPIGAGICASYHTASIL....RSTGQKAIVAYTMSLGAENSIAY
SARSr-CoV_ZC45 GCLIGAEHVNASYECDIPIGAGICASYHTASIL....RSTSQKAIVAYTMSLGAENSIAY
SARSr-CoV_RpS3 GCLIGAEHVNASYECDIPIGAGICASYHTASTL....RSVGQKSIVAYTMSLGAENSIAY
SARSr-CoV_Rs672 GCLIGAEHVNASYECDIPIGAGICASYHTASTL....RSVGQKSIVAYTMSLGAENSIAY
consensus>50 GCLIGAEHV#tSYECDIPIGAGICASYhTvs.l....RStsqksI!AYTMSLGA#sS!AY

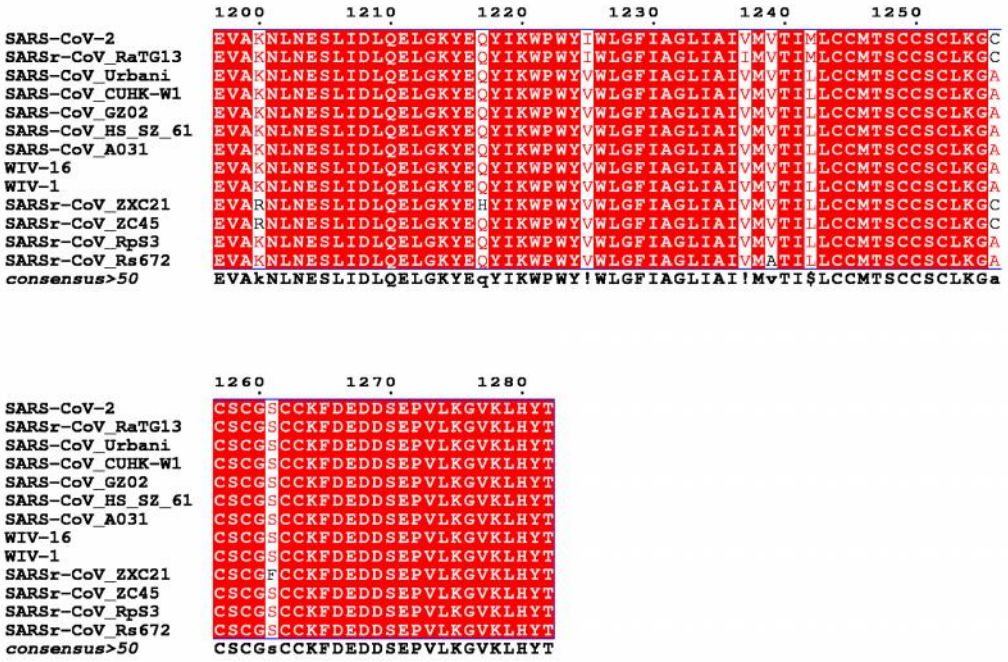

720 730 740 750 760 770
SARS-CoV-2 SNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRAL
SARSr-CoV_RaTG13 SNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRAL
SARS-CoV_Urbani SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SARS-CoV_CUHK-W1 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SARS-CoV_GZ02 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SARS-CoV_HS_SZ_61 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCRQLNRAL
SARS-CoV_A031 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCRQLNRAL
WIV-16 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
WIV-1 SNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SARSr-CoV_ZXC21 ANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSIECSNLLLQYGSFCTQLNRAL
SARSr-CoV_ZC45 ANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSIECSNLLLQYGSFCTQLNRAL
SARSr-CoV_RpS3 ANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSLECSNLLLQYGSFCTQLNRAL
SARSr-CoV_Rs672 ANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSQECSNLLLQYGSFCTQLNRAL
consensus>50 sNNtIAIPTNFsIS!TTE!$PVSMaKTSVDCnMYICGDStECaNLLLQYGSFCtQLNRAL
780 790 800 810 820 830
SARS-CoV-2 TGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
SARSr-CoV_RaTG13 TGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
SARS-CoV_Urbani SGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARS-CoV_CUHK-W1 SGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARS-CoV_GZ02 SGIAAEQDRNTREVFAQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARS-CoV_HS_SZ_61 SGIAAEQDRNTREVFVQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARS-CoV_A031 SGIAAEQDRNTREVFVQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
WIV-16 SGIAVEQDRNTREVFAQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
WIV-1 SGIAVEQDRNTREVFAQVKQMYKTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT
SARSr-CoV_ZXC21 SGIAIEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
SARSr-CoV_ZC45 SGIAIEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
SARSr-CoV_RpS3 SGIAIEQDKNTQEVFAQVKQMYKTPAIKDFGGFNFSQILPDPSKPTKRSFIEDLLFNKVT
SARSr-CoV_Rs672 TGVALEQDKNTQEVFAQVKQMYKTPAIKDFGGFNFSQILPDPSKPTKRSFIEDLLFNKVT
consensus>50 sG!AvEQDrNTrEVFaQVKQmYKTPtlKdFGGFNFSQILPDPlKPtKRSFIEDLLFNKVT
840 850 860 870 880 890
SARS-CoV-2 LADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWT
SARSr-CoV_RaTG13 LADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWT
SARS-CoV_Urbani LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARS-CoV_CUHK-W1 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARS-CoV_GZ02 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARS-CoV_HS_SZ_61 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARS-CoV_A031 LADAGFMKQYGQCLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
WIV-16 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
WIV-1 LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWT
