AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VACCINES AND ANTIVIRAL AGENTS

Cryo-electron Microscopy Structure of S-Trimer, a Subunit Vaccine Candidate for COVID-19

Jiahao Ma,[a,b] Danmei Su,[c] Yinyan Sun,[a] Xueqin Huang,[c] Ying Liang,[c] Linqiang Fang,[a] Yan Ma,[a,b] Wenhui Li,[a,b] Peng Liang,[c] Sanduo Zheng[a,b]

[a]National Institute of Biological Sciences, Beijing, China
[b]Tsinghua Institute of Multidisciplinary Biomedical Research, Tsinghua University, Beijing, China
[c]Clover Biopharmaceuticals, Chengdu, China

Jiahao Ma and Danmei Su contributed equally to this work. Author order was determined both alphabetically and in order of increasing seniority.

ABSTRACT Within a year after its emergence, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has infected over 100 million people worldwide, with a death toll over 2 million. Vaccination remains the best hope to ultimately put this pandemic to an end. In this study, using Trimer-Tag technology, we produced both wild-type (WT) and furin site mutant (MT) S-Trimers for COVID-19 vaccine studies. Cryo-electron microscopy (cryo-EM) structures of the WT and MT S-Trimers, determined at 3.2 Å and 2.6 Å, respectively, revealed that both antigens adopt a tightly closed conformation and their structures are essentially identical to that of the previously solved full-length WT S protein in detergent. The tightly closed conformation is stabilized by fatty acid and polysorbate 80 binding at the receptor binding domains (RBDs) and the N-terminal domains (NTDs), respectively. Additionally, we identified an important pH switch in the WT S-Trimer that shows dramatic conformational change and accounts for its increased stability at lower pH. These results validate Trimer-Tag as a platform technology in production of metastable WT S-Trimer as a candidate for COVID-19 subunit vaccine.

IMPORTANCE An effective vaccine against SARS-CoV-2 is critical to end the COVID-19 pandemic. In this study, using Trimer-Tag technology, we were able to produce stable and large quantities of WT S-Trimer, a subunit vaccine candidate for COVID-19 with high safety and efficacy from animal and phase 1 clinical trial studies. Cryo-EM structures of the S-Trimer subunit vaccine candidate show that it predominately adopts a tightly closed prefusion state and resembles the native and full-length spike in detergent, confirming its structural integrity. WT S-Trimer is currently being evaluated in a global phase 2/3 clinical trial. Taking into account the published structures of the S protein, we also propose a model to dissect the conformation change of the spike protein before receptor binding.

KEYWORDS COVID-19, cryo-EM, vaccine

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in late 2019 has led to a global pandemic and has disrupted lives and global economies on a scale unseen in recent human history. This is not the first time a new coronavirus has posed a major threat to public health; both SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV) have caused human infections within past 17 years (1). The fact that no licensed vaccines have ever been approved for these highly similar viruses is a reminder of the great challenges we face when hundreds of companies and institutions worldwide rush to develop COVID-19 vaccines with multiple strategies (2).

Citation Ma J, Su D, Sun Y, Huang X, Liang Y, Fang L, Ma Y, Li W, Liang P, Zheng S. 2021. Cryo-electron microscopy structure of S-Trimer, a subunit vaccine candidate for COVID-19. J Virol 95:e00194-21. https://doi.org/10.1128/JVI.00194-21.

Editor Tom Gallagher, Loyola University Chicago

Address correspondence to Peng Liang, liang.peng@cloverbiopharma.com, or Sanduo Zheng, zhengsanduo@nibs.ac.cn.

Received 4 February 2021
Accepted 4 March 2021
Accepted manuscript posted online 10 March 2021
Published 10 May 2021

A successful vaccine that could truly impact the course of the ongoing COVID-19 pandemic has to have four key characteristics: safety, efficacy, scalability (to billions of doses to meet global demand), and speed. Although protein subunit vaccines have excellent track records for the first three requirements, exemplified by the highly successful vaccine Gardasil used to prevent human papillomavirus (HPV) infections (3) and Shingrix vaccine for containing herpes zoster virus infections (4), subunit vaccine development can take years to decades to complete. Many of the difficulties reside in the manufacturing processes that have to ensure that a fully native-like antigen structure is retained, starting from subunit vaccine designs. Similar to other enveloped RNA viruses, such as HIV, respiratory syncytial virus (RSV), and influenza virus, coronaviruses, including SARS-CoV-2, use a ubiquitous trimeric viral surface antigen (spike protein) to gain entry into host cells. In the case of SARS-CoV-2, this occurs via binding to the ACE2 receptor expressed in target cells (5). Approaches for using a noncovalent trimer foldon from T4 bacteriophage fibritin to stabilize the trimer conformation while simultaneously introducing mutations in viral antigens to abolish furin cleavage and stabilize the antigen in prefusion forms (6, 7) are common strategies inherited from decades of HIV and RSV vaccine studies, but these have yet to yield a successful vaccine.

RESULTS

Construct design and negative staining. Using Trimer-Tag technology (8), we produced both soluble wild-type (WT) and a furin site mutant (MT) (R685A) forms of S-Trimer from CHO cells in serum-free fed-batch processes in bioreactors, with protein yields ranging from 0.5 to 1 g/liter. These antigen titers are 3 orders of magnitude higher than that reported previously for foldon-derived S proteins with two Pro mutations (S-2P) (6, 7), laying a solid foundation for the scalability requirement of a successful COVID-19 vaccine. Both S-Trimers consist of the ectodomain (amino acid residues 1 to 1211) of SARS-CoV-2 spike protein fused in-frame to the C-terminal region of human type I (α) collagen, which spontaneously form a disulfide-bond linked homotrimer, thereby stabilizing the antigens in trimeric forms (Fig.1A and B). Using a tailored affinity purification scheme that employs a collagen receptor Endo180-Fc fusion protein that binds to the Trimer-Tag with high affinity, the secreted S-Trimers were purified to near homogeneity in a single step (9). Reducing SDS-PAGE analysis of the purified S-Trimers revealed that the WT S-Trimer was metastable and partially cleaved precisely at S1/S2 boundary by furin, while a single point mutation (R685A) in MT S-Trimer fully abolished the protease cleavage (Fig.1*C*). WT spike proteins from live SARS-CoV-2 (10) or recombinant full-length S (11) were all previously shown to be partially cleaved apparently at the S1/S2 boundary by furin proteases. In contrast, an S-Trimer derived from wild-type SARS-CoV-1 S protein lacking a furin cleavage site produced in the same manner was essentially uncleaved by furin protease, like the MT S-Trimer from SARS-CoV-2 (Fig.1*C*). Receptor binding studies using ForteBio biolayer interferometry showed that all three S-Trimers had similarly high affinities for ACE2-Fc at 1.2 nM (Fig. 1D), similar to previous studies using purified S-2P protein (5). Negative-staining electron microscopy (EM) analysis of the MT S-Trimer revealed homogeneous particles consistent with trimeric spike proteins connected to the Trimer-Tag (Fig. 1E). Particles with similar structural features have also been observed by negative-staining EM for WT S-Trimer (9). Taken together, these results suggested that both MT and WT S-Trimer proteins produced using Trimer-Tag technology were well folded.

