

Accelerated Article Preview

Cryo-EM structure of SARS-CoV-2 postfusion spike in membrane

Received: 3 December 2022

Accepted: 31 May 2023

Accelerated Article Preview

Cite this article as: Shi, W. et al. Cryo-EM structure of SARS-CoV-2 postfusion spike in membrane. *Nature* <https://doi.org/10.1038/s41586-023-06273-4> (2023)

Wei Shi, Yongfei Cai, Haisun Zhu, Hanqin Peng, Jewel Voyer, Sophia Rits-Volloch, Hong Cao, Megan L. Mayer, Kangkang Song, Chen Xu, Jianming Lu, Jun Zhang & Bing Chen

This is a PDF file of a peer-reviewed paper that has been accepted for publication. Although unedited, the content has been subjected to preliminary formatting. Nature is providing this early version of the typeset paper as a service to our authors and readers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

Cryo-EM structure of SARS-CoV-2 postfusion spike in membrane

Wei Shi^{1,2#}, Yongfei Cai^{1,2&}, Haisun Zhu³, Hanqin Peng¹, Jewel Voyer¹, Sophia Rits-Volloch¹, Hong Cao⁴, Megan L. Mayer⁵, Kangkang Song⁶, Chen Xu⁶, Jianming Lu^{4,7}, Jun Zhang^{1,2#*} and Bing Chen^{1,2*}

¹Division of Molecular Medicine, Boston Children's Hospital, and ²Department of Pediatrics, Harvard Medical School, 3 Blackfan Street, Boston, MA 02115, USA.

³Institute for Protein Innovation, Harvard Institutes of Medicine, 4 Blackfan Circle, Boston, MA 02115, USA.

⁴Codex BioSolutions, Inc., 12358 Parklawn Drive, Rockville, MD 20852, USA.

⁵The Harvard Cryo-EM Center for Structural Biology, 250 Longwood Avenue, Boston, MA 02115, USA.

⁶Department of Biochemistry & Molecular Biotechnology, and CryoEM Core Facility, University of Massachusetts Chan Medical School, 55 Lake Avenue, Worcester, MA 01655, USA

⁷Department of Biochemistry and Molecular and Cellular Biology, Georgetown University, 3900 Reservoir Road, N.W., Washington, D.C. 20057, USA.

#These authors contributed equally to this work.

&Present address: CSL Seqirus, 225 Wyman Street, Waltham, MA 02451, USA.

*Correspondence to: Bing Chen, phone: 617-355-4625, FAX: 617-730-1967, Email: bchen@crystal.harvard.edu; Jun Zhang, phone: 857-218-5193, FAX: 617-730-1967, Email: juzhang@crystal.harvard.edu.

Abstract

Entry of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) into host cells depends on refolding of the virus-encoded spike protein from a prefusion conformation, metastable after cleavage, to a lower energy, stable postfusion conformation^{1,2}. This transition overcomes kinetic barriers for fusion of viral and target cell membranes^{3,4}. We report here a cryo-EM structure of the intact postfusion spike in a lipid bilayer that represents single-membrane product of the fusion reaction. The structure provides structural definition of the functionally critical membrane-interacting segments, including the fusion peptide and transmembrane anchor. The internal fusion peptide forms a hairpin-like wedge that spans almost the entire lipid bilayer and the transmembrane segment wraps around the fusion peptide at the last stage of membrane fusion. These results advance our understanding of the spike protein in a membrane environment and may guide development of intervention strategies.

Introduction

The COVID-19 (coronavirus disease 2019) pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has a death toll in millions and devastating socioeconomic impacts worldwide. SARS-CoV-2 is an enveloped positive-stranded RNA virus that enters a host cell after fusion of the viral and cell membranes. Although membrane fusion is energetically favorable, there are high kinetic barriers when two membranes approach each other, primarily because of repulsive hydration forces^{3,4}. Free energy, required for viral membrane fusion to overcome the kinetic barriers, comes from refolding of the virus-encoded fusion protein from a prefusion conformation, which is metastable after proteolytic cleavage to a lower energy, stable postfusion state^{1,2}. The fusion protein of SARS-CoV-2 is its spike (S) protein, which is a type I, heavily glycosylated membrane protein with a transmembrane (TM) segment embedded in the viral membrane, and another membrane-interacting region, known as “fusion peptide (FP)”, which can insert into the target cell membrane⁵. Like other class I viral fusion proteins including HIV-1 envelope glycoprotein, influenza hemagglutinin and Ebola glycoprotein^{1,2}, the S protein is synthesized as a single-chain precursor, trimerized and subsequently cleaved by a furin-like protease from the infected host cell into two fragments: the receptor-binding fragment S1 and the fusion fragment S2⁶. To initiate the next round infection, S protein binds to the receptor angiotensin converting enzyme 2 (ACE2) on the surface of a new host cell and is further cleaved at a second site in S2 (S2' site) by a host serine protease TMPRSS2 or an endosomal cysteine protease cathepsin L^{5,7,8}. It then undergoes large conformational changes to insert the FP into the target cell membrane and then refold into a hairpin-like postfusion structure, placing the TM and FP at the same end of the molecule, thereby dragging two membranes close together to complete fusion^{9,10}.

Soon after the release of the first SARS-CoV-2 genomic sequence¹¹, the structures of the S protein fragments, such as its ectodomain stabilized in the prefusion conformation^{12,13}, the receptor binding domain (RBD) in complex with ACE2¹⁴⁻¹⁷ and fragments of S2 in the postfusion conformation¹⁸, were reported. In the prefusion ectodomain structure, S1 folds into four different domains - N-terminal domain (NTD), RBD, C-terminal domain 1 (CTD-1) and C-terminal domain 2 (CTD-2), and they wrap around the prefusion conformation of S2. The structures of the purified full-length S protein in both the prefusion and postfusion conformations⁹, as well as those of the S proteins present on chemically inactivated SARS-CoV-2 virions¹⁹⁻²¹, were subsequently

determined, revealing additional structural details. The fusion fragment S2 contains several segments that play important structural and functional roles, including the FP, heptad repeat 1 (HR1), central helix (CH), connector domain (CD), heptad repeat 2 (HR2), TM and cytoplasmic tail (CT) (Extended Data Fig. 1). In the postfusion trimer of the S2 ectodomain, HR1 and CH form a central three-stranded coiled coil, $\sim 180\text{\AA}$ long⁹. Part of HR2 adopts a helical conformation and packs against the groove of the HR1 coiled-coil to form a six-helix bundle and stabilize the hairpin-like postfusion structure. It is consistent with a membrane fusion model, in which HR1 undergoes a “jack-knife” transition to insert the FP into the target cell membrane and HR2 folds back to bring the FP and TM segments close together⁶, in turn causing the two membranes to fuse into a single lipid bilayer. In all previous structures, however, the regions of the S protein near the viral membrane, while structurally and functionally critical²²⁻²⁵, are either not present or disordered.

Among the membrane-interacting regions of coronavirus spike proteins, the FP is one of the most crucial structural elements for membrane fusion, but its exact location has been the source of various competing theories²⁶. Three membranotropic regions in the SARS-CoV (severe acute respiratory syndrome coronavirus) S2 have been suggested as putative fusion peptides (Extended Data Fig. 1), including a potentially glycosylated segment upstream of the S2' cleavage site (residues 770-788, named the N-terminal FP; ref^{27,28}); the segment immediately downstream of the S2' cleavage site (residues 798-816 or 798-835, widely accepted as the “*bona fide*” FP; ref^{29,30}); and the segment immediately upstream of HR1 (residues 858-886, also known as the internal FP; ref^{24,28,31}). Early studies supported the assignment of the internal FP because it had strong membrane-perturbing capacities and mutations in this region led to inhibition of the S-mediated cell-cell fusion²⁴. Subsequent studies have shown that the *bona fide* FP, highly conserved among coronaviruses, has even a stronger calcium-dependent membrane ordering activity -- a main characteristic of a viral FP when studied as a short peptide in solution -- than the internal FP^{29,30}. Recently, several broadly neutralizing antibodies have been identified, recognizing the *bona fide* FP region of spike proteins from all known human-infecting coronaviruses, and the fusion peptide has therefore been suggested as a potential target for developing universal coronavirus vaccines^{32,33}. A fourth hydrophobic region (residues 1190-1203, also called the pretransmembrane domain (pre-TM) or aromatic domain) adjacent to the TM domain has been shown to be important in SARS-CoV fusion, and may act in concert with the FP to support membrane fusion^{34,35}.

To visualize the membrane interacting regions of postfusion SARS-CoV-2 spike protein, we have reconstituted the full-length protein in lipid-based nanodiscs, induced the conformational changes using soluble ACE2 to prepare a sample of the postfusion spike in lipid bilayer, and determined its structure by cryogenic electron microscopy (cryo-EM) to reveal structural details critical for a full understanding of the SARS-CoV-2 entry. This study shows that the true fusion peptide, as defined by target-membrane insertion, is the segment containing residues 866-910. We therefore use the designation “fusion peptide” to refer to that segment exclusively.

Results

Postfusion S2 trimer in lipid nanodiscs

We reconstituted the purified, intact protein derived from the early variant G614 (B.1) in lipid nanodiscs using a circularized membrane scaffold protein (MSP; ref^{36,37}). The reconstituted sample resolved by gel-filtration chromatography into three major peaks. SDS-PAGE analysis showed that peak I contained mainly the cleaved S protein in nanodiscs, peak II, empty nanodiscs without the S protein, and peak III, the free MSP (Fig. 1a). Negative stain EM confirmed the identity of peak I as the nanodisc-associated prefusion S trimer, showing some flexibility between the ectodomain and the TM region. We induced the conformational changes of the prefusion S trimer by incubating the reconstitution reaction or the purified trimer in nanodiscs with soluble ACE2 (Fig. 1b, Extended Data Fig. 2). The ACE2-treated sample also resolved into three peaks by gel-filtration chromatography. Peak I contained primarily the dissociated S2 in nanodiscs, peak II, empty nanodiscs co-eluting with dissociated S1 in complex with ACE2, and peak III, unbound ACE2 co-eluting with free MSP (Fig. 1b, Extended Data Fig. 2). Negative stain EM images of the peak I fraction showed a very rigid, postfusion S2 trimer projecting from the nanodisc. The S2' site in the postfusion S2 trimer appeared to be uncleaved, as judged by SDS-PAGE analysis and N-terminal sequencing. These results demonstrate that ACE2 binding is sufficient to trigger the prefusion-to-postfusion conformational transition of the purified S trimer even in the context of the restricted bilayer membrane in a nanodisc.

Structure of postfusion S2 in membrane

We determined by cryo-EM the structure of postfusion S2 trimer in nanodiscs, prepared as just described. The terminology for various segments of the S2 polypeptide chain are depicted in Fig. 2a. We recorded cryo-EM images on a Titan Krios electron microscope equipped with a Gatan K3 direct electron detector and used cryoSPARC³⁸ for particle picking, two-dimensional (2D) classification, three-dimensional (3D) classification and refinement (Extended Data Fig. 3). One major class was obtained from 3D classification and refined to 2.9Å resolution (Extended Data Figs. 3-5; Fig. 2b; Extended Data Table 1). To improve the local resolution near the nanodisc, we performed additional masked local refinement, leading to a 3.3 Å map covering the transmembrane and cytoplasmic tail regions.

The overall structure of the postfusion S2 ectodomain in a nanodisc is nearly identical to that of the protein solubilized in detergent⁹ and that of SARS-CoV¹⁰. It also fits very well in the low-resolution map of the postfusion spike on the surface of SARS-CoV-2 virion, reconstructed by cryo electron tomography (Extended Data Fig. 6; ref²⁰). As reported previously, the core structure is a long central three-stranded coiled coil made up by HR1 and CH (Fig. 2c). A three-stranded β sheet formed by a β hairpin from the connector domain (CD), and a segment (residues 718-729; $\beta^{718-729}$) from the S1/S2-S2' fragment, created by proteolytic cleavages, wraps around the C-terminal end of the CH coiled-coil. This connector β sheet and CH form the invariant structure between the prefusion and postfusion structures. Another segment (residues 737-769) in the S1/S2-S2' fragment folds into three consecutive α -helices (3H), locked by two disulfide bonds, and also tightly packing against the groove of the CH coiled-coil. In our new structure, there is additional density for the segment immediately downstream 3H, which bends back towards the TM region and attaches to the surface of the postfusion structure, becoming disordered after Gln785. The S2' site near Arg815, presumably cleaved in this conformational state, remains invisible. In the C-terminal half of S2, a β strand formed by residues 1127-1135 ($\beta^{1127-1135}$) upstream of HR2 from another protomer joins the connector β sheet to expand it into four strands, projecting HR2 towards the TM region and probably initiating the HR2 folding back. Furthermore, a two-turn helix formed by residues 1148-1155 ($\alpha^{1148-1155}$) wedges between two neighboring 3Hs and a longer helix of HR2 makes up the six-helix bundle with the HR1 coiled-coil, together reinforcing the very rigid postfusion structure.

The membrane-interacting segments of S2, missing in all previous postfusion structures of any class I viral fusion proteins, are well resolved in our new structure. There are nine membrane-spanning helices (three per protomer) in the nanodisc region, which were fully resolved in the refined maps. The FP region immediately upstream of HR1 forms a continuous α -helix extending the central coiled-coil to $\sim 218\text{\AA}$ long and well into the lipid bilayer (Fig. 2c), possibly accounting for the rigidity of the entire postfusion structure including the transmembrane region. This first FP membrane-spanning helix is followed by a sharp U-turn within the lipid bilayer, and another helix that spans through the membrane once again and sends the adjacent FPPR back to the ectodomain side of the membrane. A total of six transmembrane helices of the FP from three protomers pack tightly together to form a blunted cone shape (Fig. 2c), probably required for effectively penetrating the target cell membrane. The third membrane-spanning helix is the TM segment, which tilts relative to the plane of the membrane, gently wrapping around the blunted cone. Part of the following CT appears to be embedded horizontally in the cytosolic headgroup region of the lipid bilayer, and three copies of it form a triangle that caps the tip of the transmembrane cone (Fig. 2c). Comparison of the prefusion and postfusion structures indicates that formation of the long postfusion central helix translocates by 185\AA the FP, which reconfigures to become a hairpin-like wedge inserted into the target membrane (Fig. 2d). The FP not only inserts into the target membrane but also becomes a docking site for the TM helix in the final stages of membrane merger.