SARSr-CoV_ZXC21 LADAGFIKQYGDCLGDISARDLICAQKFNGLTVLPPLLTDEMIAAYTAALISGTATAGWT
SARSr-CoV_ZC45 LADAGFIKQYGDCLGGISARDLICAQKFNGLTVLPPLLTDEMIAAYTAALISGTATAGWT
SARSr-CoV_RpS3 LADAGFMKQYGECLGDISARDLICAQKFNGLTVLPPLLTDEMIAAYTAALVSGTATAGWT
SARSr-CoV_Rs672 LADAGFMKQYGECLGDISARDLICAQKFNGLTVLPPLLTDEMIAAYTAALVSGTATAGWT
consensus>50 LADAGFmKQYG#CLGdInARDLICAQKFNGLTVLPPLLTD#MIAaYTaALvsGTaTaGWT
900 910 920 930 940 950
SARS-CoV-2 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGK
SARSr-CoV_RaTG13 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGK
SARS-CoV_Urbani FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARS-CoV_CUHK-W1 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARS-CoV_GZ02 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARS-CoV_HS_SZ_61 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARS-CoV_A031 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
WIV-16 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
WIV-1 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARSr-CoV_ZXC21 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQESLTSTASALGK
SARSr-CoV_ZC45 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQESLTSTASALGK
SARSr-CoV_RpS3 FGAGSALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
SARSr-CoV_Rs672 FGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGK
consensus>50 FGAGaALQIPFAMQMAYRFNGIGVTQNVLYENQKqIANQFNkAIsqIQ#SLttTstALGK


960 970 980 990 1000 1010
SARS-CoV-2 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_RaTG13 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_Urbani LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_CUHK-W1 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_GZ02 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_HS_SZ_61 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARS-CoV_A031 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
WIV-16 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
WIV-1 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_ZXC21 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_ZC45 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_RpS3 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
SARSr-CoV_Rs672 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
consensus>50 LQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY


1020 1030 1040 1050 1060 1070
SARS-CoV-2 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTY
SARSr-CoV_RaTG13 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTY
SARS-CoV_Urbani VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARS-CoV_CUHK-W1 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARS-CoV_GZ02 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARS-CoV_HS_SZ_61 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARS-CoV_A031 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
WIV-16 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
WIV-1 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARSr-CoV_ZXC21 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTY
SARSr-CoV_ZC45 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTY
SARSr-CoV_RpS3 VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTY
SARSr-CoV_Rs672 VTQQLIRAAEIRASANLAATKMSECVPGQSKRVDFCGRGYHLMSFPQAAPHGVVFLHVTY
consensus>50 VTQQLIRAAEIRASANLAATKMSECVlGQSKRVDFCGkGYHLMSFPQaAPHGVVFLHVTY