Cryo-EM structure of the MT S-Trimer. To further investigate structural details of these vaccine candidates, we first sought to determine the cryo-EM structure of the MT S-Trimer, which is more tractable for structural determination. While the particles showed preferred top-view orientation when embedded in ice, they preferentially adopted side view orientation on graphene oxide (GO)-coated grids (Fig. 2A). Combination of the two data sets enabled us to obtain a structure at a 2.6-Å resolution (Fig. 2B to E and Table 1). Due to the high resolution, the EM density for side chains was clear in the structure (Fig. 3). In addition to all the glycosylation sites and disulfide bonds previously observed in the S-2P structure (PDB code 6VXX), the disulfide bond

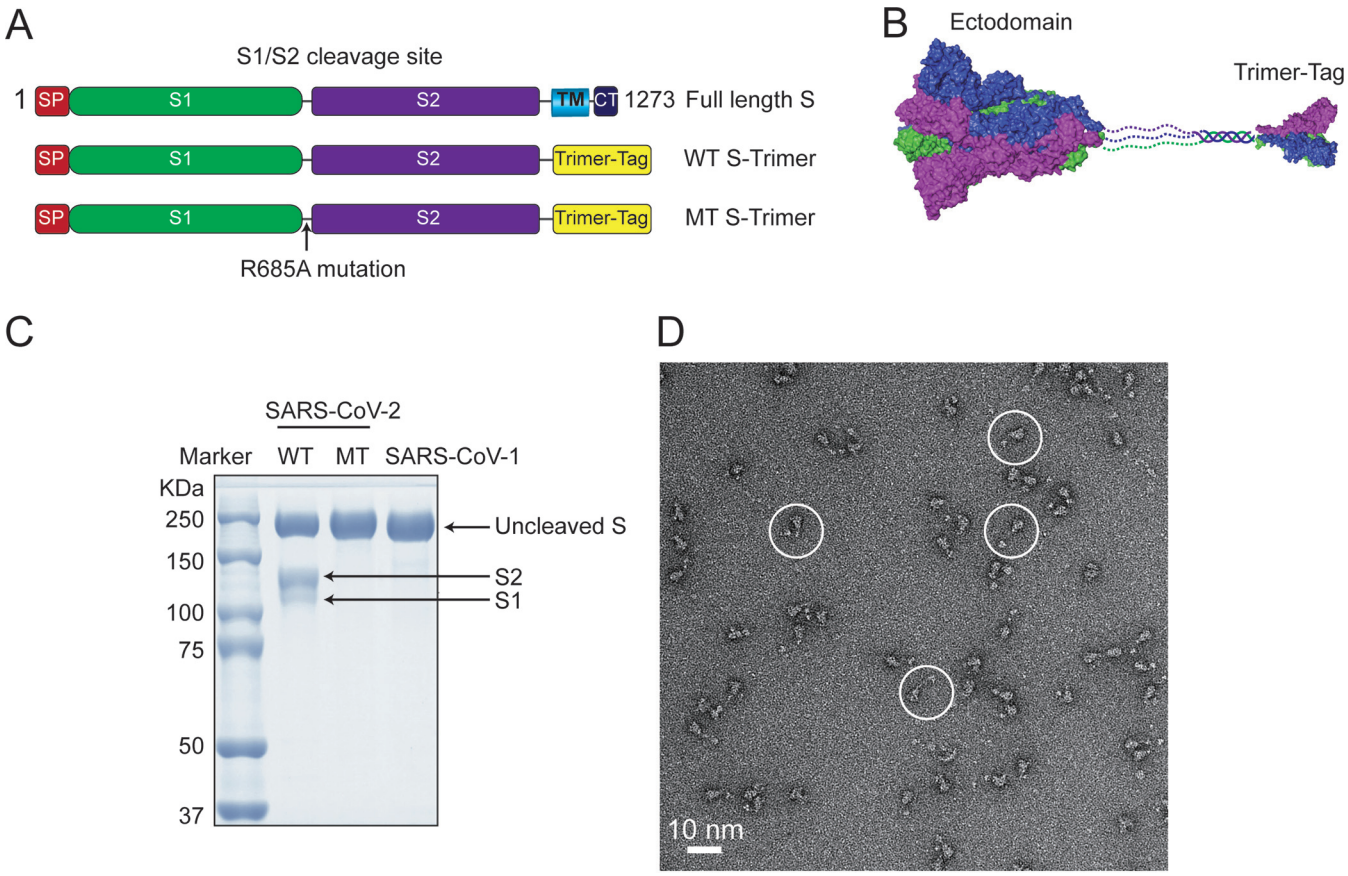

FIG 1 Structural design of the trimerized SARS-CoV-2 spike protein. (A) Schematic representation of the full-length spike protein, WT S-Trimer, and MT S-Trimer. The ectodomain of full-length S is fused with a trimer tag derived from the C-terminal domain of human type I(α) collagen to produce WT S-Trimer. A single point mutation R685A at the S1/S2 cleavage site was introduced in the WT S-Trimer to generate MT S-Trimer. The full-length S protein includes S1, S2, the S1/S2 cleavage site, transmembrane domain (TM), and cytoplasmic tail (CT). (B) Cartoon representation of the WT and MT S-Trimers. Structures of the ectodomain of the spike protein (PDB code 6VXX) and the C-terminal domain of collagen (PDB code 5K31) are shown as surface. Dotted lines represent regions not resolved in the structure. (C) The purified WT S-Trimer and MT S-Trimer of SARS-CoV-2 and wild-type S-Trimer of SARS-CoV-1 were analyzed by Coomassie-stained SDS-PAGE. (D) Receptor binding for WT and MT S-Trimers of SARS-CoV-2 and wild-type SARS-CoV-1 S-Trimer was analyzed as indicated. (E) A representative negative-stain EM image of the MT S-Trimer.

between Cys15 and Cys136 and the N-linked glycan at Asn17, both in the N-terminal domains (NTDs), were well resolved in our structure (Table 2). The Trimer-Tag was invisible in our structure due to highly flexible nature of the linker between the soluble S and the C-prodomain of collagen (Fig. 2B). After three-dimensional (3D) classification without imposing 3-fold symmetry (C3), we found that all three receptor binding domains (RBDs) in our map adopt a closed conformation (RBD down) without any open conformation (RBD up) that was previously observed in the S-2P structure (Fig. 2B) (6). Surprisingly, we observed unaccounted-for EM density in both the NTD and RBD of the S1 domain. The presence of polysorbate 80 (PS80) during the purification process suggested that the bulky EM density in the NTD region could be accounted for by PS80 (Fig. 3A and B). Indeed, PS80 can be well fitted into the density of the NTD. Mass spectrometry analysis further confirmed the presence of PS80 in the MT sample (Fig. 4). PS80 is buried deeply in the hydrophobic pocket residues, with a few hydrophilic residues, including N99, N121, R190, and H207, making hydrogen bonds with the hydroxyl group of PS80 (Fig. 5A and B). Notably, PS80 engages hydrophobic interactions with F175 and M177, which are invisible in the S-2P structure (Fig. 5C). Since no small molecule was shown to bind to the S-2P structure, PS80 likely stabilizes the disordered loops of the NTD, making them more ordered (Fig. 5B and C). The EM density in the RBD is elongated and in close proximity to R408 (Fig. 3C); we speculate that it may be oleic acid or linoleic acid, which is present in the culture medium. Although both