Definition of the SARS-CoV-2 FP and TM

The local refined maps with sidechain density for many residues in the transmembrane region allowed unambiguous modeling (Extended Data Fig. 5). It is clear that both the segment⁷⁸⁸⁻⁸⁰⁶ and segment⁸¹⁶⁻⁸³⁴ (Supplementary Figure 2) are disordered but located outside the membrane (Fig. 2c), and thus they do not function as the real fusion peptide. It is the FP that forms a hairpin-like wedge in a trimeric cone-shaped assembly to span almost the entire lipid bilayer (Fig. 3a). As expected, the membrane-inserted region of the wedge is largely hydrophobic with Trp886 and Phe888 at the blunted tip while there are several charged residues (Asp867, Glu868 and Arg905) at the base of the wedge located in the headgroup region of lipid bilayer. Thus, total 43 residues (Asp867 - Ile909) insert into membrane and make up the functional fusion peptide. Interestingly, Pro897 introduces a kink in the long central helix, making the region near the tip of the fusion wedge deviated from a canonical coiled-coil structure and also allowing the tight packing of the

returning helix (residues 870-882) through an ⁸⁷⁶AXXXG⁸⁸⁰ motif to form the blunted cone shape (Fig. 3a). The interactions among the FPs from the three protomers are primarily mediated by polar residues (Thr881 and Gln 895) before the Pro897-introduced kink, but by hydrophobic residues (Phe898, Met902, Phe906 and Ile909) after the kink (Extended Data Fig. 7).

For the transmembrane anchor, the TM helix begins at Tyr1215 and ends at Leu1234, and Trp1212 and Trp1217 from the pre-TM are embedded entirely in the lipid bilayer, indicating that the pre-TM is part of the TM domain, not a separate structural element. In addition, Cys1235 and Cys1236 are clearly not part of the TM helix and mark the beginning of the CT (Fig. 3a). These two cysteine residues, possibly palmitoylated in the virus^{39,40}, are also very close to Cys1247 and Cys1248 from the CT of a neighboring protomer, although formation of disulfide bonds between them on the cytosolic side is unlikely.

Interactions among the membrane regions

It is postulated that the FP of a class I viral fusion protein may interact with its TM domain in the final stage of the fusion process, but there has been no direct structural evidence for this interaction in an intact protein in the context of membrane. The structure reported here shows an intimate FP-TM interaction in the SARS-CoV-2 postfusion spike. The TM segment packs against the fusion peptide from another protomer with two major contacts. First, Trp1212, Trp1217 and Phe1220 near the ectodomain side pack against the base of the fusion wedge; the contact may also involve some phospholipids, as suggested by strong elongated density in the vicinity, aligned perpendicular to the membrane plane (Fig. 3b). Second, Leu1234 near the CT side interacts with Trp886 and Phe888 from two different protomers at the tip of the fusion wedge (Fig. 3c). Three copies of the CT fragment come together to cap the tip of the fusion wedge.

The segment (residues 857-866) immediately upstream of the FP lies in the headgroup region of the lipid bilayer with several hydrophobic residues pointing towards the hydrophobic core of the membrane. This configuration probably helps orient the adjacent disulfide-locked FPPR (residues 835-856) in a way with its C-terminal half embedded in the membrane as well (Fig. 3d, 3e). The N-terminal half of the FPPR, directly connected to the disordered s⁸¹⁶⁻⁸³⁴, projects away from the membrane. In addition, the FPPR packs against the FP wedge and Asp848 from the FPPR forms

a salt bridge with Arg905 of the FP. Thus, the FPPR, which facilitates clamping down of the RBD in the prefusion conformation⁹, may also contribute to anchoring the membrane-interacting regions of the postfusion S2 structure.

Functional effect of designed mutations

To test the functional roles of the membrane-interacting structural elements, we have assessed impacts of several structure-guided mutations in the s⁸¹⁶⁻⁸³⁴, FP, FPPR and CT regions. As illustrated in Fig. 4a, we substituted multiple hydrophobic residues in the b-FP, which would in principle be critical for membrane interaction, with charged residues (Mut-1); we mutated the key residues that participate in the FPPR interaction with the FP (Mut-2); we introduced multiple charged residues in the FP (Mut-3 and Mut-4) and mutated its AXXXG motif (Mut-5 and Mut-6); and we changed the two cysteine residues that mark the beginning of the CT (Mut-7).

When transiently transfected in 293T cells, these mutants expressed the same level of S, but the extent of cleavage between S1 and S2 was substantially reduced for Mut-3, Mut-4, Mut-5 and Mut-6, all of which are in the FP (Extended Data Fig. 8a). Like fusion peptides and fusion loops of many other viral fusion proteins⁴¹⁻⁴³, the FP is buried in the prefusion conformation, in the present case by a segment near the N-terminus of S2 from another promoter^{9,12,13}. When detected by monoclonal antibodies using flow cytometry (Extended Data Fig. 8b), the recognition patterns by RBD-specific antibodies (C63C8, G32B6, C63C7, REGN10933 and REGN10987; ref^{44,45}), NTD-specific antibodies (C12C9, 4A8 and C81D6; ref^{44,46}), S2-specific antibodies (C163E6 and 3H3; ref^{44,46}) and ACE2 for Mut-1, Mut-2 and Mut-7 were very similar to that of the wildtype S, except that the overall binding levels were substantially lower for Mut-1, suggesting that the mutations had some impact on protein secretion. The secretion level of Mut-3 was very low, as judged by binding to all antibodies and ACE2, indicating that this mutant protein might not be folded correctly. The patterns for binding to the antibodies and ACE2 were similar among Mut-4, Mut-5 and Mut-6, but they bound more weakly to RBD-specific antibodies and the neutralizing NTD-specific antibody, 4A8, while much more strongly to S2-specific antibodies and to the nonneutralizing NTD-specific antibody, C81D6, than did the wildtype S, consistent with large structural rearrangements caused by the mutations in these proteins. These results suggest that the FP does not tolerate multiple mutations in the prefusion conformation, consistent with its buried

position. In a S-mediated cell-cell fusion assay⁹, Mut-1 showed reduced fusion activity while Mut-7 behaved like a wildtype spike; all other mutants showed no detectable fusion activity (Fig. 4b). Thus, as our structure shows, the s⁸¹⁶⁻⁸³⁴ is not the fusion peptide and does not interact with membrane, as introducing multiple charged residues does not abolish membrane fusion. Moreover, the two cysteine residues in the CT close to the TMD, which have the potential to form interchain disulfide bonds, are dispensable for membrane fusion. Mut-3, Mut-4, Mut-5 and Mut-6 lost their membrane fusion activity, but their mutations also had substantial impact on the prefusion structures of the mutant proteins. We next created additional constructs (Mut-8 – Mut-15), each with a single point mutation in the FP, and identified three mutants (Mut-8, Mut-11 and Mut-14) that preserved the prefusion conformation but showed almost no detectable membrane fusion (Fig. 4c; Extended Data Fig. 8c,d). Similar results were obtained when the mutations of Mut-1, Mut-2, Mut-8, Mut-11 and Mut-14 were introduced in the Delta variant spike (Fig. 4d; Extended Data Fig. 8e,f). These results confirm that the FP is a key functional structural element in the SARS-CoV-2 spike.

Discussion

Our structure of the SARS-CoV-2 postfusion S2, with membrane-associated elements interacting with a lipid bilayer, has resolved a long-standing contentious issue in the coronavirus field regarding the functional identity of the fusion peptide. Although both the s⁷⁸⁸⁻⁸⁰⁶ and FP were suggested as candidates, the s⁸¹⁶⁻⁸³⁴ subsequently became widely accepted as “*bona fide*” because it was more conserved among coronaviruses and showed a stronger activity than the other two in perturbing membrane structure as a synthetic peptide^{29,30}. Our results show unambiguously that only the FP interacts with membrane. They underscore the risk of studying fusion peptides when taken out of the context of intact viral fusion proteins, because many hydrophobic peptides can induce structural changes of a lipid bilayer. For example, the FPPR, also identified as the “fusion peptide 2” for SARS-CoV in some studies³⁰, projects partway into the membrane headgroup region and could therefore modify the local membrane structure, but it does not insert into the hydrophobic part of the lipid bilayer. From our structure, neither the s⁷⁸⁸⁻⁸⁰⁶ nor the s⁸¹⁶⁻⁸³⁴ inserts into the membrane in the postfusion structure and therefore neither is likely to function as a real “fusion peptide”. The conclusion is further supported by the evidence that introducing multiple charged residues in the s⁸¹⁶⁻⁸³⁴ does not abolish the membrane fusion activity of the full-length S

and by the presence of a conserved N-linked glycosylation site in the s⁷⁸⁸⁻⁸⁰⁶. We cannot rule out the possibility that the s⁷⁸⁸⁻⁸⁰⁶ and s⁸¹⁶⁻⁸³⁴ may transiently interact with the membrane during the fusion process. Nevertheless, membrane insertion of the FP does not seem to depend on the s⁷⁸⁸⁻⁸⁰⁶ or s⁸¹⁶⁻⁸³⁴ since it is probably directly coupled with the refolding of adjacent HR1 into the coiled-coil. Moreover, a fusion peptide may not have to be fully conserved in its amino acid sequence as long as it maintains the ability to insert correctly into the target membrane. Indeed, there is a four-residue deletion at the tip of the FP in some coronaviruses, including HCoV-OC43 and HCoV-229E (Extended Data Fig. 6), but it is not likely to alter the overall shape of the fusion wedge and thus its conserved function.

The substantial lipid-bilayer interaction surface of the fusion wedge observed here is consistent with the stability of the extended intermediate, shown to persist for many minutes, even after release of S1 and its associated ACE2, when SARS-CoV-2 binds at the cell surface, unless the pH falls below neutral⁴⁷. Fusion peptides with less extensive hydrophobic contacts, such as those postulated for the influenza virus fusion peptide⁴⁸, would probably have shorter lifetimes.

Our structure also provides a more accurate structural definition of the TMD, as well as the direct evidence for interactions between the FP and TM in the merged membrane for a class I viral fusion protein. The previously designated “pre-TM” segment, now shown to be part of the TM, contributes to packing of the TM against the FP, explaining its functional role in membrane fusion. One study reported that a single residue insertion into the TM of SARS-CoV S led to a complete loss in membrane fusion⁴⁹, and the observation was difficult to interpret based on the presumptive trimeric TMD structure in the prefusion conformation⁵⁰. The insertion was between G1201 and F1202 of SARS-CoV, which correspond to G1219 and F1220 in SARS-CoV-2. Such an insertion would cause a rotation of the C-terminal portion of the TM helix possibly together with the CT in the postfusion structure, thereby disrupting the interactions between the FP and the TM and CT and blocking the last step of membrane fusion. These results, together with our new structure, highlight the functional importance of the interactions among the membrane-interacting elements within the lipid bilayer, which have not been recognized previously. Nevertheless, it remains an open question how and to what extent these interactions functionally contribute to membrane fusion.

Finally, our structure suggests that the real fusion peptide of SARS-Cov-2 spike is unlikely a useful vaccine target since it is well protected and not accessible by antibodies in the prefusion conformation^{9,12,13}, nor is it exposed after inserting into the target cell membrane in the prehairpin intermediate state. The s⁸¹⁶⁻⁸³⁴, targeted by several broadly neutralizing antibodies^{32,33}, may be conserved for other structural or functional reasons, which will require further investigation. Moreover, the interactions among the FP and FPPR, TM and CT shown in our structure, may provide some novel therapeutic targets for developing peptide-based or small-molecule fusion inhibitors. In particular, mutations in the FPPR, designed to disrupt its interaction with the FP, completely abolish fusion activity, raising the possibility of identifying therapeutic candidates that target this interaction and block viral infection.

ACCELERATED ARTICLE PREVIEW

References

- 1 Harrison, S. C. Viral membrane fusion. *Virology* **479-480**, 498-507, doi:10.1016/j.virol.2015.03.043 (2015).
- 2 Kielian, M. Mechanisms of Virus Membrane Fusion Proteins. *Annu Rev Virol* **1**, 171-189, doi:10.1146/annurev-virology-031413-085521 (2014).
- 3 Rand, R. P. & Parsegian, V. A. Physical force considerations in model and biological membranes. *Can J Biochem Cell Biol* **62**, 752-759 (1984).
- 4 Parsegian, V. A., Fuller, N. & Rand, R. P. Measured work of deformation and repulsion of lecithin bilayers. *Proc Natl Acad Sci U S A* **76**, 2750-2754 (1979).
- 5 Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat Rev Mol Cell Biol* **23**, 3-20, doi:10.1038/s41580-021-00418-x (2022).
- 6 Tortorici, M. A. & Veesler, D. Structural insights into coronavirus entry. *Adv Virus Res* **105**, 93-116, doi:10.1016/bs.aivir.2019.08.002 (2019).
- 7 Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278, doi:10.1016/j.cell.2020.02.052 (2020).
- 8 Zhao, M. M. *et al.* Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduct Target Ther* **6**, 134, doi:10.1038/s41392-021-00558-8 (2021).
- 9 Cai, Y. *et al.* Distinct conformational states of SARS-CoV-2 spike protein. *Science* **369**, 1586-1592, doi:10.1126/science.abd4251 (2020).
- 10 Fan, X., Cao, D., Kong, L. & Zhang, X. Cryo-EM analysis of the post-fusion structure of the SARS-CoV spike glycoprotein. *Nat Commun* **11**, 3618, doi:10.1038/s41467-020-17371-6 (2020).
- 11 Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269, doi:10.1038/s41586-020-2008-3 (2020).
- 12 Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263, doi:10.1126/science.abb2507 (2020).
- 13 Walls, A. C. *et al.* Structure, function and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*, DOI: 10.1016/j.cell.2020.1002.1058 (2020).
- 14 Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220, doi:10.1038/s41586-020-2180-5 (2020).
- 15 Yan, R. *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444-1448, doi:10.1126/science.abb2762 (2020).
- 16 Shang, J. *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224, doi:10.1038/s41586-020-2179-y (2020).
- 17 Wang, Q. *et al.* Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **181**, 894-904 e899, doi:10.1016/j.cell.2020.03.045 (2020).
- 18 Xia, S. *et al.* Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res* **30**, 343-355, doi:10.1038/s41422-020-0305-x (2020).
- 19 Ke, Z. *et al.* Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* **588**, 498-502, doi:10.1038/s41586-020-2665-2 (2020).
- 20 Yao, H. *et al.* Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **183**, 730-738 e713, doi:10.1016/j.cell.2020.09.018 (2020).