1080 1090 1100 1110 1120 1130
SARS-CoV-2 VPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD
SARSr-CoV_RaTG13 VPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGSCD
SARS-CoV_Urbani VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCD
SARS-CoV_CUHK-W1 VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCD
SARS-CoV_GZ02 VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCD
SARS-CoV_HS_SZ_61 VPSQERNFTTAPAICHEGKAYFPREGVFVFSGTSWFITQRNFFSPQIITTDNTFVSGNCD
SARS-CoV_A031 VPSQERNFTTAPAICHEGKAYFPREGVFVFSGTSWFITQRNFFSPQIITTDNTFVSGNCD
WIV-16 VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGSCD
WIV-1 VPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGSCD
SARSr-CoV_ZXC21 IPSQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD
SARSr-CoV_ZC45 IPSQEKNFTTAPAICHEGKAHFPREGVFVSNGTHWFVTQRNFYEPKIITTDNTFVSGNCD
SARSr-CoV_RpS3 VPSQERNFTTAPAICHEGKAYFPREGVFVSNGTSWFITQRNFYSPQIITTDNTFVAGSCD
SARSr-CoV_Rs672 VPSQEKNFTTAPAICHEGKAYFPREGVFVSNGTSWFITQRNFYSPQIITTDNTFVAGNCD
consensus>50 !PsQErNFTTAPAICH#GKAyFPREGVFVfnGTsWF!TQRNF%sPqIITTDNTFVsGnCD

1140 1150 1160 1170 1180 1190
SARS-CoV-2 VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_RaTG13 VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARS-CoV_Urbani VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARS-CoV_CUHK-W1 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARS-CoV_GZ02 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQEEIDRLN
SARS-CoV_HS_SZ_61 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQEEIDRLN
SARS-CoV_A031 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQEEIDRLN
WIV-16 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
WIV-1 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_ZXC21 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDIDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_ZC45 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDIDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_RpS3 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
SARSr-CoV_Rs672 VVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLN
consensus>50 VVIGI!NNTVYDPLQPELDSFKEELDKYFKNHTSPD!DLGDISGINASVVNIQkEIDRLN


1200 1210 1220 1230 1240 1250
SARS-CoV-2 EVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGC
SARSr-CoV_RaTG13 EVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIIMVTIMLCCMTSCCSCLKGC
SARS-CoV_Urbani EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARS-CoV_CUHK-W1 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARS-CoV_GZ02 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARS-CoV_HS_SZ_61 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARS-CoV_A031 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
WIV-16 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
WIV-1 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARSr-CoV_ZXC21 EVARNLNESLIDLQELGKYEHYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGC
SARSr-CoV_ZC45 EVARNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGC
SARSr-CoV_RpS3 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGA
SARSr-CoV_Rs672 EVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMATILLCCMTSCCSCLKGA
consensus>50 EVAkNLNESLIDLQELGKYEqYIKWPWY!WLGFIAGLIAI!MvTI$LCCMTSCCSCLKGa

1260 1270 1280
SARS-CoV-2 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_RaTG13 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_Urbani CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_CUHK-W1 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_GZ02 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_HS_SZ_61 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARS-CoV_A031 CSCGSCCKFDEDDSEPVLKGVKLHYT
WIV-16 CSCGSCCKFDEDDSEPVLKGVKLHYT
WIV-1 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_ZXC21 CSCGFCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_ZC45 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_RpS3 CSCGSCCKFDEDDSEPVLKGVKLHYT
SARSr-CoV_Rs672 CSCGSCCKFDEDDSEPVLKGVKLHYT
consensus>50 CSCGsCCKFDEDDSEPVLKGVKLHYT

Figure S1. Amino acid sequence alignment of sarbecovirus S glycoproteins. The following sequences were used: SARS-CoV-2 (), SARSr-CoV RaTG13 (QHR63300.1), SARS-CoV Urbani (AAP13441.1), SARS-CoV CUHK-W1 (AAP13567.1), SARS-CoV GZ02 (AAS00003.1), SARS-CoV HS/SZ/61/03 (AY515512.1), SARS-CoV A031 (AAV97988.1), WIV-16 (ALK02457.1), WIV-1 (AGZ48828.1), SARSr-CoV ZXC21 (AVP78042.1), SARSr-CoV ZC45 (AVP78031.1), SARSr-CoV Rp3 (Q3I5J5.1), SARSr-CoV Rs672 (ACU31032.1).