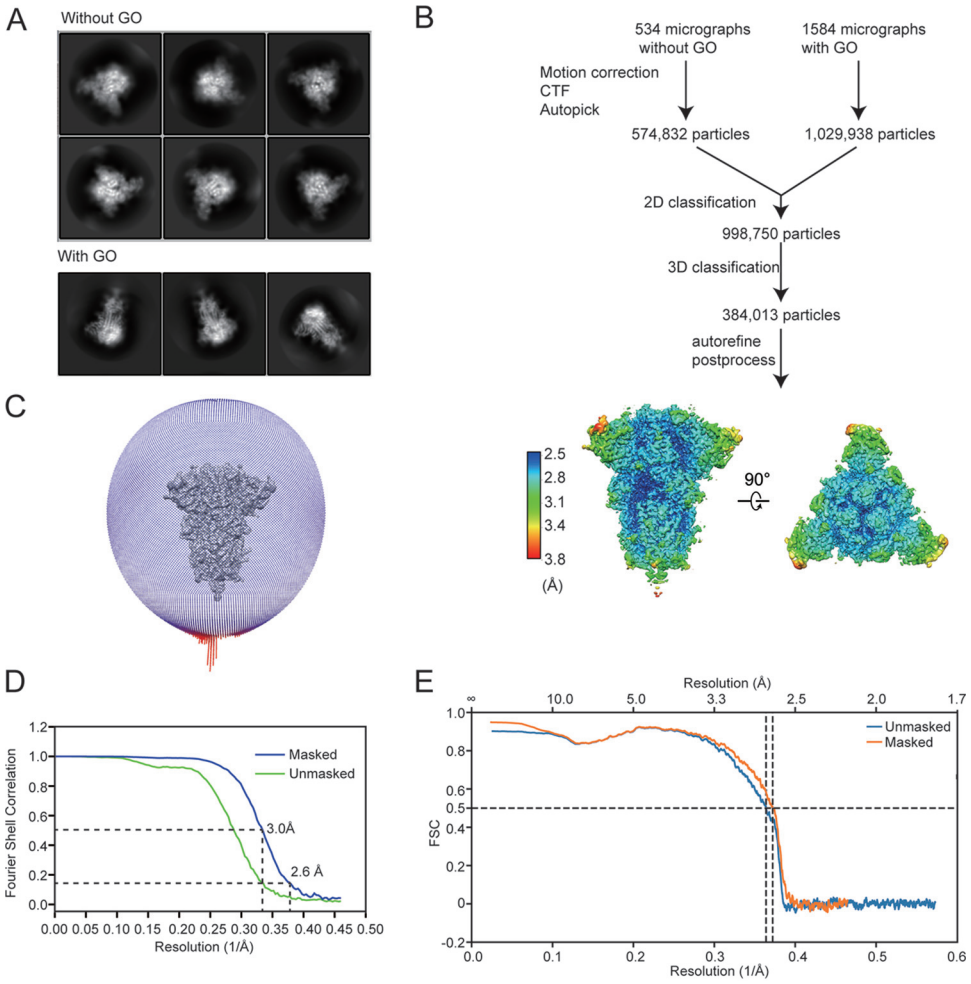

FIG 2 Cryo-EM data processing and analysis for MT S-Trimer. (A) Representative 2D class averages of images collected from grids with and without graphene oxide (GO). (B) Cryo-EM data processing workflow. (C) Angular distribution of the particles used for 3D reconstruction. (D) Gold standard FSC curves for the MT S-Trimer. Resolutions at FSC values of 0.5 and 0.143 are indicated on the masked FSC curve. (E) FSC curves of model to map.

oleic acid and linoleic acid can be ambiguously fitted into the density of the RBD owing to the excellent quality of EM map, mass spectrometry analysis indicated both MT and WT samples contained a higher level of oleic acid than linoleic acid (Fig 4B and C). Recent studies revealed the presence of linoleic acid at the same binding site but with a slightly different position (7, 12). The oleic acid in our structure, located in the hydrophobic pocket of the RBD, engaged a salt bridge interaction with R408 and a hydrogen bond interaction with Q409 at the adjacent protomer through its carboxylic acid group, bringing the RBD domain in close proximity and resulting in the tightly closed conformation (Fig. 5D).

Cryo-EM structure of the WT S-Trimer. Recent studies have shown that low pH can stabilize the S-Trimer (13). Indeed, negative-staining EM analysis of WT S-Trimer at pH 5.5 revealed more homogenous trimer than that at physiological pH (9). In light of this finding, we were able to determine the cryo-EM structure of the WT S-Trimer at a 3.2-Å resolution at pH 5.5 (Fig. 6 and Table 1). The structure of the WT S-Trimer resembled that of the MT form, with a root mean square deviation (RMSD) of 0.5 Å over 2,773 Cα atoms (Fig. 7A). Oleic acid was well resolved in the WT structure but the density for the PS80 was weak, likely due to the low resolution or low occupancy. As a result, the NTD of the WT S-Trimer was less well resolved than that of MT (Fig. 7A). It

TABLE 1 Cryo-EM data collection and refinement statistics

Parameter	Datum for indicated protein	
	Covid-19 WT S-trimer	Covid-19 MT S-trimer
EM data collection statistics		
EMDB no.	EMD-30999	EMD-30998
Microscope	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300
Detector	Gatan K3	Gatan K3
Magnification (nominal)	64,000	64,000
Pixel size (Å/pix)	1.087	1.087
Flux (e−/pix/s)	22	22
Frames per exposure	32	32
Exposure (e−/Å^2)	50	50
Defocus range (μm)	0.8–3	0.6–2.8
Micrographs collected	1,199	2,114
Particles extracted/final	752,204/240,705	998,750/384,013
Symmetry imposed	C3	C3
Map sharpening B-factor	−126.5	−80.9
Unmasked resolution at 0.5/0.143 FSC (Å)	4.2/3.6	3.5/3.0
Masked resolution at 0.5/0.143 FSC (Å)	3.6/3.2	3.0/2.6
Model refinement and statistics		
PDB	7E7B	7E7D
Composition		
Amino acids	3,240	3,306
Glycans	87	96
Ligand	3	6
Water	0	33
RMSD bonds (Å)	0.008	0.006
RMSD angles (°)	0.913	0.665
Mean B-factors		
Amino acids	64.62	45.70
Glycans	102.42	73.38
Water	0	31.17
Ramachandran		
Favored (%)	96.81	97.44
Allowed (%)	3.19	2.56
Outliers (%)	0.00	0.00
Rotamer outliers (%)	2.24	2.92
Clash score	7.53	4.92
C-beta outliers (%)	0.00	0.10
CC (mask)	0.84	0.87
MolProbity score	1.87	1.72
EMRinger score	3.08	4.47

has been shown that a pH-dependent switch domain (residues 824 to 848; pH switch 1) undergoes dramatic conformational changes at different pH values (13). However, this region was nearly identical between our two structures (Fig. 7B). The fatty acid binding in our WT S-Trimer structure may stabilize the conformation of pH switch 1, accounting for conformational difference from the previous study (13). In contrast, a fragment (residues 617 to 639) we named pH switch 2 at the CTD1 region of the S1 domain before the furin cleavage site displays significant structural arrangement. Whereas this region appeared disordered in the MT structure at physiological pH conditions, it was well ordered and formed a helix-turn-helix structural motif in the WT structure at pH 5.5 (Fig. 7C). From a structural perspective, lower pH likely contributes to this structural arrangement. At physiological pH, R319 forms salt bridge interactions with D737 and D745 (Fig. 7D). At lower pH, the protonation of D737 and D745 weakens these interactions. As a result, R319 flips to the other side and makes hydrophobic interactions with W633 and L629 through its aliphatic chain, leading to the ordered

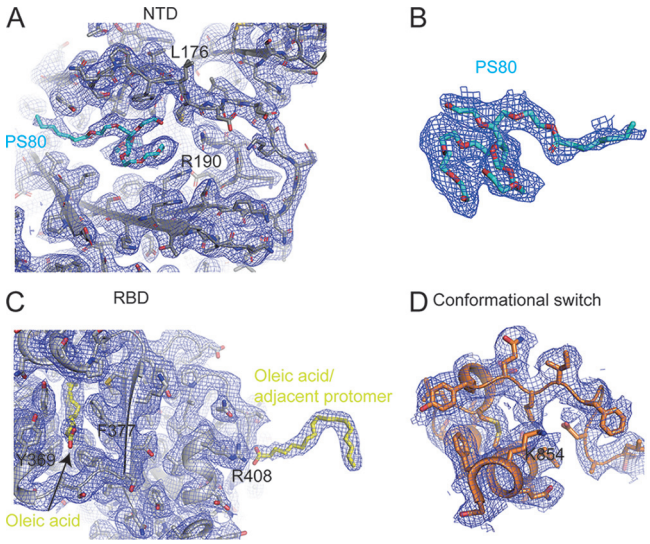

FIG 3 Representative B-factor sharpened EM density map for the NTD (A), polysorbate 80 (B), the RBD with oleic acid (C), and the conformational switch (D).

helix-turn-helix motif (Fig. 7E). The newly formed structural motif makes direct contact with the previously identified pH switch 1 of the adjacent protomer, accounting for the enhanced stability of the WT S-Trimer at lower pH. The structural arrangement of pH switch 2 at different pHs was also observed in previous studies (13), further supporting the notion that the conformational change between the MT and WT S-Trimer structures was due to the different pH but not to the mutation in the furin site. The biological relevance of the structural arrangement of pH switch 2 may require further investigation.