- 21 Liu, C. *et al.* The Architecture of Inactivated SARS-CoV-2 with Postfusion Spikes Revealed by Cryo-EM and Cryo-ET. *Structure* **28**, 1218-1224.e1214 (2020).
- 22 Bosch, B. J., de Haan, C. A., Smits, S. L. & Rottier, P. J. Spike protein assembly into the coronavirus: exploring the limits of its sequence requirements. *Virology* **334**, 306-318, doi:10.1016/j.virol.2005.02.001 (2005).
- 23 Lontok, E., Corse, E. & Machamer, C. E. Intracellular targeting signals contribute to localization of coronavirus spike proteins near the virus assembly site. *J Virol* **78**, 5913-5922, doi:10.1128/JVI.78.11.5913-5922.2004 (2004).
- 24 Petit, C. M. *et al.* Genetic analysis of the SARS-coronavirus spike glycoprotein functional domains involved in cell-surface expression and cell-to-cell fusion. *Virology* **341**, 215-230, doi:10.1016/j.virol.2005.06.046 (2005).
- 25 Ye, R., Montalto-Morrison, C. & Masters, P. S. Genetic analysis of determinants for spike glycoprotein assembly into murine coronavirus virions: distinct roles for charge-rich and cysteine-rich regions of the endodomain. *J Virol* **78**, 9904-9917, doi:10.1128/JVI.78.18.9904-9917.2004 (2004).
- 26 Belouzard, S., Millet, J. K., Licitra, B. N. & Whittaker, G. R. Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses* **4**, 1011-1033, doi:10.3390/v4061011 (2012).
- 27 Sainz, B., Jr., Rausch, J. M., Gallaher, W. R., Garry, R. F. & Wimley, W. C. Identification and characterization of the putative fusion peptide of the severe acute respiratory syndrome-associated coronavirus spike protein. *J Virol* **79**, 7195-7206, doi:10.1128/JVI.79.11.7195-7206.2005 (2005).
- 28 Basso, L. G., Vicente, E. F., Crusca, E., Jr., Cilli, E. M. & Costa-Filho, A. J. SARS-CoV fusion peptides induce membrane surface ordering and curvature. *Sci Rep* **6**, 37131, doi:10.1038/srep37131 (2016).
- 29 Madu, I. G., Roth, S. L., Belouzard, S. & Whittaker, G. R. 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, doi:10.1128/JVI.00079-09 (2009).
- 30 Lai, A. L., Millet, J. K., Daniel, S., Freed, J. H. & Whittaker, G. R. The SARS-CoV Fusion Peptide Forms an Extended Bipartite Fusion Platform that Perturbs Membrane Order in a Calcium-Dependent Manner. *J Mol Biol* **429**, 3875-3892, doi:10.1016/j.jmb.2017.10.017 (2017).
- 31 Bosch, B. J. *et al.* Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. *Proc Natl Acad Sci U S A* **101**, 8455-8460, doi:10.1073/pnas.0400576101 (2004).
- 32 Dacon, C. *et al.* Broadly neutralizing antibodies target the coronavirus fusion peptide. *Science* **377**, 728-735, doi:10.1126/science.abq3773 (2022).
- 33 Low, J. S. *et al.* ACE2-binding exposes the SARS-CoV-2 fusion peptide to broadly neutralizing coronavirus antibodies. *Science* **377**, 735-742, doi:10.1126/science.abq2679 (2022).
- 34 Guillen, J., Perez-Berna, A. J., Moreno, M. R. & Villalain, J. Identification of the membrane-active regions of the severe acute respiratory syndrome coronavirus spike membrane glycoprotein using a 16/18-mer peptide scan: implications for the viral fusion mechanism. *J Virol* **79**, 1743-1752, doi:10.1128/JVI.79.3.1743-1752.2005 (2005).

- 35 Sainz, B., Jr., Rausch, J. M., Gallaher, W. R., Garry, R. F. & Wimley, W. C. The aromatic domain of the coronavirus class I viral fusion protein induces membrane permeabilization: putative role during viral entry. *Biochemistry* **44**, 947-958, doi:10.1021/bi048515g (2005).
- 36 Grinkova, Y. V., Denisov, I. G. & Sligar, S. G. Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Eng Des Sel* **23**, 843-848, doi:10.1093/protein/gzq060 (2010).
- 37 Johansen, N. T. *et al.* Circularized and solubility-enhanced MSPs facilitate simple and high-yield production of stable nanodiscs for studies of membrane proteins in solution. *FEBS J* **286**, 1734-1751, doi:10.1111/febs.14766 (2019).
- 38 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).
- 39 Ramadan, A. A. *et al.* Identification of SARS-CoV-2 Spike Palmitoylation Inhibitors That Results in Release of Attenuated Virus with Reduced Infectivity. *Viruses* **14**, doi:10.3390/v14030531 (2022).
- 40 Tien, C. F. *et al.* Glycosylation and S-palmitoylation regulate SARS-CoV-2 spike protein intracellular trafficking. *iScience* **25**, 104709, doi:10.1016/j.isci.2022.104709 (2022).
- 41 Modis, Y., Ogata, S., Clements, D. & Harrison, S. C. Structure of the dengue virus envelope protein after membrane fusion. *Nature* **427**, 313-319, doi:10.1038/nature02165 (2004).
- 42 Steinhauer, D. A., Wharton, S. A., Skehel, J. J. & Wiley, D. C. Studies of the membrane fusion activities of fusion peptide mutants of influenza virus hemagglutinin. *J Virol* **69**, 6643-6651, doi:10.1128/JVI.69.11.6643-6651.1995 (1995).
- 43 Wilson, I. A., Skehel, J. J. & Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366-373, doi:10.1038/289366a0 (1981).
- 44 Tong, P. *et al.* Memory B cell repertoire for recognition of evolving SARS-CoV-2 spike. *Cell* **184**, 4969-4980 e4915, doi:10.1016/j.cell.2021.07.025 (2021).
- 45 Baum, A. *et al.* Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* **369**, 1014-1018, doi:10.1126/science.abd0831 (2020).
- 46 Chi, X. *et al.* A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **369**, 650-655, doi:10.1126/science.abc6952 (2020).
- 47 Kreuzberger, A. J. B. *et al.* SARS-CoV-2 requires acidic pH to infect cells. *Proc Natl Acad Sci U S A* **119**, e2209514119, doi:10.1073/pnas.2209514119 (2022).
- 48 Lorieau, J. L., Louis, J. M. & Bax, A. The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface. *Proc Natl Acad Sci U S A* **107**, 11341-11346, doi:10.1073/pnas.1006142107 (2010).
- 49 Corver, J., Broer, R., van Kasteren, P. & Spaan, W. Mutagenesis of the transmembrane domain of the SARS coronavirus spike glycoprotein: refinement of the requirements for SARS coronavirus cell entry. *Virol J* **6**, 230, doi:10.1186/1743-422X-6-230 (2009).
- 50 Fu, Q. & Chou, J. J. A Trimeric Hydrophobic Zipper Mediates the Intramembrane Assembly of SARS-CoV-2 Spike. *J Am Chem Soc* **143**, 8543-8546, doi:10.1021/jacs.1c02394 (2021).

- 51 Zhang, J. *et al.* Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science* **372**, 525-530, doi:10.1126/science.abf2303 (2021).
- 52 Xiao, T. *et al.* A trimeric human angiotensin-converting enzyme 2 as an anti-SARS-CoV-2 agent. *Nat Struct Mol Biol* **28**, 202-209, doi:10.1038/s41594-020-00549-3 (2021).
- 53 Chen, J. *et al.* HIV-1 ENVELOPE. Effect of the cytoplasmic domain on antigenic characteristics of HIV-1 envelope glycoprotein. *Science* **349**, 191-195, doi:10.1126/science.aaa9804 (2015).
- 54 Ritchie, T. K. *et al.* Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol* **464**, 211-231, doi:10.1016/S0076-6879(09)64011-8 (2009).
- 55 Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006 (2012).
- 56 Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51, doi:10.1016/j.jsb.2005.07.007 (2005).
- 57 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).
- 58 Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530, doi:10.1107/S2059798318002425 (2018).
- 59 Morin, A. *et al.* Collaboration gets the most out of software. *Elife* **2**, e01456, doi:10.7554/eLife.01456 (2013).
- 60 Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636-W641, doi:10.1093/nar/gkz268 (2019).
- 61 Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324, doi:10.1093/nar/gku316 (2014).

Figure Legends

Figure 1. Preparation of the membrane-bound SARS-CoV-2 spike protein in lipid nanodiscs.

(a) The purified full-length SARS-CoV-2 S protein was reconstituted in lipid nanodiscs and resolved by gel-filtration chromatography on a Superose 6 column. Three major peaks (not at the void volume) are peak I containing the prefusion S trimer in nanodiscs, peak II containing empty nanodiscs, and peak III containing free membrane scaffold protein (MSP), as analyzed by SDS-PAGE. Representative 2D averages by negative stain EM of the peak I fractions are also shown. The box size of 2D averages is $\sim 880\text{\AA}$. (b) The purified full-length SARS-CoV-2 S protein was first reconstituted in nanodiscs, incubated with soluble ACE2 and then resolved by gel-filtration chromatography on a Superose 6 column. Three non void-volume peaks are peak I containing the postfusion S2 trimer in nanodiscs, peak II containing empty nanodiscs and dissociated S1 in complex with ACE2, and peak III containing unbound ACE2 and free MSP, as analyzed by SDS-PAGE. Representative 2D averages by negative stain EM of the peak I fractions are also shown. The box size of 2D averages is $\sim 880\text{\AA}$. The experiments in both (a) and (b) were repeated at least four times independently with similar results. For gel source data, see Supplementary Figure 1.

Figure 2. Cryo-EM structure of the SARS-CoV-2 S2 in the postfusion conformation.

(a) Schematic representation of the SARS-CoV-2 S2 fragment. Various segments include: S1/S2, S1/S2 cleavage site; S2', S2' cleavage site; $\beta^{718-729}$, a β -strand formed by residues 718-729 in the S1/S2-S2' fragment; 3H, three-helix segment; s816-834, a segment containing residues 816-834; FPPR, fusion peptide proximal region; FP, the fusion peptide; HR1, heptad repeat 1; CH, central helix region; CD, connector domain; $\beta^{1127-1135}$, a β -strand formed by residues 1127-1135; $\alpha^{1148-1155}$, an α -helix formed by residues 1148-1155; HR2, heptad repeat 2; TM, transmembrane anchor; CT, cytoplasmic tail; and tree-like symbols for glycans. (b) The structure of the S2 trimer fits into a 2.9\AA density map. Three protomers (A, B, C) are colored in green, blue and red, respectively. (c) Overall structure of the S2 trimer in the postfusion conformation shown in ribbon diagram. Various structural components in the color scheme shown in (a) include 3H, FPPR, FP, HR1, CH, CD, HR2, TM and CT. The S2' cleavage site is in a disordered segment between Gln785 and Gly832, indicated by a dashed line. A top view of the transmembrane region in the nanodisc is also shown.

(d) One protomer each of the S2 trimer from the prefusion and the postfusion conformations superposed by the invariant CH and CD region. The locations of HR2, TM and CT disordered in the prefusion structure, as well as the b-FP in the postfusion structure are indicated by colored rectangles. The N-terminal end of HR1, connected directly to the FP, translocates by $\sim 185\text{\AA}$ during the prefusion-to-postfusion transition.

Figure 3. Structural details of the postfusion SARS-CoV-2 spike. (a) Left, the transmembrane region of the postfusion S2 fits in the EM density with one protomer colored: HR1 in yellow, FPPR in green, FP in red, HR2 in orange, TM in magenta and CT in light green. Right, the atomic models for the FP and TM segment with key residues indicated. (b) Interactions between the TM and FP near the ectodomain side of the membrane with key interface residues shown in stick model. The elongated density potentially for ordered phospholipids are shown in gray. (c) Interactions between the TM and FP near the CT side of the membrane with key interface residues shown in stick model. (d) Attachment of the FPPR and the segment of residues 857-866 to the lipid bilayer with the hydrophobic residues facing towards the hydrophobic core of the membrane highlighted in stick model. (e) Interactions between the FPPR and FP on the ectodomain side of the membrane with key interface residues shown in stick model.

Figure 4. Effect of mutations in the membrane-interacting segments on membrane fusion. (a) Structure-guided design of mutations in the S⁸¹⁶⁻⁸³⁴, FPPR, FP and CT regions. The wildtype sequences are shown in black and the mutated residues in red. (b) and (c) Cell-cell fusion mediated by the G614 S and designed mutants. HEK293T cells transfected with the full-length S protein expression plasmids were fused with ACE2-expressing cells. Cell-cell fusion led to reconstitution of α and ω fragments of β -galactosidase to form an active enzyme, and the fusion activity was then quantified by a chemiluminescent assay. No ACE2 were negative controls. The experiment was repeated four times with similar results. Error bars, mean \pm s.d.. (d) Cell-cell fusion mediated by the Delta variant S and its mutants. Negative controls include no spike/no ACE2, no spike, no ω fragment of β -galactosidase, and no α fragment of β -galactosidase. The experiment was repeated three times with similar results. Error bars, mean \pm s.d..

Methods

Expression constructs

The expression constructs for the full-length spike (S) protein (residue 1-1273) of SARS-CoV-2 (G614) and monomeric soluble ACE2 protein were described previously^{51,52}. To produce membrane scaffold protein (MSP), a modified gene encoding MSP2N2^{36,37} fused with a 7×His tag and a TEV site at N-terminal end and a sortase eSrt site and a 6×His tag at C-terminal end, was synthesized and cloned into the pET28a vector between *Nco* I and *Xho* I restriction sites to create the construct pET28a-IsMSP2N2 by GenScript (Piscataway, NJ). The expression constructs for the mutant spike proteins were generated by standard PCR methods and verified by DNA sequencing of the entire coding regions.