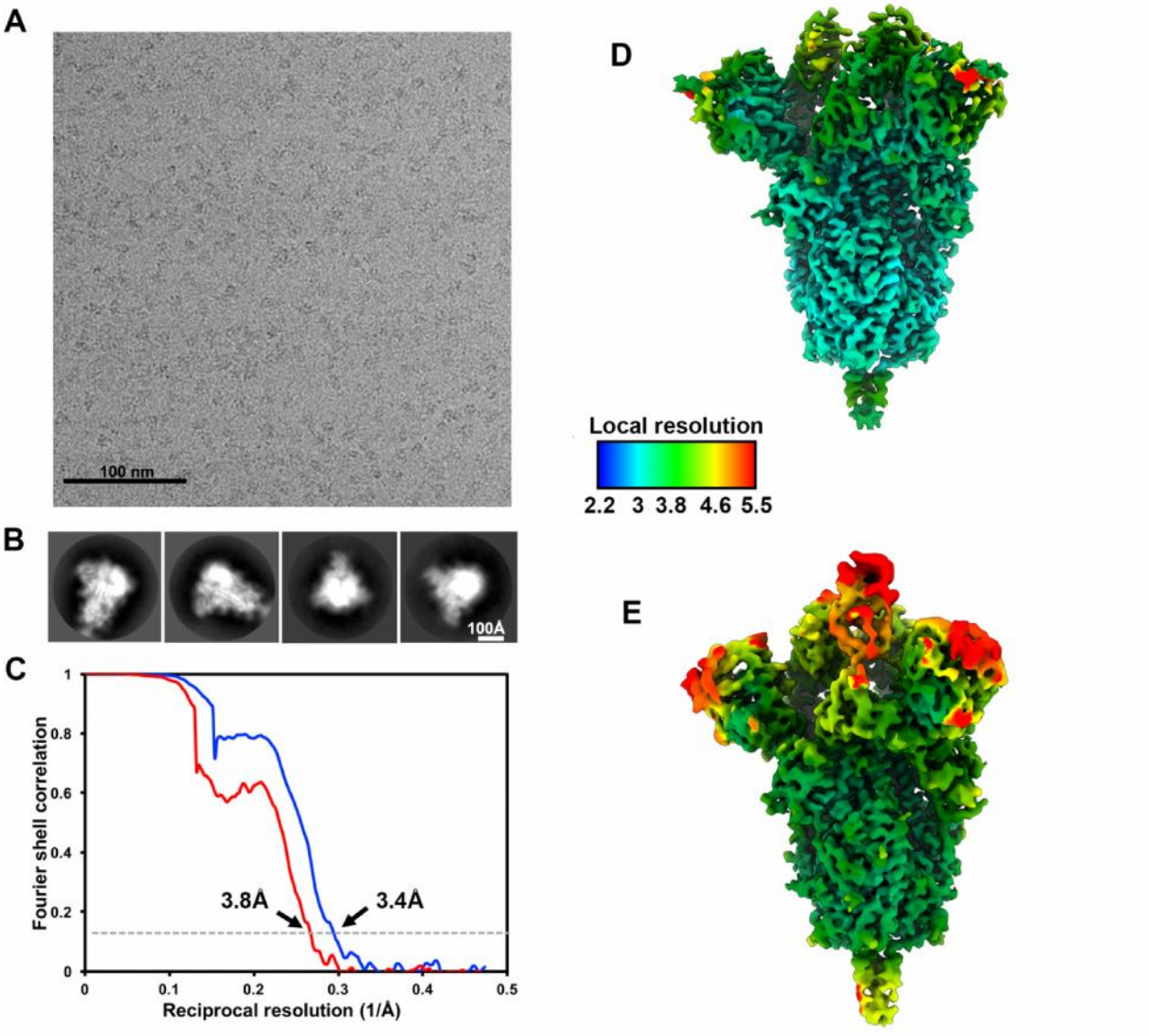


Figure S2. CryoEM data processing and validation. A-B. Representative electron micrograph (A) and class averages (B) of SARS-CoV-2 S embedded in vitreous ice. **C.** Gold-standard Fourier shell correlation curves for the closed (blue) and partially open trimers (red). The 0.143 cutoff is indicated by horizontal dashed lines. **D-E.** Local resolution map calculated using cryoSPARC for the closed (D) and partially open (E) reconstructions.