Conformational change of the S protein before receptor binding. In contrast to the structural differences described above for S-2P protein, both of our WT and MT S-Trimer proteins were nearly identical to the recently published structures of full-length wild-type S (PDB code 6XR8) (11) and 3Q-2P-FL with two proline mutations (PDB code 7JJI) (7, 14) purified in detergent from HEK293 and sf9 insect cell membranes, respectively. When revisiting the electron density map for full-length wild-type S protein (Electron Microscopy Data Bank [EMDB] no. 22292), we spotted unassigned density at the same position as oleic acid, which was absent in the S-2P map (EMDB no. 21452). Therefore, fatty acid binding at the RBD stabilizes the tightly closed conformation, accounting for the conformational difference from S-2P. Based on the structural similarity among all published structures of SARS-CoV-2 S protein, we can classify them into three distinct states: a tightly closed state, a loosely closed state, and an open state (Fig. 8A). A conformational switch (residues 835 to 858) previously known as the pH switch or the fusion peptide-proximal region (FPPR) is critical for conformational transition from the tightly closed state to the open state (Fig. 8B). In the tightly closed state, which is less accessible to the receptor, three RBDs are down and in closer prox-

TABLE 2 Summary of the glycosylation sites and disulfide bond observed in the structure

Feature	Sites or bonds
Glycosylation sites (Asn linked)	17, 61, 122, 149, 165, 234, 282, 331, 343, 603, 616, 657, 709, 717, 801, 1074, 1098, 1134
Disulfide bonds	Cys15-Cys136, Cys131-Cys166, Cys291-Cys301, Cys379-Cys432, Cys336-Cys361, Cys391-Cys525, Cys480-Cys488, Cys538-Cys590, Cys617-Cys649, Cys662-Cys671, Cys743-Cys749, Cys738-Cys760, Cys840-Cys851, Cys1032-Cys1043, Cys1082-Cys1126

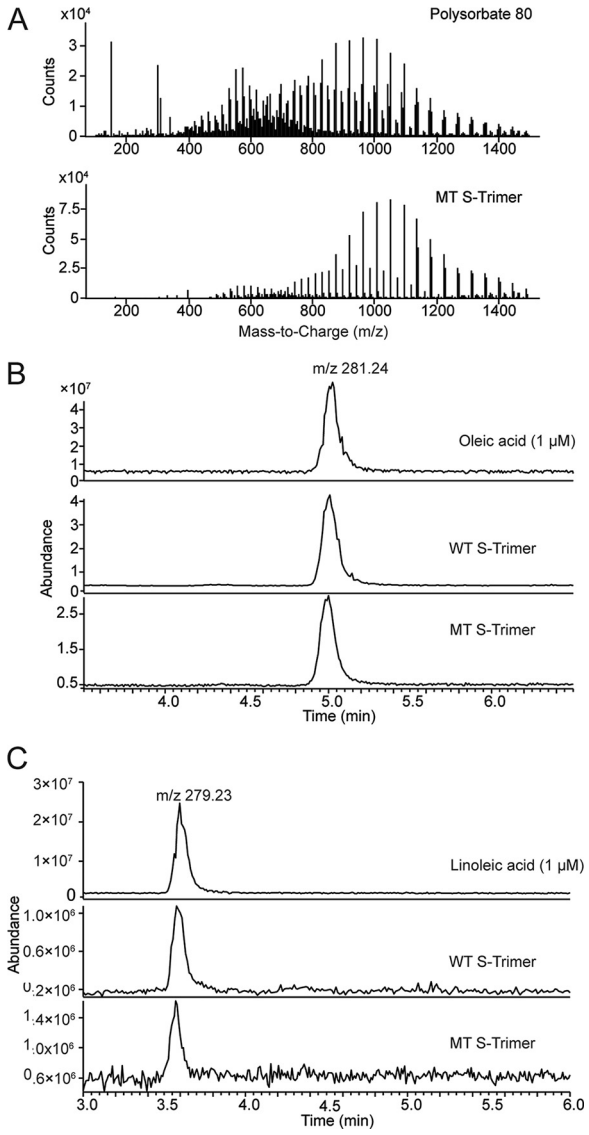

FIG 4 Small-molecule identification of the purified S-Trimer sample by mass spectrometry. (A) Mass spectra of polysorbate 80 and small molecules extracted from the MT S-Trimer sample revealed that they are similar. (B) Extracted ion chromatograms (EICs) for small molecules from both the WT and MT S-Trimers with 1 μM oleic acid (B) and linoleic acid (C) as references showed that S-Trimers were enriched with oleic acid.

imity to one another. The conformational switch region is well ordered and stabilizes this tightly closed state. Recently, the D614G mutation has become predominant over the ancestral form worldwide and has been shown to increase viral infections (15, 16). In the tightly closed state, D614 makes a salt bridge interaction with K854 at the conformational switch region (Fig. 8C). From the tightly to the loosely closed state, the conformational switch undergoes a large conformational arrangement and becomes disordered (Fig. 8D). K854 flips to the opposite side and interacts with D568 and D574 of the CTD1, causing the S1 to move downwards relative to the S2 (Fig. 8A). Finally, the CTD1 domain further moves downwards and causes the RBD to adopt an open conformation for receptor binding (see Movie S1 in the supplemental material). Therefore, the D614G mutation abolishes the salt bridge interaction with K854, which, in turn, engages the CTD1 domain, thus favoring the open state. Consistent with this model, SARS-CoV-2 pseudovirus (PV) infection experiments showed that the K854G mutation

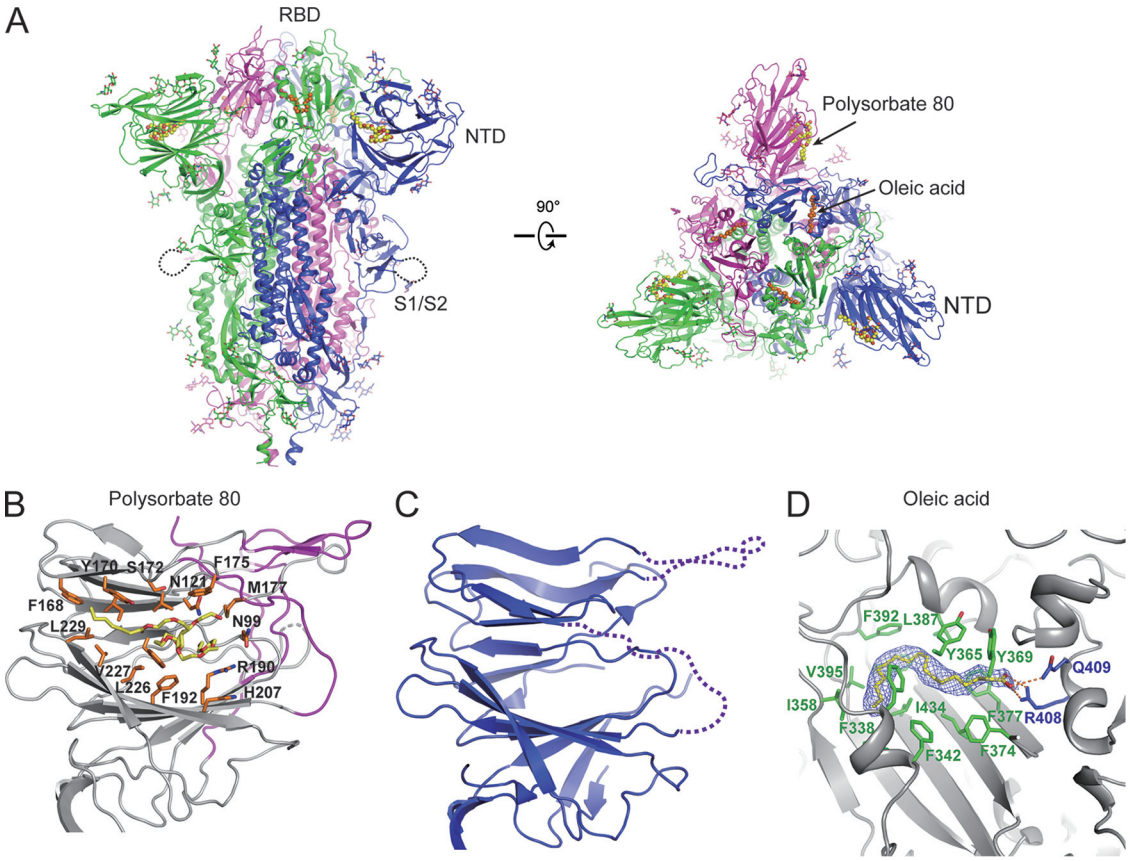

FIG 5 Cryo-EM structure of the MT S-Trimer. (A) Ribbon representation of the MT structure colored by subunit from two orthogonal views. Oleic acid and PS80 are shown as spheres and colored in orange and yellow, respectively. (B) Detailed view of the NTD bound to PS80. Structure colored in magenta corresponds to the flexible region at the NTD of the S-2P structure. (C) Structure of the NTD of the S-2P (PDB code 6VXX). The dotted line represents a disordered region. (D) Oleic acid bridges the two adjacent RBD domains. Oleic acid is colored in yellow, and the side chains of the two adjacent RBDs are colored in green and blue, respectively. The EM density for oleic acid is shown.

significantly reduced the virus infectivity in both the wild-type and D614G mutant backgrounds (Fig. 8E).