Expression and purification of recombinant proteins

A stable cell line was generated using HEK 293T cells for large-scale production of the full-length G614 S protein following the published protocol⁵³. Purification of the full-length S protein was carried out as previously described⁵¹. Briefly, the stably transfected cells were grown in Expi293 expression medium (Thermo Fisher Scientific, Waltham, MA) containing 1% Pen Strep (Thermo Fisher Scientific) and 1.0 µg/ml puromycin (Thermo Fisher Scientific) to a density of $\sim 4-5 \times 10^6$ /ml, then lysed in a lysis buffer containing Buffer A (100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA) and 1% (w/v) n-dodecyl-β-D-maltopyranoside (DDM) (Anatrace, Inc. Maumee, OH), EDTA-free complete protease inhibitor cocktail (Roche, Basel, Switzerland), and incubated at 4°C for one hour. After centrifugation at 27,000 ×g for 30 min, the supernatant was loaded onto a streptactin (IBA Lifesciences, Göttingen, Germany) column equilibrated with the lysis buffer. The column was washed with 50 column volumes of Buffer A and 0.3% DDM, followed by additional washes with 50 column volumes of Buffer A and 0.1% DDM, and with 50 column volumes of Buffer A and 0.02% DDM. The S protein was eluted by Buffer A containing 0.02% DDM and 5 mM desthiobiotin (IBA Lifesciences). The protein was further purified by gel filtration chromatography on a Superose 6 10/300 column (GE Healthcare, Chicago, IL) in a buffer containing 25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.02% DDM.

Another stable cell line was also generated for large-scale production of monomeric soluble ACE2 protein. Purification of the ACE2 protein was carried out as previously described⁵². Briefly, the stably transfected cells were incubated in Expi293 expression medium containing 1% Pen Strep

and 1.0 µg/ml puromycin to a density of $\sim 4.5 \times 10^6$ /mL, the cell supernatant was harvested by centrifugation at 3,000 xg for 30 min and loaded onto a column packed with Ni Sepharose excel resin (Cytiva Life Sciences, Marlborough, MA). The column was washed with a buffer containing 20 mM Tris-HCl, pH 7.5, 300 mM NaCl and 10 mM imidazole. The ACE2 protein was eluted with a buffer containing 300 mM imidazole, and further purified by gel filtration chromatography on HiLoad 16/600 Superdex 200 pg column (GE Healthcare) in 25 mM Tris-HCl, pH 7.5 and 150 mM NaCl.

To produce the linear MSP2N2 protein, the expression construct was transformed into *E. coli* BL21 Star (DE3) (Thermo Fisher Scientific), and the cells were grown at 37 °C in LB medium with 50 µg/mL kanamycin to an OD₆₀₀ of 0.8 and then induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG; Sigma-Aldrich, St. Louis, MO) at 30 °C for 5 h, and harvested by centrifugation. Purification and circularization of the lsMSP2N2 protein were carried out as previously described^{37,54}. Briefly, the cells were completely resuspended in a buffer (40 mM Tris-HCl, pH 8.0, 300 mM NaCl) and lysed by sonication. The cell lysate was clarified by centrifugation at 30,000 ×g for 30 min and the supernatant was loaded onto a column packed with Ni-NTA agarose (Qiagen, Hilden, Germany). The column was sequentially washed with buffer I (40 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1% Trion X-100), buffer II (40 mM Tris-HCl, pH 8.0, 300 mM NaCl, 50 mM sodium cholate, 20 mM imidazole), and buffer III (40 mM Tris-HCl, pH 8.0, 300 mM NaCl, 50 mM imidazole). The lsMSP2N2 protein was eluted with 40 mM Tris-HCl, pH 8.0, 300 mM NaCl, 400 mM imidazole. To remove the N-terminal His-tag, the protein was treated with TEV protease in the TEV buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT) at room temperature for 3 hrs. To produce circularized protein - csMSP2N2 while removing its C-terminal His-tag, the sortase eSrt was mixed with the TEV-treated lsMSP2N2 with a molar ratio of 1:20 in a circularization buffer (30 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM CaCl₂, 1 mM DTT), and incubated at 37 °C for 3 hrs. The circularization reaction was passed through a Ni-NTA agarose column to remove the histagged protein and collect the flow-through, which contained the csMSP2N2 protein. The protein was further purified by anion exchange chromatography on a HiTrap Q HP column (Cytiva Life Sciences), dialyzed against 20 mM Tris-HCl, pH 7.5 and 100 mM NaCl, flash frozen in liquid nitrogen and stored at -80 °C.

The trimeric ACE2 protein and S-specific monoclonal antibodies were described previously^{44,52}

Reconstitution of the spike protein in nanodiscs

To reconstitute the spike protein in nanodiscs, soy extract polar lipid (Avanti, Birmingham, AL) in chloroform was first dried under a nitrogen stream and dried further overnight in a vacuum desiccator. Lipid films were dissolved in a buffer containing 25 mM Tris 7.5, 150 mM NaCl, 2% DDM, 0.4% CHS, vortexed vigorously and sonicated until the solution became clear. The S protein and lipids were mixed and incubated on ice for 30 min. The csMSP2N2 was added to the mixture with a spike:csMSP2N2:lipid molar ratio of 1:8:700 and incubated on ice for another 30 min. Bio-beads SM2 (Bio-Rad, Hercules, CA) were added to remove detergents from the mixture and initiate the reconstitution. After gentle rotation at 4 °C overnight, Bio-beads were removed through filtration. To induce the conformational transition of the S protein from the prefusion state to the postfusion state in the nanodiscs, the reconstitution mixture was added with soluble ACE2 protein at 6 μM and incubated at room temperature for 30 min. A control experiment confirmed the prefusion S protein was stable at room temperature in absence of ACE2. The ACE2-induced postfusion S-nanodisc sample was purified by gel filtration chromatography on Superose 6 Increase 10/300 column (GE Healthcare) in a buffer containing 25 mM Tris-HCl, pH 7.5, 150 mM NaCl. Peak fractions containing the postfusion S-nanodisc sample were used for cryo-EM analysis.

Negative stain EM

To prepare grids, 4 μl of freshly purified prefusion or postfusion S nanodisc sample was adsorbed to a glow-discharged carbon-coated copper grid (Electron Microscopy Sciences, Hatfield, PA), washed with deionized water, and stained with freshly prepared 1.5% uranyl formate. Images were recorded at room temperature on the Phillips CM10 transmission electron microscope with a nominal magnification of 52,000×. Particles were auto-picked and 2D class averages were generated using RELION 4.0.0⁵⁵.

Cryo-EM sample preparation and data collection

To prepare cryo EM grids, the postfusion S-nanodisc sample at 1.5 mg/ml was applied to a 1.2/1.3 Quantifoil gold grid (Electron Microscopy Sciences), which had been glow discharged with a PELCO easiGlow™ Glow Discharge Cleaning system (Ted Pella, Inc.) for 60 s at 15

mA. Grids were immediately plunge-frozen in liquid ethane using a Vitrobot Mark IV (ThermoFisher Scientific), and excess protein was blotted away using grade 595 filter paper (Ted Pella, Inc.) with a blotting time of 4 s, a blotting force of -12 at 4°C with 100% humidity. The grids were screened for ice thickness and particle distribution. Selected grids were used to acquire images with a Titan Krios transmission electron microscope (ThermoFisher Scientific) operated at 300 keV and equipped with a BioQuantum GIF/K3 direct electron detector. Automated data collection was carried out using SerialEM version 4.0.5⁵⁶ at a nominal magnification of 105,000× and the K3 detector in counting mode (calibrated pixel size, 0.825 Å) at an exposure rate of ~13.8 electrons per pixel per second. Each movie adds a total accumulated electron exposure of ~50.6/51.3 e-/Å², fractionated in 50/51 frames. Data sets were acquired using a defocus range of 0.8-2.2 μm.

Image processing and 3D reconstructions

All data were processed using cryoSPARC v.3.3.1³⁸ and RELION. Drift correction for cryo-EM images was performed using patch mode, and contrast transfer function (CTF) estimated by patch mode. Motion corrected sums with dose-weighting were used for all other image processing. Blob picking was performed for particle picking. 18,339,830 particles were extracted from 142,99 images using a box size of 600 Å (downsizing to 128Å) from data set I, and 57,656,504 particles from 17,028 images from data set II. The two sets of particles were processed separately and each subjected to 8-10 rounds of 2D classification, giving 1,532,458 and 3,885,698 good particles, respectively. An initial model was produced using ab-initio reconstruction in cryoSPARC based on selected 2D class averages from data set I.

For data set I, good particles from 2D classification were used for one round of heterogeneous classification with six copies of the initial model as the reference in C1 symmetry. Three major classes with total 57.6% of the particles showing clear structural features were subjected to two additional rounds of heterogeneous refinement with six copies of the initial model as the reference in C1 symmetry. The major class with a nanodisc shape at one end was re-extracted to an unbinned box size of 600Å and subjected to one round of 2D classification to remove bad particles, followed by one round of non-uniform refinement in C1 symmetry to produce a map

at 3.18Å resolution from 153,658 particles. Another round of non-uniform refinement in C3 symmetry was performed and improved the resolution to 3.0Å. To further improve the local resolution in the nanodisc, these particles were transferred to RELION and subjected to one round of focus-classification with a mask covering the entire protein, leading to a major class (78.9% particles) with continuous density at the transmembrane region. This class with 123,298 particles was imported back to cryoSPARC, and subjected to another round of non-uniform refinement in C3 symmetry, giving a map at 2.96Å resolution. One round of particle-subtraction was performed to remove noises from the nanodiscs, followed by one round of local refinement with an overall mask in C3 symmetry to produce a map at 2.84Å resolution. Another round of local refinement in C3 symmetry focused on the transmembrane region was performed subsequently to improve its resolution, giving a map at 3.35Å resolution. These overall and locally refined maps showed clear density for transmembrane helices, but the resolution was not high enough for unambiguously modeling the structure. Data set II was therefore collected with the same batch of the S-nanodisc sample.

For data set II, the selected particles from 2D classification were used for one round of heterogeneous classification with six copies of the initial model as the reference in C1 symmetry. Two major classes with total 37.8% of the particles showing clear structural features were subjected to two additional rounds of heterogeneous refinement with six copies of the initial model as the reference in C1 symmetry. The major class was re-extracted to an unbinned box size of 600Å and subjected to one round of 2D classification, followed by one round of duplicates removing, non-uniform refinement in C1 symmetry, giving a map at 3.4Å resolution from 131,604 particles. These particles were combined with the particle set from data set I after the focus-classification, and subjected to one round of non-uniform refinement in C3 symmetry, CTF refinement, and another round of non-uniform refinement in C3 symmetry, giving a final overall map at 2.86Å resolution from the combined 254,902 particles. Particle-subtraction was also performed to reduce noises from the nanodiscs, followed by two rounds of local refinement with a different mask in size each time, giving a map at 3.26Å resolution in the transmembrane region with showing clear sidechain density for the transmembrane helices. The best maps from data set I and the combined data set were all used for model building.

All resolutions were reported from the gold-standard Fourier shell correlation (FSC) using the 0.143 criterion. Density maps were corrected from the modulation transfer function of the K3 detector and sharpened by applying a temperature factor that was estimated using sharpening tools in cryoSPARC. Local resolution was also determined using cryoSPARC.

Model building

The initial template for model building was our postfusion S structure (PDB ID: 6XRA). Several rounds of manual building were performed in Coot. The model around the transmembrane region was first refined in Phenix⁵⁷ against the locally refined maps at 3.26/3.35 Å resolution, and further refined in Phenix against the 2.9 Å overall map. Iteratively, refinement was performed in both Phenix (real space refinement) and ISOLDE⁵⁸, and the Phenix refinement strategy included `minimization_global`, `local_grid_search`, and `adp`, with `rotamer`, `Ramachandran`, and `reference-model` restraints, using 6XRA as the reference models. The local map from data set I had stronger density for the CT region and was used for modeling the CT; the map from the combined data set map had better sidechain density for the transmembrane helices and thus used for modeling them. The refinement statistics are summarized in Table S1. Structural biology applications used in this project were compiled and configured by SBGrid⁵⁹.

The intact S2 proteins from various coronaviruses are aligned by Clustal Omega⁶⁰ and displayed using the online server ESPript 3.0⁶¹.

Western blot

Full-length S protein samples were resolved in 4-15% Mini-Protean TGX gel (Bio-Rad) and transferred onto PVDF membranes (Millipore, Billerica, MA) by an Iblot2 (Invitrogen by Thermo Fisher Scientific). Membranes were blocked with 5% skimmed milk in PBS for 1 hour and incubated with a SARS-CoV-2 (2019-nCoV) Spike RBD Antibody (Sino Biological Inc., Beijing, China, Cat: 40592-T62) at the concentration of 1 µg/ml for another hour at room temperature. Alkaline phosphatase conjugated anti-Rabbit IgG (1:5000) (Sigma-Aldrich, St. Louis, MO) was used as a secondary antibody. Proteins were visualized using one-step NBT/BCIP substrates (Promega, Madison, WI).

Flow cytometry

Expi293F cells were grown in Expi293 expression medium. Cell surface display DNA constructs for the SARS-CoV-2 G614 or its mutants or S2 together with a plasmid expressing blue fluorescent protein (BFP) were transiently transfected into Expi293F cells using ExpiFectamine 293 reagent (ThermoFisher Scientific) per manufacturer's instruction. Two days after transfection, the cells were stained with primary antibodies at 5 µg/ml concentration. An Alexa Fluor 647 conjugated donkey anti-human IgG Fc F(ab')₂ fragment (Jackson ImmunoResearch, West Grove, PA) was used as secondary antibody at the concentration of 5 µg/ml. Cells were run through an Intellicyt iQue Screener Plus flow cytometer. Cells gated for positive BFP expression were analyzed for antibody binding.

Cell-cell fusion assay

The cell-cell fusion assay, based on the α -complementation of *E. coli* β -galactosidase, was conducted to quantify the fusion activity mediated by SARS-CoV-2 S protein, as described previously⁹. Briefly, the full-length G614 S or its mutants (10 µg) and the α fragment of *E. coli* β -galactosidase construct (10 µg), or the full-length ACE2 construct (5 µg) together with the ω fragment of *E. coli* β -galactosidase construct (10 µg), were transfected in HEK 293T cells using polyethylenimine (PEI) (80 µg). After incubation at 37°C for 5 hrs, the medium was aspirated and replaced with complete DMEM (1% Pen Strep, 1% GlutaMax and 10% FBS), followed by incubation at 37°C for additional 19 hrs. The cells were detached using PBS and resuspended in complete DMEM. 50 µl S-expressing cells (1.0×10^6 cells/ml) were mixed with 50 µl ACE2-expressing cells (1.0×10^6 cells/ml) to allow cell-cell fusion to proceed at 37°C for 4 hrs. Cell-cell fusion activity was quantified using a chemiluminescent assay system, Gal-Screen (Applied Biosystems, Foster City, CA), following the standard protocol recommended by the manufacturer. The substrate was added to the mixture of the cells and allowed to react for 90 min in dark at room temperature. The luminescence signal was recorded with a Synergy Neo plate reader (Biotek, Winooski, VT).