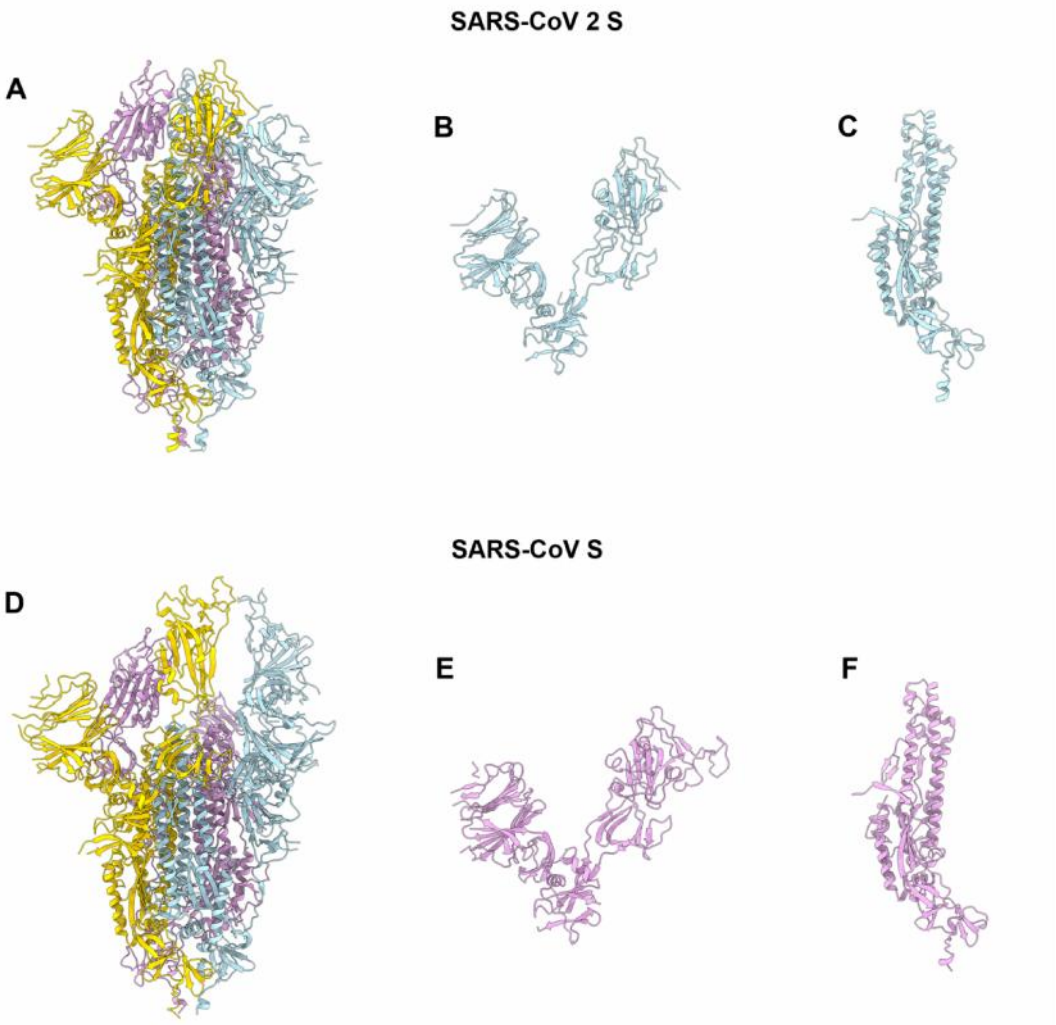


Figure S3. Comparison of the SARS-CoV-2 and SARS-CoV S structures. A-D. Ribbon diagrams of the SARS-CoV-2 S (A) and SARS-CoV S (PDB 6NB6, D) ectodomain cryoEM structures. **B-E,** The SARS-CoV-2 (C) and SARS-CoV (E) S_1 subunits. **C-F,** The SARS-CoV-2 (C) and SARS-CoV (F) S_2 subunits.