DISCUSSION

To our knowledge, this is the first cryo-EM structure of the wild-type S protein in soluble and cleavable form without the transmembrane domain, confirming the structural integrity of this metastable wild-type form of COVID-19 subunit vaccine candidate. Whereas most vaccine candidates currently in clinical trials incorporated furin site and double proline mutations, our S-Trimer vaccine candidate used wild-type sequence without any mutagenesis, providing an important alternative option. TrimerTag technology has proven in this study to be able to rapidly produce large quantities of native-like S-Trimer antigen and may offer a platform technology for subunit vaccine development for enveloped RNA viruses that use ubiquitous trimeric antigens to invade host cells.

Like for the previously reported structure of full-length WT S protein purified in detergent micelles, it is unclear whether the furin cleavage site in our resolved WT S-Trimer structure is cleaved. Moreover, we could not exclude the possibility that other conformational states exist in the WT sample that were not captured in our cryo-EM study, since partial cleavage of the furin site may lead to some S1 dissociation from S-Trimer. Nevertheless, we are certain that the highly purified WT S-Trimer

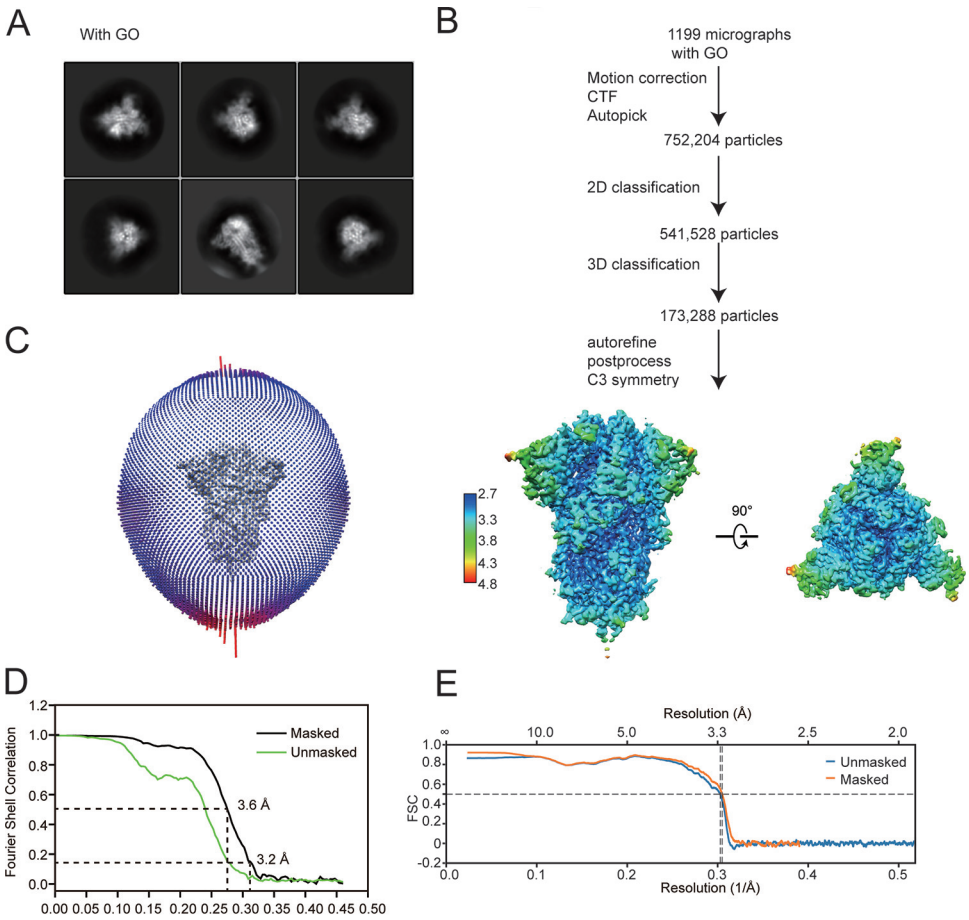

FIG 6 Cryo-EM data processing and analysis for the WT S-Trimer. (A) Representative 2D class averages of images collected from grids with GO. (B) Cryo-EM data processing workflow. (C) Angular distribution of the particles used for 3D reconstruction. (D) Gold standard FSC curves for the WT S-Trimer. (E) FSC curves of model to map.

predominately adopts a prefusion state, unlike the full-length wild-type spike protein, which forms both pre- and postfusion states in the presence of detergent (11).

Similar to the subunit vaccine candidate with two proline mutations from Novavax, WT S-Trimer largely adopts tightly closed conformation, which is stabilized by fatty acid binding (14). The biological relevance of the fatty acid binding awaits further investigation. One may argue that the tightly closed conformation of WT S-Trimer prevents the production of antibodies recognizing receptor binding region at the RBD, since they are buried in the closed conformation. The binding affinity of the spike protein in open conformation for the receptor is similar to that in tightly closed state, suggesting energy barrier from tightly closed state to open state is relatively low (5). Therefore, it is likely that WT S-Trimer adopts multiple conformational states when injected into the human body. The phase 3 clinical results for the vaccine candidate from Novavax with fatty acid binding and closed conformation showed promising efficacy, indicating that fatty acid binding has little influence on production of neutralizing antibodies.

Preclinical studies of the WT S-Trimer vaccine candidate revealed that it resulted in rapid and high-level induction of both neutralizing antibodies and Th1-biased cellular immune responses in mice, rats, and nonhuman primates. Nonhuman primates immunized with WT S-Trimer were fully protected from SARS-CoV-2 viral challenge (9). The phase 1 clinical trial studies also showed great immunogenicity and safety (17). Global phase 2/3 clinical trials are currently ongoing.

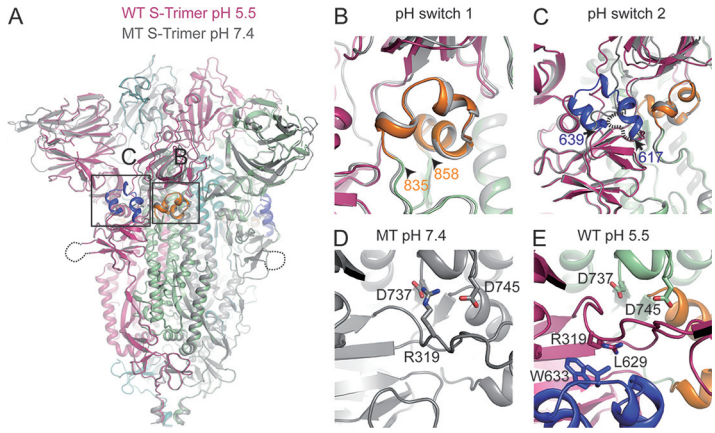

FIG 7 Cryo-EM structure of the WT S-Trimer at pH 5.5. (A) Structural overlay of the MT and WT S-Trimer. MT is colored in gray, and WT is colored by subunits. Two pH switches in the WT S-Trimer structure are boxed and colored in blue and orange, respectively. (B and C) Comparison of pH switch 1 (B) and pH switch 2 (C) between MT and WT S-Trimers. The dashed line represents disorder region in pH switch 2 of the MT structure. (D) pH switch 2 is flexible in the MT structure, where R319 makes electrostatic interactions with D737 and D745. (E) At low pH, switch 2 forms a helix-turn-helix motif in the WT structure in which R319 flips and stabilizes it.

MATERIALS AND METHODS

Protein expression and purification. Endo180-Fc expression vector was generated by subcloning a PCR-amplified cDNA encoding soluble human Endo180 (amino acid residues 1 to 1394) into the HindIII site of pGH-hFc expression vector (GenHunter, Nashville, TN) to allow in-frame fusion to human IgG Fc. The expression vector was transfected into the GH-CHO (*dhfr*$^{-/-}$) cell line (GenHunter, Nashville, TN) using FuGENE 6 (Roche) and grown in Iscove's modified Dulbecco's medium (IMDM) with 10% fetal bovine serum (FBS). After stepwise gene amplification with increasing concentrations (0.0 to 10 nM) of methotrexate (MTX) (Sigma), a high-titer clone was then adapted to SFM-4CHO serum-free medium (GE BioSciences). The secreted End0180-Fc fusion protein was then produced in a 15-liter bioreactor (Applikon) under a fed-batch process with Cell Boost 2 supplement (GE HyClone) as instructed by the manufacturer. Endo180-Fc was purified to homogeneity by protein A affinity chromatography using MabSelect PrismA (GE Healthcare) followed by Capto QXP resins (GE BioSciences) in a flowthrough mode to remove any host cell DNA and residual host cell proteins (HCP).