Methods-only References:

- 51 Zhang, J. *et al.* Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science* **372**, 525-530, doi:10.1126/science.abf2303 (2021).
- 52 Xiao, T. *et al.* A trimeric human angiotensin-converting enzyme 2 as an anti-SARS-CoV-2 agent. *Nat Struct Mol Biol* **28**, 202-209, doi:10.1038/s41594-020-00549-3 (2021).
- 53 Chen, J. *et al.* HIV-1 ENVELOPE. Effect of the cytoplasmic domain on antigenic characteristics of HIV-1 envelope glycoprotein. *Science* **349**, 191-195, doi:10.1126/science.aaa9804 (2015).
- 54 Ritchie, T. K. *et al.* Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol* **464**, 211-231, doi:10.1016/S0076-6879(09)64011-8 (2009).
- 55 Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006 (2012).
- 56 Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51, doi:10.1016/j.jsb.2005.07.007 (2005).
- 57 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S09074444909052925 (2010).
- 58 Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530, doi:10.1107/S2059798318002425 (2018).
- 59 Morin, A. *et al.* Collaboration gets the most out of software. *Elife* **2**, e01456, doi:10.7554/eLife.01456 (2013).
- 60 Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636-W641, doi:10.1093/nar/gkz268 (2019).
- 61 Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324, doi:10.1093/nar/gku316 (2014).

Acknowledgments: We thank the SBGrid team for computing support, computing resources from S. Harrison, and S. Harrison for critical reading of the manuscript. We acknowledge support for COVID-19 related structural biology research at Harvard from the Nancy Lurie Marks Family Foundation and the Massachusetts Consortium on Pathogen Readiness (MassCPR). This work was supported by COVID-19 Awards by MassCPR (to B.C.), Fast grant by Emergent Ventures (to B.C.), and NIH grants AI127193 (to B.C. and James Chou), AI147884 (to B.C.) and AI141002 (to B.C.).

Author Contribution: B.C., J.Z. and W.S. conceived the project. W.S. produced the full-length S protein and reconstituted in nanodiscs with help from Y.C. and H.P.. Y.C. also created the new MSP construct and established initial conditions for nanodisc reconstitution. W.S. and J.Z. prepared cryo grids and performed EM data collection with contributions from M.L.M., K.S. and C.X.. J.Z. processed the cryo-EM data, built and refined the atomic models. H.C. and J.L. created the mutant constructs. H.Z. performed the flow cytometry experiment. J.Z. and W.S. designed the mutant constructs, and carried out the cell-cell fusion assay. J.V. and S.R.V. contributed to plasmid preparation, cell culture and protein production. All authors analyzed the data. B.C., J.Z. and W.S. wrote the manuscript with input from all other authors.

Competing Interests: All authors declare no competing interests.

Data Availability: The atomic structure coordinates and EM map are deposited in the EMDataBank under the accession number: 8FDW, EMD-29016, EMD-29017 and EMD-29018. All other related data generated during and/or analyzed during the current study, such as raw cryo-EM images, are available from the corresponding authors on reasonable request. Source data are also available.

Extended Data Figure Legends and Table

Extended Data Figure 1. Schematic representations of the full-length SARS-CoV-2 and SARS-CoV spike proteins. Segments of S1 and S2 include: NTD, N-terminal domain; RBD, receptor-binding domain; CTD1, C-terminal domain 1; CTD2, C-terminal domain 2; S1/S2, the furin cleavage site at the S1/S2 boundary; S2', S2' cleavage site; FP, fusion peptide; FPPR, fusion peptide proximal region; HR1, heptad repeat 1; CH, central helix region; CD, connector domain; HR2, heptad repeat 2; TM, transmembrane segment; CT, cytoplasmic tail; and tree-like symbols for glycans. Based on studies of the SARS-CoV S protein, three membranotropic regions in the SARS-CoV S2 have been suggested as putative fusion peptides, including a segment of residues 770-788, equivalent of the N-terminal FP in SARS-CoV S; a segment of residues 798-816 or 798-835, widely accepted as the “bona fide” FP; and a segment of residues 858-886, also equivalent of the internal FP in SARS-CoV S. A segment of residues 1190-1203 is called the pretransmembrane domain (pre-TM).

Extended Data Figure 2. Additional details for preparation of the cryo-EM sample. The prefusion S trimer was first reconstituted in nanodiscs (a) and further purified by gel filtration chromatography (b). Incubation of the stable prefusion S trimer in nanodiscs with soluble ACE2 led to S1 dissociation, and formation of the postfusion S2 trimer in nanodiscs (c). Samples from each step were analyzed by SDS-PAGE and negative stain EM. The box size of 2D averages is ~880Å. The experiment was repeated at least three times independently at different scales with similar results. For gel source data, see Supplementary Figure 1.

Extended Data Figure 3. Cryo-EM analysis of the postfusion S2 trimer. Top, representative micrograph out of ~30,000 similar micrographs, and 2D averages (box dimension: 500Å) of the cryo-EM particle images of the postfusion S2 trimer in nanodiscs. Bottom, data processing workflow for structure determination.

Extended Data Figure 4. Analysis of the postfusion S2 trimer structures by cryoSPARC. Gold standard FSC curves of the three refined 3D reconstructions of the postfusion S2 trimer and the corresponding cryoSPARC output for particle distribution of each reconstruction.

Extended Data Figure 5. Additional analysis of the postfusion S2 trimer structure. (a) 3D reconstructions of the postfusion S2 trimer from the overall refinement and local refinement are colored according to local resolution estimated by ResMap with the FSC=0.143 criterion. (b) Representative density in gray surface representation from the EM maps of the postfusion S2.

Extended Data Figure 6. Fit of the postfusion structure into the density map of the postfusion spike on the virion surface. The postfusion structure was fit into the map of the viral postfusion spike reconstructed by cryo-ET (EMD-30428) at two different contour levels.

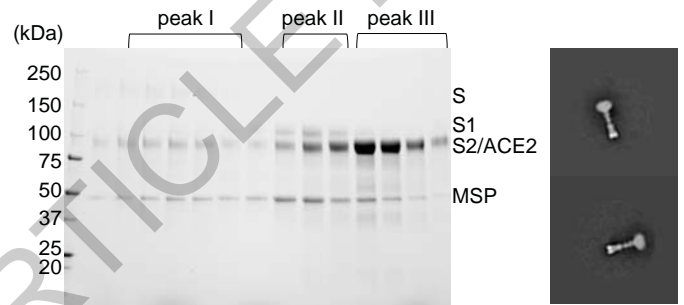
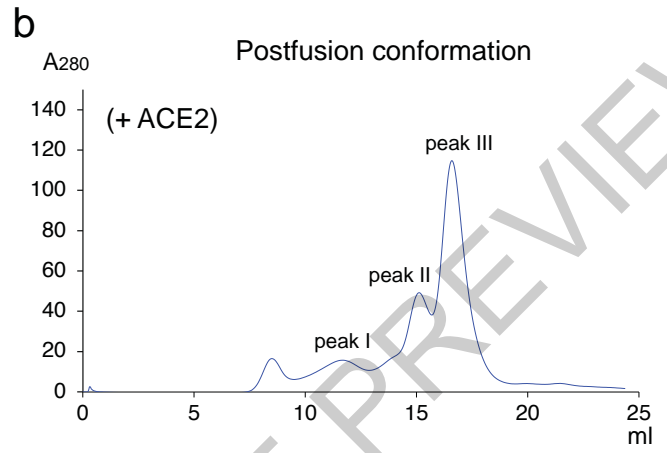
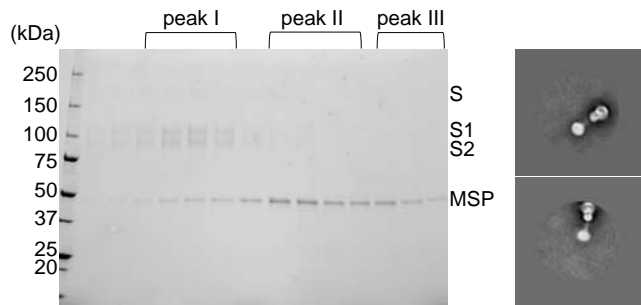
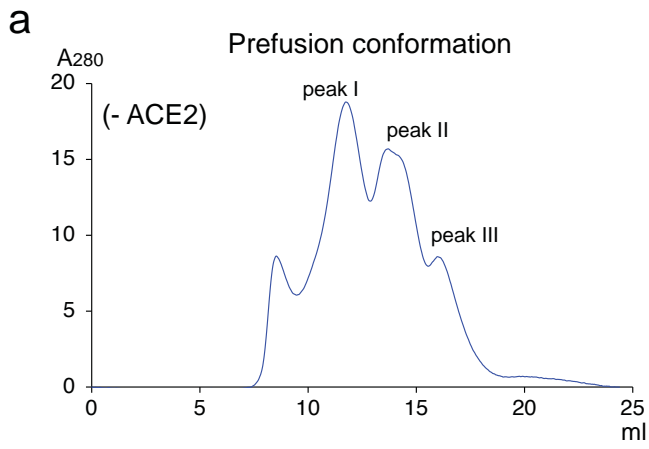
Extended Data Figure 7. Interactions among three FP protomers in the postfusion SARS-CoV-2 spike. Left, side view of packing of three FP protomer with each colored differently shown in both ribbon diagram and stick model. Residues mediating the formation of the three-stranded coiled coil near the three-fold axis are indicated. Right, top views of representative residues that mediate the formation of the three-stranded coiled coil.

Extended Data Figure 8. Expression and antigenic properties of the mutant spike proteins assessed by flow cytometry. (a) Expression and processing of the full-length G614 S construct and its mutants (Mut-1 – Mut-7) in HEK293 cells. S protein samples prepared from HEK293 cells transiently transfected with 10 µg of the full-length S expression plasmids were detected by anti-RBD polyclonal antibodies. Bands for the uncleaved S and S1 fragment are indicated. The experiment was repeated four times with similar results. (b) Antibody binding to the full-length G614 S protein and its mutants (Mut-1 – Mut-7), as well as an S2 construct expressed on the cell surfaces analyzed by flow cytometry. The antibodies and their targets are indicated. A designed ACE2-based fusion inhibitor ACE2₆₁₅-foldon-T27W was used for detecting receptor binding⁵². MFI, mean fluorescent intensity. The error bars represent standard errors of mean from measurements using three independently transfected cell samples. The flow cytometry assays were repeated three times with essentially identical results. (c) Expression and processing of the full-length G614 S construct and its additional mutants (Mut-8 – Mut-15) in HEK293 cells. The experiment was repeated three times with similar results. (d) Antibody binding to the full-length G614 S protein and its additional mutants (Mut-8 – Mut-15), as well as an S2 construct expressed on the cell surfaces analyzed by flow cytometry. The error bars represent standard errors of mean from measurements using three independently transfected cell samples. The flow cytometry assays

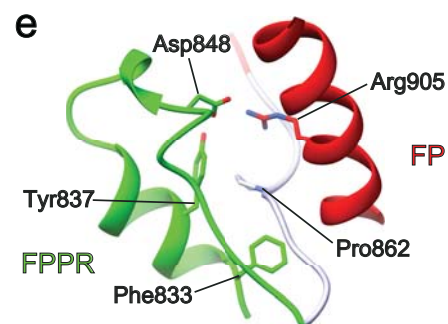
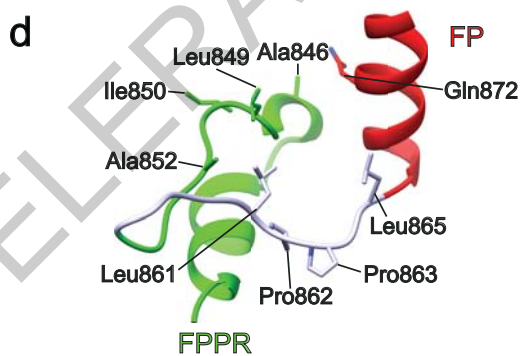
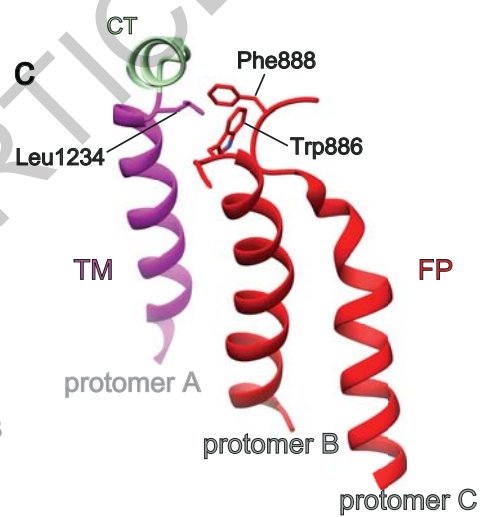
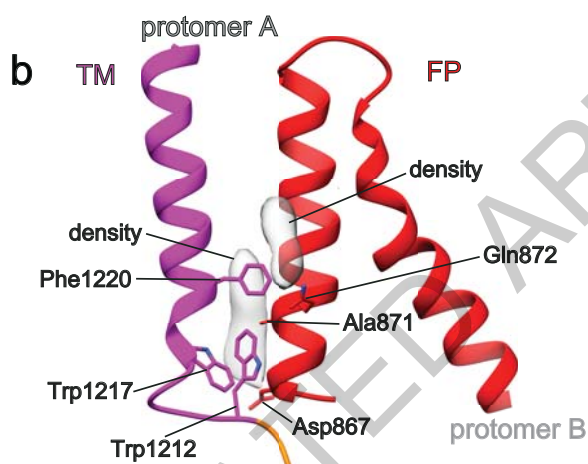
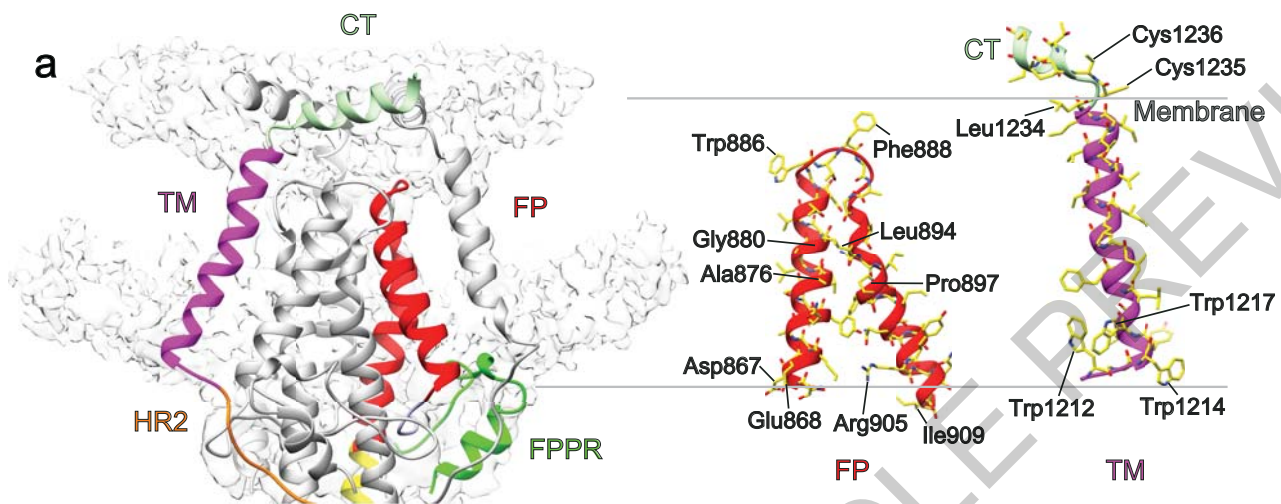
were repeated three times with essentially identical results. (e) Expression and processing of the full-length Delta S construct and its mutants in HEK293 cells. The experiment was repeated four times with similar results. (f) Antibody binding to the full-length Delta S protein and its mutants, as well as an S2 construct expressed on the cell surfaces analyzed by flow cytometry. The error bars represent standard errors of mean from measurements using three independently transfected cell samples. The flow cytometry assays were repeated three times with essentially identical results. For western blot source data, see Supplementary Figure 1.