To produce S-Trimer fusion proteins, cDNA encoding the ectodomain of either wild-type SARS-CoV-2 S protein (amino acid residues 1 to 1211) (GenBank accession number MN908947.3) or a SARS-CoV (amino acid residues 1 to 1193) (GenBank accession number AAS00003.1) were synthesized using -optimized codons for *Cricetulus griseus* (CHO cells) by GenScript. The cDNAs were subcloned into pTRIMER expression vector (GenHunter, Nashville, TN) at HindIII and BglII sites to allow in-frame fusion of the soluble S protein to Trimer-Tag [residues 1156 to 1406 from human type I(α) collagen] as described previously (8). Furin cleavage site mutant S-Trimer (R685A) was generated by site-directed mutagenesis using the QuikChange kit (Stratagene). The expression vectors were transfected into the GH-CHO (*dhfr*$^{-/-}$) cell line (GenHunter) using FuGENE 6 (Roche) and grown in IMDM with 10% FBS. After stepwise gene amplification with increasing concentrations (0.0 to 10 nM) of MTX (Sigma), clones producing the highest S-Trimer titer were adapted to SFM-4CHO serum-free medium (GE BioSciences). The secreted S-Trimer fusion proteins were produced in a 15-liter bioreactor (Applikon) under a fed-batch process with Cell Boost 2 supplement (GE HyClone) as instructed by the manufacturer.

Cell culture medium was clarified by depth filtration (Millipore) to remove cell and debris. S-Trimers were purified to homogeneity by consecutive chromatographic steps, including a protein A affinity column using MabSelect PrismA (GE Healthcare) which was preloaded with Endo180-Fc at 3 mg/ml to capture S-Trimer, based on the high-affinity binding between Endo180 and Trimer-Tag (18). After unbound contaminating proteins were washed off, S-Trimers were purified to near homogeneity in a single step using 0.5 M NaCl in phosphate-buffered saline (PBS). For S MT and SARS-CoV S-Trimer, the proteins were dialyzed against PBS plus 0.02% polysorbate 80 before analysis. After 1 h of low-pH (pH 3.5) viral inactivation (VI) using acetic acid, the pH was adjusted to neutral range, and WT S-Trimer was further purified on a Capto QXP resin (GE BioSciences) in a flowthrough mode to remove any host cell DNA and residual HCP. A final preventative viral removal (VR) step was performed using a nanofiltration cartridge (AsahiKASEI) before final buffer exchange to PBS plus 0.02% polysorbate 80 by ultrafiltration/diafiltration (UF/DF) (Millipore). For EM sample preparation, the protein was further purified by size exclusion chromatography on an Agilent 1260 Infinity high-performance liquid chromatograph (HPLC) with a TSK gel G3000 SWxL column (Tosoh) in PBS.

ACE2-Fc expression vector was generated by subcloning a gene-synthesized cDNA template (GenScript) encoding soluble human ACE2 (amino acid residues 1 to 738; GenBank accession number NM_001371415.1) into HindIII and BglII sites of pGH-hFc expression vector (GenHunter, Nashville, TN) to

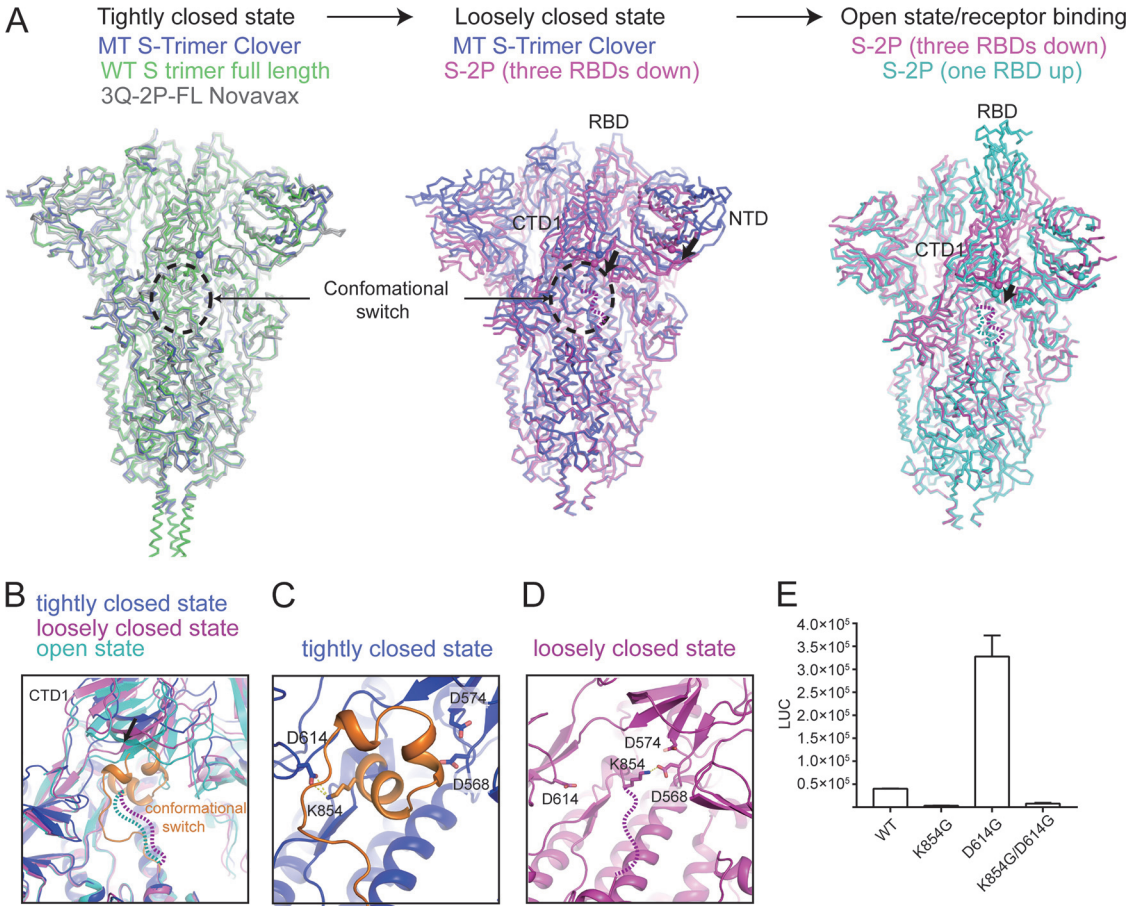

FIG 8 Conformational change of the SARS-CoV-2 spike protein before engaging the ACE2 receptor. (A) Conformation transition from the tightly closed conformation WT/MT structure/WT S trimer full length (PDB code 6XR8)/3Q-2P-FL from Novavax (PDB code 7JJI) to the loosely closed conformation (PDB code 6VXX) and, finally, to the open conformation (PDB code 6VYB). The direction of movement of the NTD and CTD1 of the S1 domain during conformational change is marked with an arrow. The conformation switch region boxed is disorder in both the loosely closed and open conformation and is represented as dashed line. (B) A close-up view of the conformational switch in three conformations. The ordered switch would clash with CTD1 domain in the loosely closed and open conformation (C) K854 in the switch region engages D614 in the tightly closed conformation. (D) In the loosely closed conformation, the switch region undergoes structural arrangement, in which K854 flips and pulls the S1 domain downwards through interacting with D574 and D568 in the CTD1 region. (E) The K854G mutation reduced the virus infectivity. TMPRSS2-293T cells expressing hACE2 on a 96-well plate were infected with lentivirus PV expressing luciferase and pseudotyped with the indicated SARS-CoV-2 spike glycoprotein. Luciferase activity (LUC) was measured at 48 h postinfection.

allow in-frame fusion to human IgG Fc. The expression vector was then stably transfected into the GH-CHO (*dhfr*$^{-/-}$) cell line, high-expression clones were selected and adapted to SFM-4-CHO (HyClone) serum-free medium, and ACE2-Fc was produced in a 15-liter bioreactor essentially as described for S-Trimer above. ACE2-Fc was purified to homogeneity from the conditioned medium using a PoRos XQ column (Thermo Fisher) following the manufacturer's instructions.

Studies of binding of S-Trimers to human receptor ACE2-Fc. The avidity of binding of different S-Trimers to the SARS-CoV-2 receptor ACE2 was assessed by biolayer interferometry measurements on a ForteBio (Pall, NY) Octet QKe platform. ACE2-Fc (10 μg/ml) was immobilized on protein A (ProA) biosensors (Pall). Real-time receptor binding curves were obtained by applying the sensor in 2-fold serial dilutions of S-Trimer from 22.5 to 36 μg/ml in PBS. Kinetic parameters (K_{on} and K_{off}) and affinities (dissociation constants [K_D]) were analyzed using Octet software, version 12.0. Dissociation constants were determined using steady-state analysis, assuming a 1:1 binding model for S-Trimer to ACE2-Fc.

Negative staining. Negative staining was performed as previously described (19). In brief, 3 μl of purified S-trimer at a concentration of about 0.01 mg/ml was deposited on a glow-discharged carbon-coated copper grid for 30 s before being blotted with filter paper. Grids were quickly washed with 2 drops of water and 1 drop of 2% (wt/vol) uranium acetate. Grids were kept touching to the last drop of 2% (wt/vol) uranium acetate for 100 s and blotted with filter paper. Data collection was performed on a Talos L120C electron microscope operated at 120 kV equipped with an FEI Ceta 4K detector. Images were collected at a magnification of $\times$57,000 and a defocus of 1.5 μm.

Cryo-EM sample preparation and data collection. Purified MT S-trimer proteins diluted to 0.2 and 0.5 mg/ml in PBS were applied to glow-discharged gold holey carbon 1.2/1.3 300-mesh grids with

and without graphene oxide (GO), respectively. Grids were blotted for 2 to 4 s at a blotting force of 4 and plunge-frozen in liquid ethane using a MarkIV Vitrobot (Thermo Fisher Scientific). The chamber was maintained at 8°C and 100% humidity during freezing. A sample (0.3 mg/ml) of WT S-trimer in low-pH buffer (100 mM sodium citrate [pH 5.5] and 100 mM NaCl) was deposited on glow-discharged gold holey carbon 1.2/1.3 300-mesh grids with graphene oxide. Grids were blotted and vitrified under the same condition.

All movies were collected using a Titan Krios microscope (Thermo Fisher Scientific) equipped with a BioQuantum GIF/K3 direct electron detector (Gatan). The detector was operated in superresolution mode. A complete description of cryo-EM data collection parameters is summarized in Table 1.

Cryo-EM image processing. For MT S-Trimer protein, motion correction for cryo-EM images and contrast transfer function (CTF) estimation were performed using motioncorr2 (20) and CTFFIND4 (21), respectively. A total of 574,832 particles were automatically picked from 534 images collected on grids without GO using Laplacian-of-Gaussian in Relion 3.0.7 (22), and 1,029,938 particles were automatically picked from 1,584 images collected on grids with GO. Extract particles from two data sets were downsized 2-fold and subjected to 2D classification separately, resulting in 438,205 and 560,545 good particles. Particles from GO grids were recentered using scripts written by Kai Zhang before 3D classification. Good particles from both data sets were combined and subjected to 3D classification using the initial model generated from the model of S protein with C1 symmetry (PDB code 6VXX). Two major classes accounting for 26.7% and 25.8% of the particles showed clear and complete structural features. These particles were autorefined, followed by local 3D classification with C1 symmetry to generate two classes with similar structures. The better class was subjected to 3D refinement with C3 symmetry and postprocessed using mask on the entire molecule to yield a 2.9-Å map. CTF refinement was performed to further increase the resolution to 2.6 Å.

For image processing of WT S-Trimer, the MT S-Trimer map was low-pass-filtered to a 20-Å resolution and used as the 3D reference template for autopicking. A total of 752,204 articles were picked from 1,199 images collected on a GO-coated grid and used for two rounds of 2D classification. A total of 541,528 particles were selected from good 2D classes and used for 3D classification with C1 symmetry. One class accounting for 32.2% showing a well-defined structure was refined with C3 symmetry and postprocessed using mask on the entire molecule to give a map at a 3.4-Å resolution. CTF refinement followed by another round of 3D refinement improved the resolution to 3.2 Å. Reported resolutions were calculated based on the gold standard Fourier shell correlation (FSC) at the 0.143 criterion.

Model building. The soluble ectodomain structure S-2P (PDB code 6VXX) was used as a template for model building. The missing region in the S-2P structure can be *de novo* modeled in the MT S-Trimer map, owing to its high resolution. The model was manually built in COOT (23), and real space refinement was performed in Phenix (24) using rotamer, Ramachandran, and secondary-structure restraints. The model for WT S-trimer was built in COOT based on the MT structure and refined in Phenix using the same strategy. Refinement statistics are summarized in Table 1.

Small-molecule extraction from protein samples. Twenty-microliter volumes of WT and MT S-Trimer samples at 1.21 mg/ml and 1.02 mg/ml, respectively, were transferred to Eppendorf tubes and placed in a heater block at 100°C for 5 min. Samples were then extracted with 100 μl of methanol. The tubes were vortexed and centrifuged at 15,000 rpm and 4°C for 10 min. The supernatant was collected and transferred to glass vials, and 1 μl was injected for liquid chromatography-mass spectrometry (LC-MS) analysis. For oleic acid and linoleic acid reference samples, 10 mM stock solutions of compounds were prepared in dimethyl sulfoxide (DMSO) and diluted to 1 μM by methanol, and 1 μl was injected for the LC-MS analysis.

LC-MS analysis. LC-MS analyses of PS80 and MT S-Trimer sample were performed by an Agilent 1290 ultrahigh-performance liquid chromatograph (UHPLC) coupled to an Agilent quadrupole-time of flight (QTOF) mass spectrometer via an electrospray ionization source (ESI) with JetStream technology. The separation was performed by flow injection using isocratic flow of a solvent composed of 0.1% formic acid in 60% acetonitrile and 40% water. The flow rate was set at 0.2 ml/min for 0.5 min. Mass spectra were recorded in the positive ionization mode over a mass range from m/z 100 to 1,500. The scan parameters were capillary voltage of 4.0 kV and fragmentor voltage of 135 V. The nitrogen pressure and flow rate on the nebulizer were 40 lb/in^2 and 5 liters/min, respectively. Other ion source parameters included drying gas temperature of 325°C, sheath gas temperature of 350°C, and sheath gas flow rate of 10 liters/min.

The relative quantitation for oleic acid and linoleic acid was performed by a Thermo Vanquish UHPLC coupled to a Thermo Q Exactive HF-X hybrid quadrupole-Orbitrap mass spectrometer. The chromatographic separation was performed using a Waters CSH C_{18} column (2.1 by 100 mm; 1.7 μm) and maintained at 40°C. The separation was performed using isocratic flow of a solvent composed of 90% acetonitrile, 10% water, and 2 mM ammonium acetate.

SARS-CoV-2 pseudovirus production and infection. SARS-CoV-2 pseudovirus (PV) was produced using the lentiviral system. HEK293T cells were transfected at ~80% confluency in a 6-well plate using Lipofectamine 3000 with 4.8 μg of total DNA. The ratio of 4:3:1 by mass was used for the lentivirus packaging plasmid (pHIV-Luc) encoding luciferase, the pSPAX2 plasmid expressing the lentivirus *gag*, *pol*, and *rev* genes, and a plasmid expressing the full-length spike protein of SARS-CoV-2 or mutated S proteins, including D614G, K854G, or D614G/K854G double mutations. The SARS-CoV-2 S protein gene used in the production of pseudovirus was codon optimized containing a C9 tag at its C terminus. Medium was changed with Dulbecco's modified Eagle's medium (DMEM) containing 2% FBS (D2) at 6 h posttransfection, and pseudovirus-containing culture supernatants were collected at 48 h posttransfection, cleared through 0.45-μm filters, and either purified or aliquoted and frozen at −80°C immediately.

Infection assays were performed using hACE2/TMPRSS2-293T cells seeded on a 96-well plate with equal pseudovirus each well. Inoculated plates were incubated for 24 h in a CO_2 incubator and medium was replaced with DMEM containing 2% FBS. Infection levels were assessed 48 h postinfection by measuring luciferase activity using the Bright-Glo luciferase assay system.