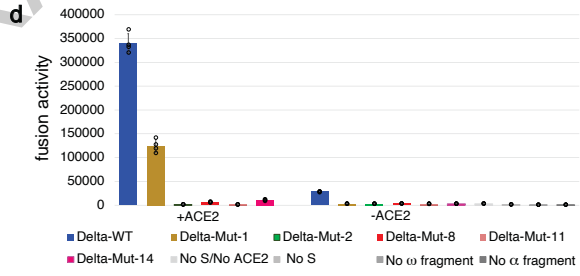
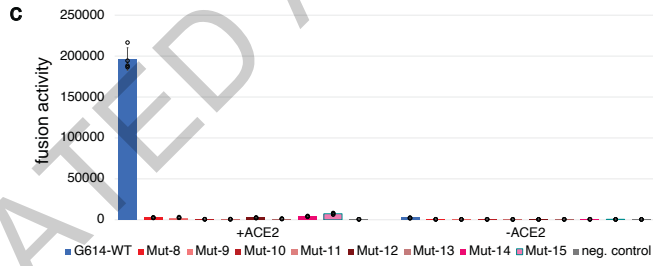
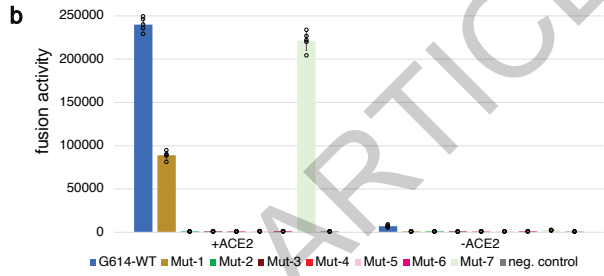
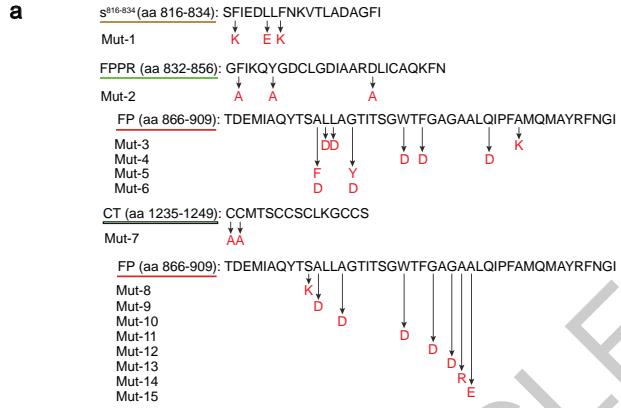
Extended Data Table 1. Cryo-EM data collection, refinement and validation statistics.

ACCELERATED ARTICLE PREVIEW

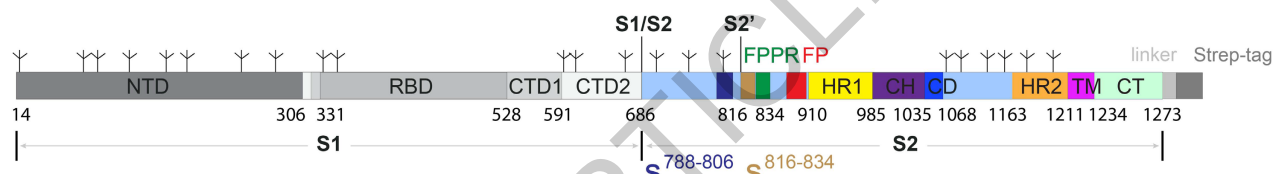


ACCELERATED ARTICLE PREVIEW

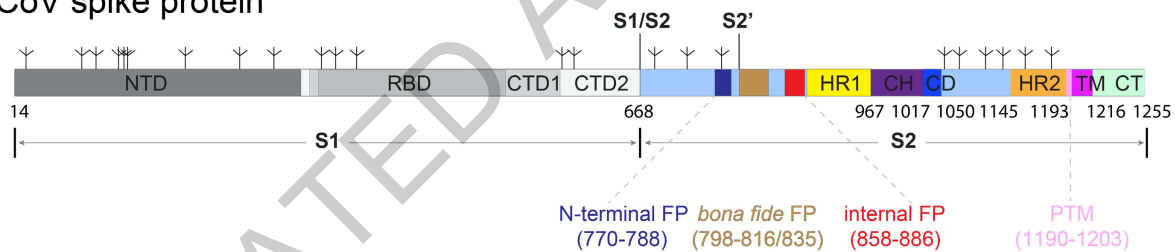




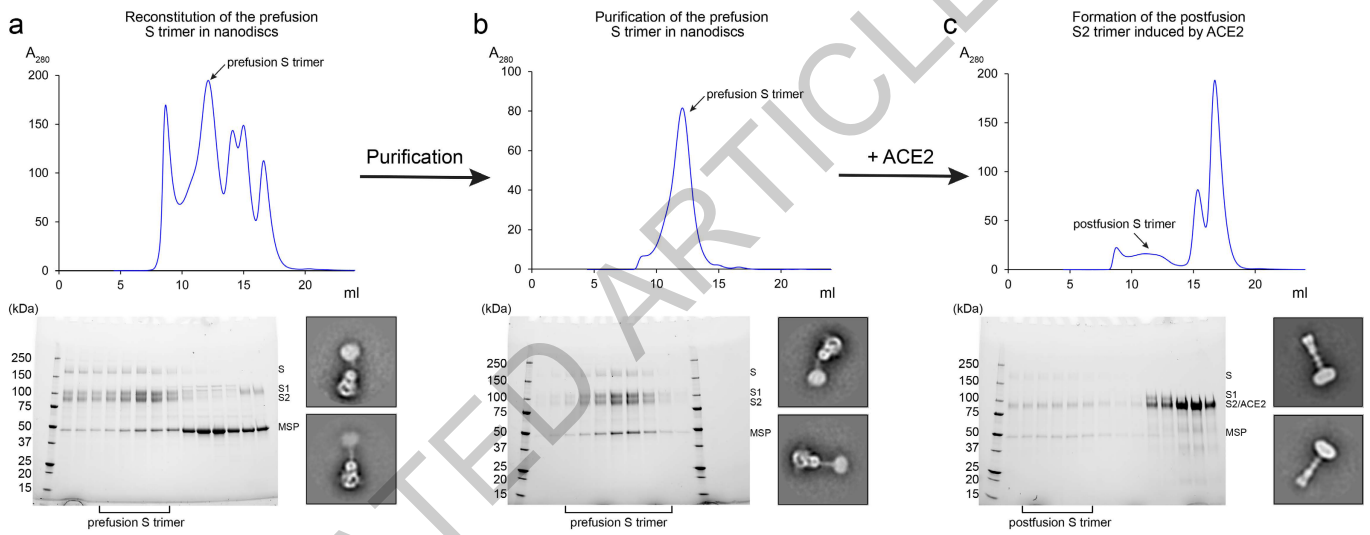
SARS-CoV-2 spike protein (expression construct)



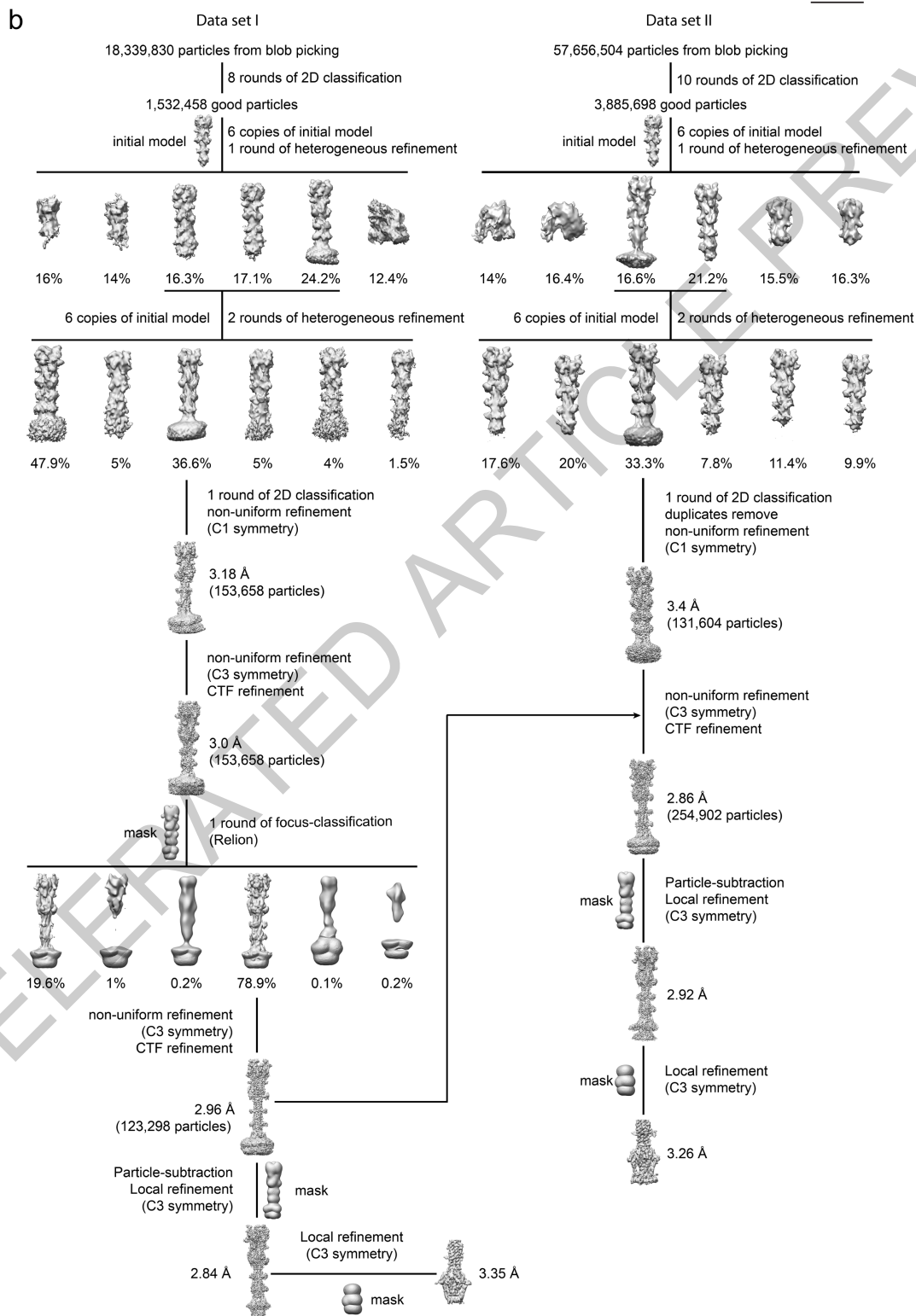
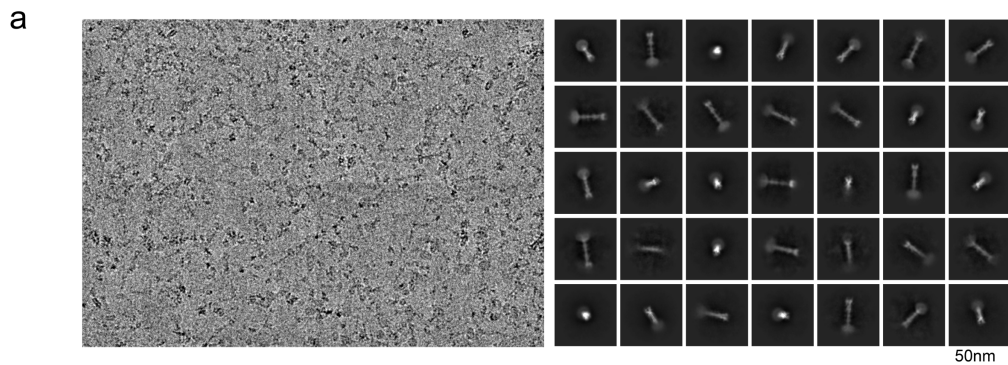
SARS-CoV spike protein



Extended Data Fig. 1

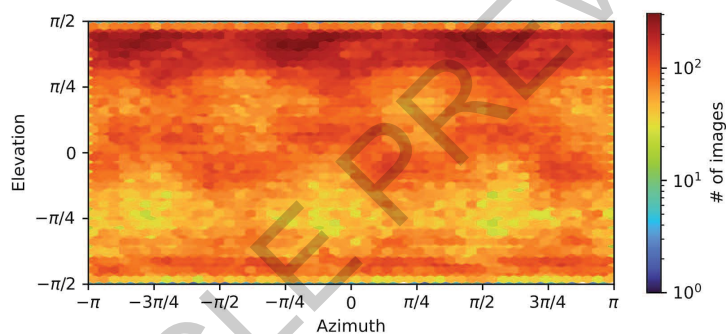
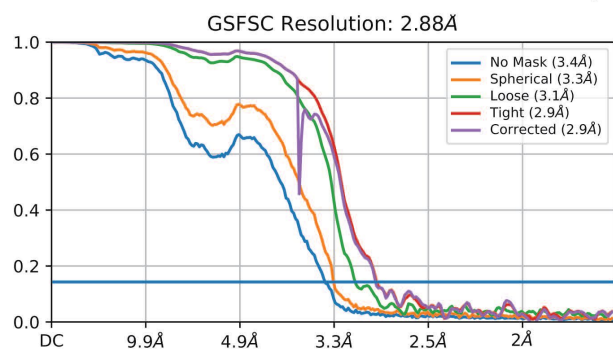


Extended Data Fig. 2

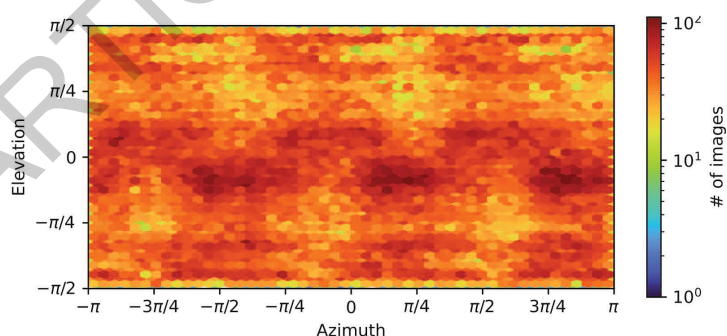
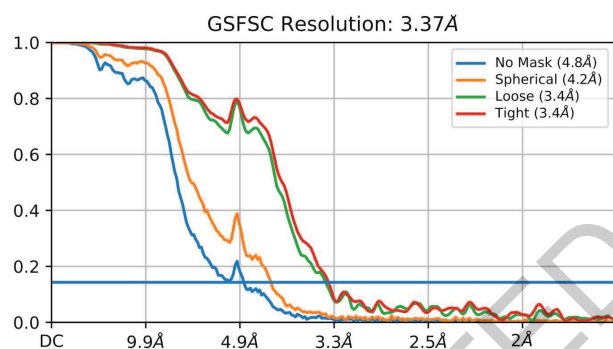


Extended Data Fig. 3

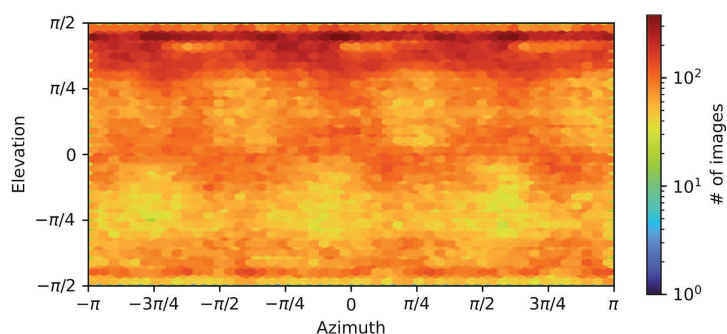
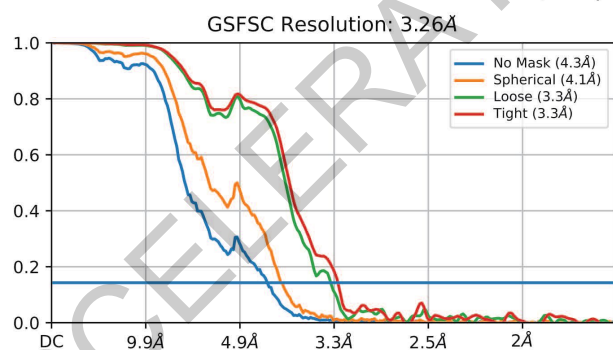
C3 symmetry, overall map from the combined data set
254,902 particles; GSFSC Resolution 2.88Å



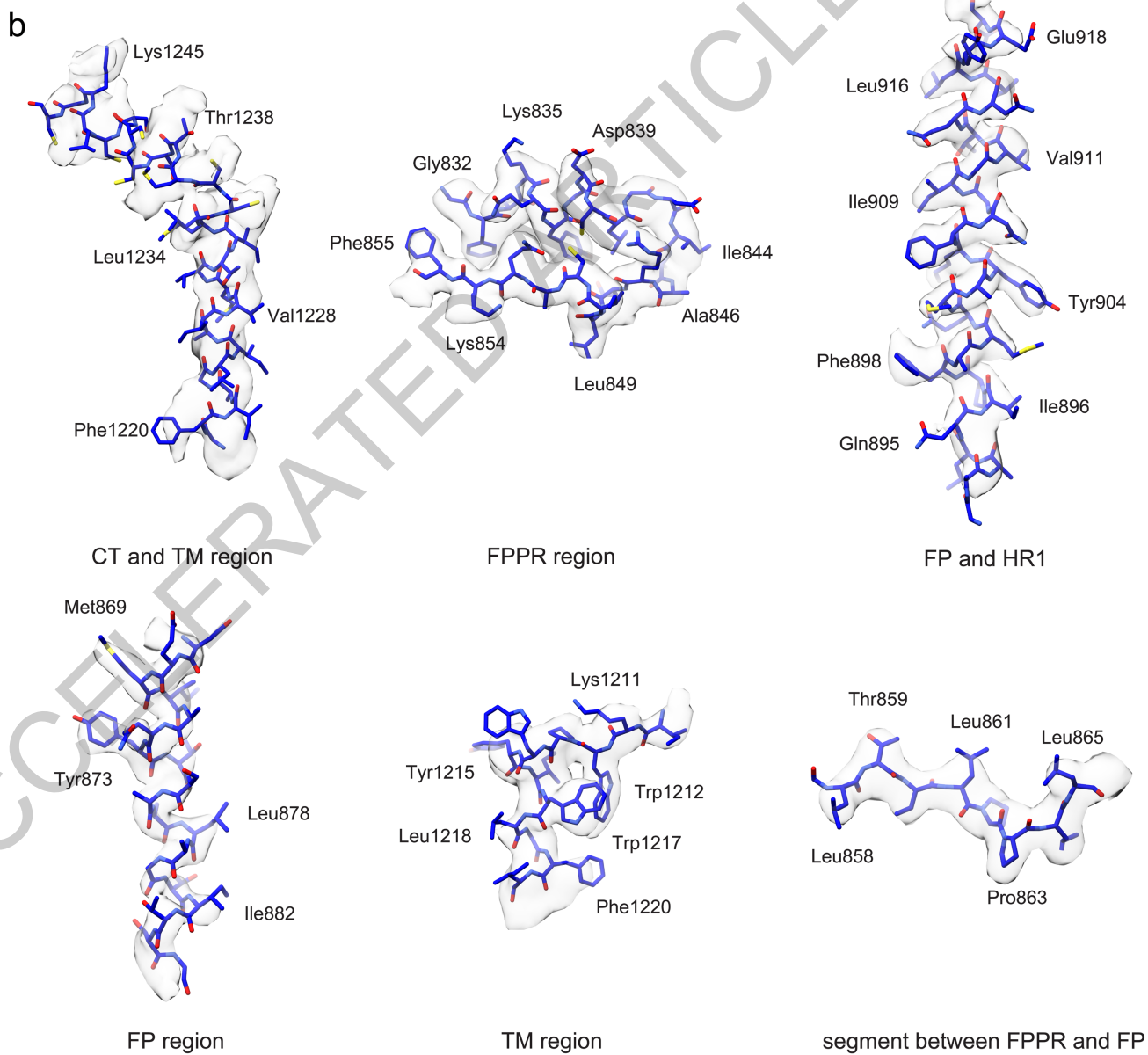
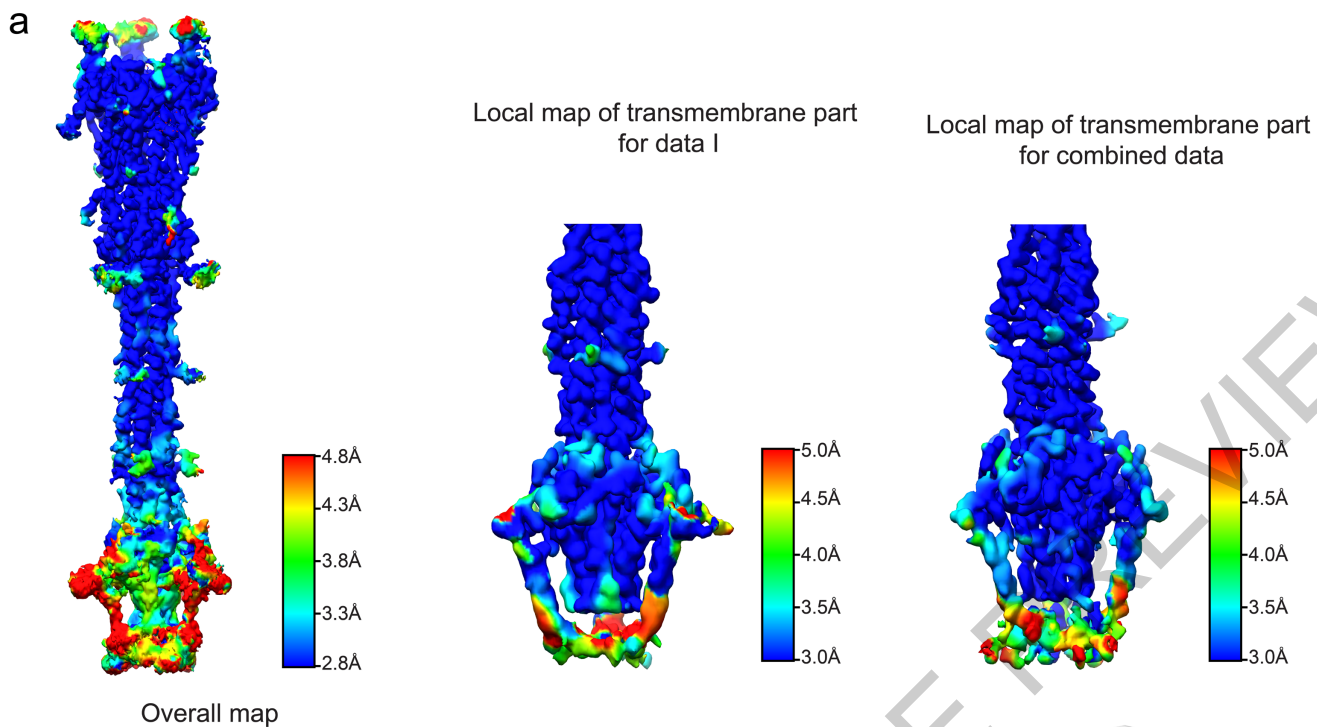
C3 symmetry, local map of the transmembrane region from data set I
123,298 particles; GSFSC Resolution 3.37Å



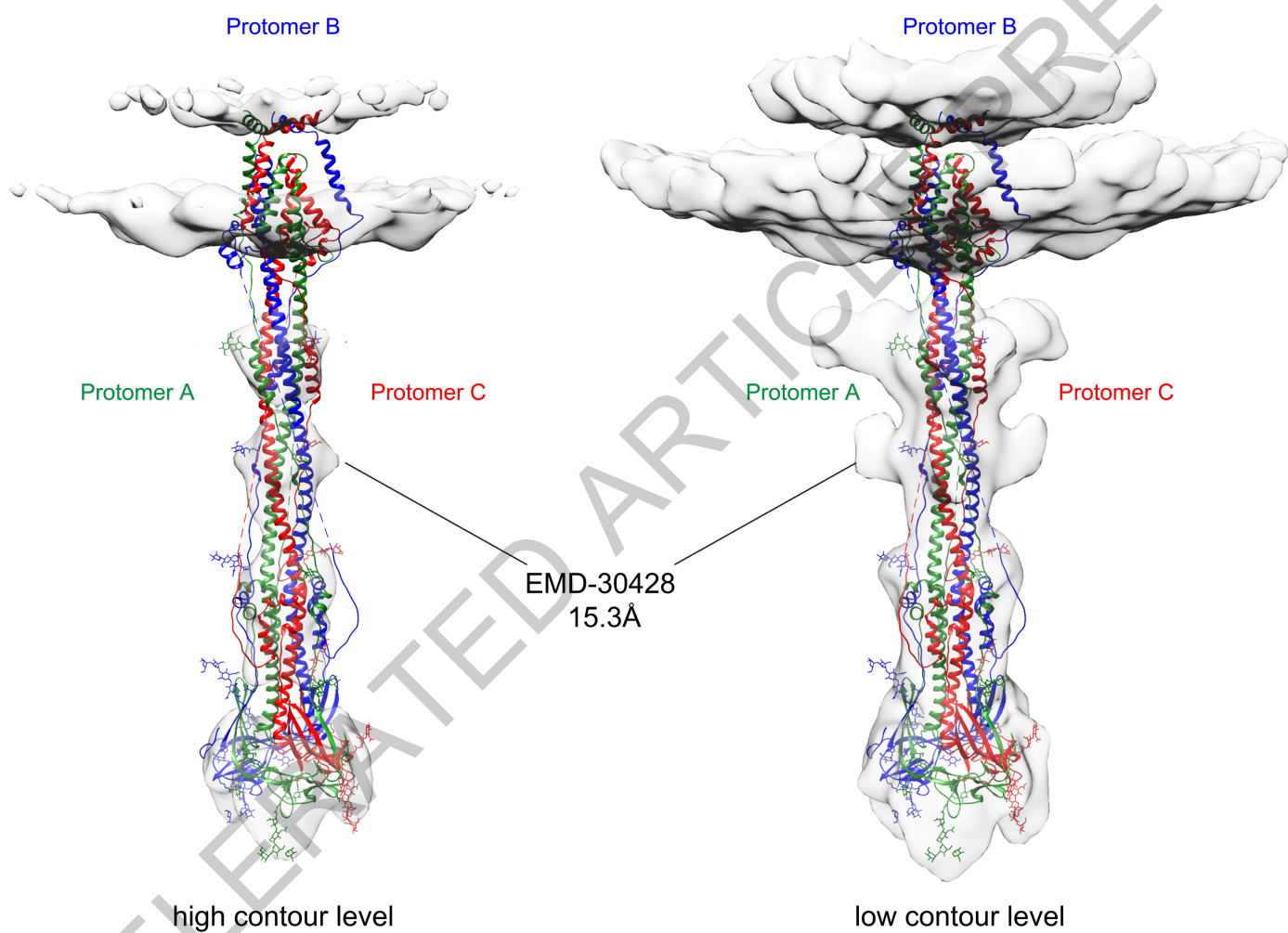
C3 symmetry, local map of the transmembrane region from the combined data set
131,604 particles; GSFSC Resolution 3.26Å