Data availability. Data are available in the Protein Data Bank under accession numbers 7E7B and 7E7D and in EMDB under accession numbers EMD-30998 and EMD-30999.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, MOV file, 0.4 MB.

ACKNOWLEDGMENTS

We thank Xiaodong Wang for his coordination and input in this study. We thank Maofu Liao and Andrew C. Kruse for critical readings of the manuscript. We thank Hongwei Wang for providing GraFuture grids (reduced graphene oxide-coated grids). We also thank staff at Shuimu BioSciences for their assistance with cryo-EM data collection. All EM data were collected at Shuimu BioSciences.

This work was supported by grants from the Coalition for Epidemic Preparedness Innovations (CEPI), Chinese Ministry of Science and Technology (2020YFA0707401), the Beijing Municipal Commission of Science and Technology, Tsinghua University, and the Chengdu Bureau of Science & Technology (2020-YF08-00024-GX).

D.S. expressed and purified the S-Trimers. J.M. collected negative-stain EM images. J.M. and S.Z. prepared cryo-grids and collected cryo-EM data. X.H. and Y.L. performed ForteBio affinity analysis. S.Z. performed cryo-EM data processing and model building. Y.M. did mass spectrometry experiments. Y.S., L.F., and W.L. performed the pseudovirus infection experiments. S.Z. and P.L. conceived this project, directed the experiments, and wrote the manuscript with input from all other authors.

P.L. has ownership interest in Clover Biopharmaceuticals. All other authors declare no competing interests.

REFERENCES

1. de Wit E, van Doremalen N, Falzarano D, Munster VJ. 2016. SARS and MERS: recent insights into emerging coronaviruses. Nat Rev Microbiol 14:523–534. https://doi.org/10.1038/nrmicro.2016.81.
2. Corey L, Mascola JR, Fauci AS, Collins FS. 2020. A strategic approach to COVID-19 vaccine R&D. Science 368:948–950. https://doi.org/10.1126/science.abc5312.
3. Einstein MH, Baron M, Levin MJ, Chatterjee A, Edwards RP, Zepp F, Carletti I, Dessy FJ, Trofa AF, Schuind A, Dubin G, HPV-010 Study Group. 2009. Comparison of the immunogenicity and safety of Cervarix and Gardasil human papillomavirus (HPV) cervical cancer vaccines in healthy women aged 18–45 years. Hum Vaccin 5:705–719. https://doi.org/10.4161/hv.5.10.9518.
4. Bharucha T, Ming D, Breuer J. 2017. A critical appraisal of 'Shingrix', a novel herpes zoster subunit vaccine (HZ/Su or GSK1437173A) for varicella zoster virus. Hum Vaccin Immunother 13:1789–1797. https://doi.org/10.1080/21645515.2017.1317410.
5. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. 2020. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 181:281–292.e6. https://doi.org/10.1016/j.cell.2020.02.058.
6. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367:1260–1263. https://doi.org/10.1126/science.abb2507.
7. Bangaru S, Ozorowski G, Turner HL, Antanasijevic A, Huang D, Wang X, Torres JL, Diedrich JK, Tian JH, Portnoff AD, Patel N, Massare MJ, Yates JR, Nemazee D, Paulson JC, Glenn G, Smith G, Ward AB. 2020. Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. bioRxiv https://doi.org/10.1101/2020.08.06.234674.
8. Liu H, Su D, Zhang J, Ge S, Li Y, Wang F, Gravel M, Roulston A, Song Q, Xu W, Liang JG, Shore G, Wang X, Liang P. 2017. Improvement of pharmacokinetic profile of TRAIL via trimer-tag enhances its antitumor activity in vivo. Sci Rep 7:8953. https://doi.org/10.1038/s41598-017-09518-1.
9. Liang JG, Su D, Song T-Z, Zeng Y, Huang W, Wu J, Xu R, Luo P, Yang X, Zhang X, Luo S, Liang Y, Li X, Huang J, Wang Q, Huang X, Xu Q, Luo M, Huang A, Luo D, Zhao C, Yang F, Han J-B, Zheng Y-T, Liang P. 2021. S-Trimer, a COVID-19 subunit vaccine candidate, induces protective immunity in nonhuman primates. Nat Commun 12:1346. https://doi.org/10.1038/s41467-021-21634-1.
10. Wang H, Zhang Y, Huang B, Deng W, Quan Y, Wang W, Xu W, Zhao Y, Li N, Zhang J, Liang H, Bao L, Xu Y, Ding L, Zhou W, Gao H, Liu J, Niu P, Zhao L, Zhen W, Fu H, Yu S, Zhang Z, Xu G, Li C, Lou Z, Xu M, Qin C, Wu G, Gao GF, Tan W, Yang X. 2020. Development of an inactivated vaccine candidate, BBIBP-CorV, with potent protection against SARS-CoV-2. Cell 182:713–721.e9. https://doi.org/10.1016/j.cell.2020.06.008.
11. Cai Y, Zhang J, Xiao T, Peng H, Sterling SM, Walsh RM, Jr, Rawson S, Rits-Volloch S, Chen B. 2020. Distinct conformational states of SARS-CoV-2 spike protein. Science 369:1586–1592. https://doi.org/10.1126/science.abd4251.
12. Toelzer C, Gupta K, Yadav SKN, Borucu U, Davidson AD, Kavanagh Williamson M, Shoemark DK, Garzoni F, Staufer O, Milligan R, Capin J, Mulholland AJ, Spatz J, Fitzgerald D, Berger I, Schaffitzel C. 2020. Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. Science 370:725–730. https://doi.org/10.1126/science.abd3255.
13. Zhou T, Tsybovsky Y, Olia AS, Gorman J, Rapp MA, Cerutti G, Katsamba PS, Nazzari A, Schon A, Wang PD, Bimela J, Shi W, Teng IT, Zhang B, Boyington JC, Chuang GY, Sampson JM, Sastry M, Stephens T, Stuckey J, Wang S, Friesner RA, Ho DD, Mascola JR, Shapiro L, Kwong PD. 2020. A pH-dependent switch mediates conformational masking of SARS-CoV-2 spike. bioRxiv https://doi.org/10.1101/2020.07.04.187989.
14. Bangaru S, Ozorowski G, Turner HL, Antanasijevic A, Huang D, Wang X, Torres JL, Diedrich JK, Tian JH, Portnoff AD, Patel N, Massare MJ, Yates JR, III, Nemazee D, Paulson JC, Glenn G, Smith G, Ward AB. 2020. Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. Science 370:1089–1094. https://doi.org/10.1126/science.abe1502.

15. Zhang L, Jackson CB, Mou H, Ojha A, Rangarajan ES, Izard T, Farzan M, Choe H. 2020. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. bioRxiv https://doi.org/10.1101/2020.06.12.148726.
16. Grubaugh ND, Hanage WP, Rasmussen AL. 2020. Making sense of mutation: what D614G means for the COVID-19 pandemic remains unclear. Cell 182:794–795. https://doi.org/10.1016/j.cell.2020.06.040.
17. Richmond P, Hatchuel L, Dong M, Ma B, Hu B, Smolenov I, Li P, Liang P, Han HH, Liang J, Clemens R. 2021. Safety and immunogenicity of S-Trimer (SCB-2019), a protein subunit vaccine candidate for COVID-19 in healthy adults: a phase 1, randomised, double-blind, placebo-controlled trial. Lancet 397:682–694. https://doi.org/10.1016/S0140-6736(21)00241-5.
18. Thomas EK, Nakamura M, Wienke D, Isacke CM, Pozzi A, Liang P. 2005. Endo180 binds to the C-terminal region of type I collagen. J Biol Chem 280:22596–22605. https://doi.org/10.1074/jbc.M501155200.
19. Zheng S, Abreu N, Levitz J, Kruse AC. 2019. Structural basis for KCTD-mediated rapid desensitization of GABAB signalling. Nature 567:127–131. https://doi.org/10.1038/s41586-019-0990-0.
20. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods 14:331–332. https://doi.org/10.1038/nmeth.4193.
21. Rohou A, Grigorieff N. 2015. CTFFIND4: fast and accurate defocus estimation from electron micrographs. J Struct Biol 192:216–221. https://doi.org/10.1016/j.jsb.2015.08.008.
22. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife 7:e42166. https://doi.org/10.7554/eLife.42166.
23. Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132. https://doi.org/10.1107/S0907444904019158.
24. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66:213–221. https://doi.org/10.1107/S0907444909052925.