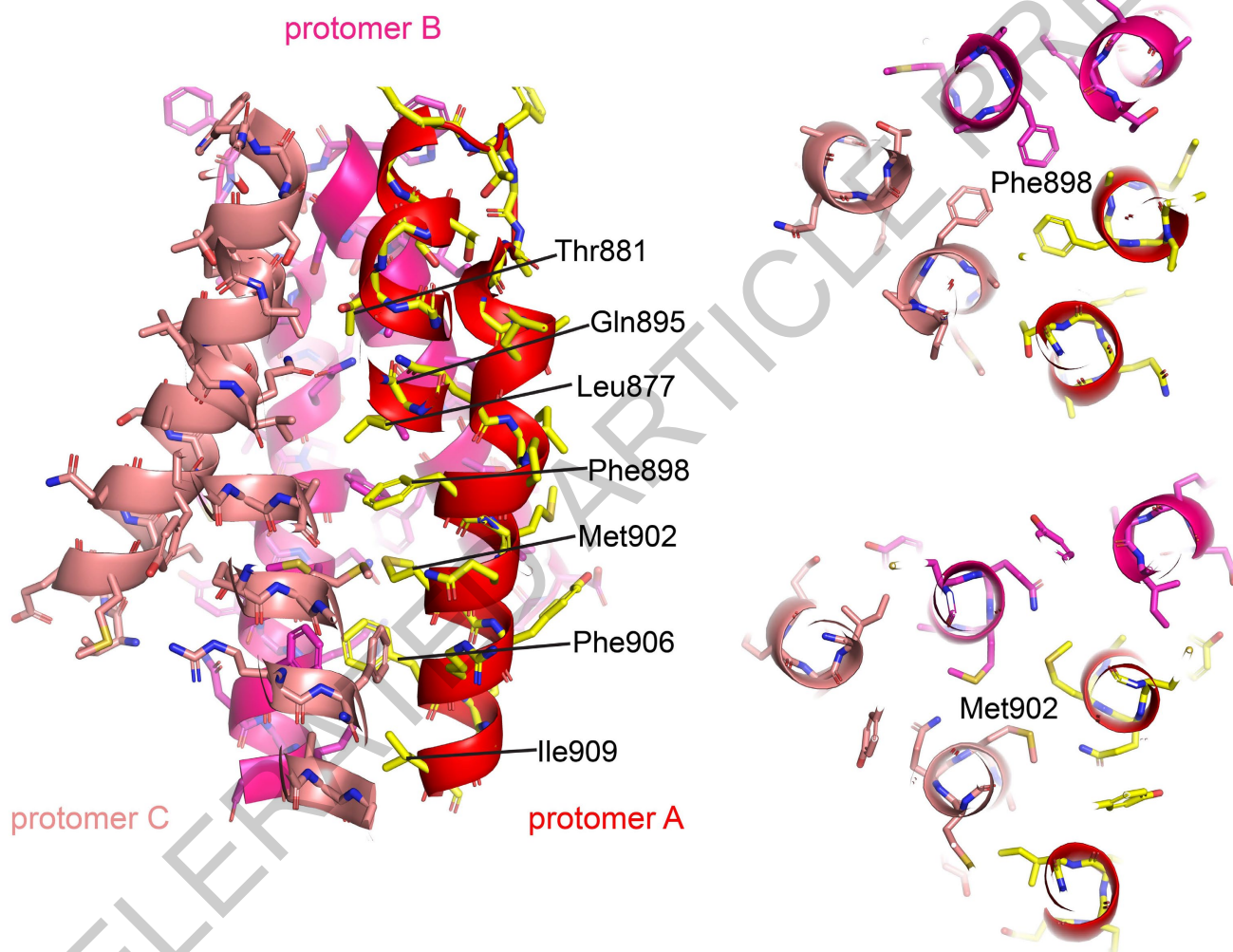
Extended Data Fig. 4



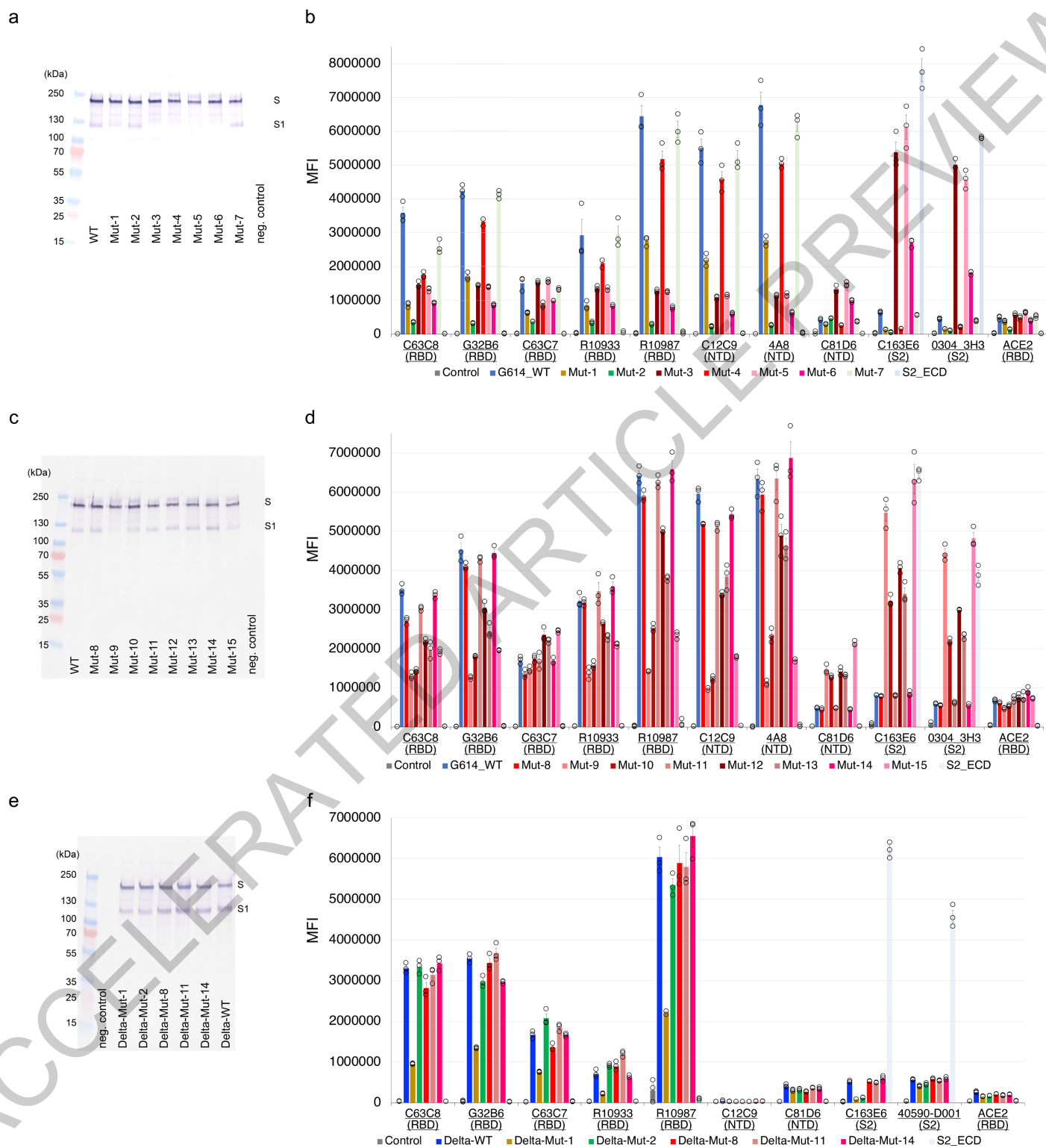
Extended Data Fig. 5



Extended Data Fig. 6



Extended Data Fig. 7



Extended Data Fig. 8

	SARS-CoV-2 postfusion spike protein in nanodisc (EMDB-29016) (PDB 8FDW)	Local map from data set I (EMDB- 29017)	Local map from the combined data sets (EMDB- 29018)
Data collection and processing			
Magnification	105,000	105,000	105,000
Voltage (kV)	300	300	300
Electron exposure (e-/Å ²)	50.6, 51.3	50.6	50.6,51.3
Defocus range (µm)	-0.8 to -2.2	-1.2 to -2.2	-0.8 to -2.2
Pixel size (Å)	0.825	0.825	0.825
Symmetry imposed	C3	C3	C3
Initial particle images (no.)	75,996,334	18,339,830	75,996,334
Final particle images (no.)	254,902	123,298	254,902
Map resolution (Å)	2.9	3.4	3.3
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	1.7-8.0	1.8-11.0	2.0-7.7
Refinement			
Initial model used (PDB code)	6XRA		
Model resolution (Å)	4.23		
FSC threshold	0.5		
Model resolution range (Å)	2.14-5.80		
Map sharpening <i>B</i> factor (Å ²)	-93.1	-112.3	-122.5
Model composition			
Non-hydrogen atoms	12456		
Protein residues	1524		
Ligands	24		
<i>B</i> factors (Å ²)			
Protein	75		
Ligand	99		
R.m.s. deviations			
Bond lengths (Å)	0.012		
Bond angles (°)	1.603		
Validation			
MolProbity score	2.23		
Clashscore	9.93		
Poor rotamers (%)	1.92		
Ramachandran plot			
Favored (%)	91.80		
Allowed (%)	8.07		
Disallowed (%)	0.13		

Extended Data Table 1

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

- | | |
|-----------------|--|
| Data collection | SerialEM version 3.8.6 |
| Data analysis | cryo-SPARC-v.3.3.2, Relion 4.0, Coot 0.9.8.7, Chimera 1.16, ChimeraX 1.5, Phenix 1.19.1, PyMOL Version 2.5.2, Clustal Omega BETA version, ESPript 3.0, iQue Forecyt® Enterprise Client Edition 9.0 (R3). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The atomic structure coordinates and EM map are deposited in the EMDDataBank under the accession number: 8FDW, EMD-29016, EMD-29017 and EMD-29018. All other related data generated during and/or analyzed during the current study, such as raw cryo-EM images, are available from the corresponding authors on reasonable request. Source data are also available.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	<input type="text" value="N/A"/>
Population characteristics	<input type="text" value="N/A"/>
Recruitment	<input type="text" value="N/A"/>
Ethics oversight	<input type="text" value="N/A"/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<input type="text" value="No statistical methods were used to predetermine sample size. All our biochemical and functional assays have been repeated 3-4 times independently. They have excellent reproducibility and additional repeats do not add any additional useful information."/>
Data exclusions	<input type="text" value="No data were excluded from analyses."/>
Replication	<input type="text" value="All experiments have been repeated 3-4 times independently with excellent reproducibility."/>
Randomization	<input type="text" value="Experimental groups are not needed for this work, therefore randomization is not relevant."/>
Blinding	<input type="text" value="The investigators were not blinded to group allocation during data collection and/or analysis because groups were not necessary for this work."/>

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	<input type="text" value="In-house antibodies include C63C8, G32B6, C63C7, R10933, R10987, C12C9, 4A8, C81D6, C163E6, 0304-3H3. Alexa Fluor 647 conjugated donkey anti-human IgG Fc F(ab')2 fragment (Jackson ImmunoResearch, West Grove, PA), cat# 709-606-098; Sars-Cov-2(2019-nCov) Spike RBD, Rabbit PAb, From Sino Biological, cat#40592-T62; and Anti-Rabbit whole IgG-Alkaline Phosphatase antibody from Sigma Cat#A3687; 40590-D001 from SinoBiological, Cat: 40590-D001."/>
Validation	<input type="text" value="The non-commercial antibodies used in the current manuscript have been characterized in previous studies (Tong et al., Cell. 2021 Jul 23:S0092-8674(21)00884-9; Zhang et al., Science. 2021 Apr 30;372(6541):525-530). Alexa Fluor 647 conjugated donkey anti-"/>

human IgG Fc F(ab')₂ fragment was validated by immunoelectrophoresis and/or ELISA (<https://www.jacksonimmuno.com/catalog/products/709-606-149/Donkey-Human-IgG-HL-Alexa-Fluor-647>); SARS-CoV-2 (2019-nCoV) Spike RBD Antibody was validated by western blot and ELISA (<https://www.sinobiological.com/antibodies/cov-spike-40592-t62>); Anti-Rabbit whole IgG-Alkaline Phosphatase antibody was validated by ELISA and western blot (<https://www.sigmaldrich.com/US/en/product/sigma/a3687>); 40590-D001 was validated by ELISA, FCM and ICC/IF (<https://www.sinobiological.com/antibodies/cov-spike-40590-d001>).

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	HEK 293 and HEK 293T cells were purchased from ATCC; Expi293F from Thermo Fisher Scientific.
Authentication	HEK 293, HEK293T and Expi293F cells used for protein production were authenticated for protein production by western blot and protein purification.
Mycoplasma contamination	Mycoplasma contamination is routinely tested for our cell culture and no contaminated cells were ever used for our studies.
Commonly misidentified lines (See ICLAC register)	None.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Expi293 cells were transiently transfected with SPIKE variants and BFP.
Instrument	Sartorius iQue Plus
Software	iQue Forecyt® Enterprise Client Edition 9.0 (R3)
Cell population abundance	Stained cells were evaluated in the flow cytometer, but not sorted. Approximately 2000 BFP+ cells from each sample were analyzed.
Gating strategy	Cell population was identified using FSC-H/SSC-H gate. Singlet cells were further defined using FSC-H/FSC-A gates. Positive transfected cells were identified using the VL-1 channel gating for the co-transfected BFP. Antibody bindings were determined by Alexa Fluor 647-labeled secondary antibodies using the RL-1 channel.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.