Molecular architecture of the SARS-CoV-2 virus

Hangping Yao[1,2,9], Yutong Song[3,4,9], Yong Chen[3,4,9], Nanping Wu[1,2,9], Jialu Xu[3,4,5,9], Chujie Sun[3,4,5], Jiaxing Zhang[3,4], Tianhao Weng[1,2], Zheyuan Zhang[3,4], Zhigang Wu[1,2], Linfang Cheng[1,2], Danrong Shi[1,2], Xiangyun Lu[1,2], Jianlin Lei[3,4], Max Crispin[6], Yigong Shi[3,4,5,7,8], Lanjuan Li[1,2,*] and Sai Li[3,4,5,10,*]

[1]State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, Hangzhou 310003, China

[2]National Clinical Research Center for Infectious Diseases, First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China

[3]School of Life Sciences, Tsinghua University, Beijing 100084, China

[4]Beijing Advanced Innovation Center for Structural Biology & Frontier Research Center for Biological Structure, Beijing 100084, China

[5]Tsinghua University-Peking University Joint Center for Life Sciences, Beijing 100084, China

[6]School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, UK

[7]Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, 18 Shilongshan Road, Hangzhou 310024, Zhejiang Province, China

[8]Institute of Biology, Westlake Institute for Advanced Study, 18 Shilongshan Road, Hangzhou 310024, Zhejiang Province, China

[9]These authors contributed equally

[10]Lead Contact

[*]Correspondence: sai@tsinghua.edu.cn (S.L.), ljli@zju.edu.cn (L.L.)

SUMMARY

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an enveloped virus responsible for the COVID-19 pandemic. Despite recent advances in the structural elucidation of SARS-CoV-2 proteins and the complexes of the spike (S) proteins with the cellular receptor ACE2 or neutralizing antibodies, detailed architecture of the intact virus remains to be unveiled. Here we report the molecular assembly of the authentic SARS-CoV-2 virus using cryo-electron tomography (cryo-ET) and subtomogram averaging (STA). Native structures of the S proteins in both pre- and postfusion conformations were determined to average resolutions of 8.7-11 Å. Compositions of the N-linked glycans from the native spikes were analyzed by mass-spectrometry, which revealed highly similar overall processing states of the native glycans to that of the recombinant glycoprotein glycans. The native conformation of the ribonucleoproteins (RNP) and its higher-order assemblies were revealed. Overall, these characterizations have revealed the architecture of the SARS-CoV-2 virus in unprecedented detail, and shed lights on how the virus packs its ~30 kb long single-segmented RNA in the ~80 nm diameter lumen.

INTRODUCTION

As of August 10th, 2020, a total of over 20 million cases of COVID-19 were reported and more than 700 thousand lives were claimed globally (https://covid19.who.int). The causative pathogen, SARS-CoV-2, is a novel β-coronavirus (Lu et al., 2020; Wu et al., 2020; Zhou et al., 2020). SARS-CoV-2 encodes at least 29 proteins in its (+) RNA genome, four of which are structural proteins: the spike (S), membrane (M), envelope (E) and nucleocapsid (N) proteins (Kim et al., 2020).

The ~600 kDa, trimeric S protein, one of the largest known class-I fusion proteins, is heavily glycosylated with 66 N-linked glycans (Walls et al., 2020; Watanabe et al., 2020a; Wrapp et al., 2020). Each S protomer comprises the S1 and S2 subunits, and a single transmembrane (TM) anchor (Wrapp et al., 2020). The S protein binds to the cellular surface receptor angiotensin-converting enzyme-2 (ACE2) through the receptor binding domain (RBD), an essential step for membrane fusion (Hoffmann et al., 2020; Lan et al., 2020; Shang et al., 2020; Wang et al., 2020; Yan et al., 2020; Zhou et al., 2020). The activation of S requires cleavage of S1/S2 by furin-like protease and undergoes the conformational change from prefusion to postfusion (Belouzard et al., 2009; Kirchdoerfer et al., 2018; Simmons et al., 2004; Simmons et al., 2013; Song et al., 2018). Several prefusion conformations have been resolved for the S protein, wherein the three RBDs display distinct orientations, "up" or "down" (Walls et al., 2020; Wrapp et al., 2020). The receptor binding sites expose, only when the RBDs adopt an 'up' conformation. The "RBD down", "one RBD up" and "two-RBD up" conformations have been observed in recombinantly expressed S proteins of SARS-CoV-2 (Henderson et al., 2020; Walls et al., 2020; Wrapp et al., 2020). Upon activation, S follows a classic pathway among class-I fusion proteins (Rey and Lok, 2018): it undergoes dramatic structural rearrangements involving shedding its S1 subunit and inserting the fusion peptide (FP) into the target cell membrane (Cai et al., 2020). Following membrane fusion, S transforms to a needle-shaped postfusion form, having three helixes entwining coaxially (Cai et al., 2020). Despite the efforts in elucidating the SARS-CoV-2 virus host recognition and entry mechanism at near-atomic resolution using recombinant proteins, high-resolution information regarding the *in situ* structures and landscape of the authentic virus is in-demand.

Coronavirus has the largest genome among all RNA viruses. It is enigmatic how the N protein oligomerizes, organizes, and packs the ~30 kb long single-stranded RNA in the viral lumen. Early negative-staining electron microscopy of coronaviruses showed single-strand helical

RNPs with a diameter of ~15 nm (Caul and Eggleston, 1979). Cryo-ET of SARS-CoV revealed that RNPs organized into lattices underneath the envelope at ~4-5 nm resolution (Neuman et al., 2006). However, such ultrastructure was not observed in the Mouse hepatitis virus (MHV), the prototypic β-coronavirus (Barcena et al., 2009). No molecular model exists so far for the coronavirus RNP, and little is known about the architecture, assembly and RNA packaging of the RNPs of other (+) RNA viruses.

To address these questions, we combined cryo-ET and STA for the imaging analysis of 2,294 intact virions propagated from an early viral strain (Yao et al., 2020). To our knowledge, this is the largest cryo-ET data set of SARS-CoV-2 virus to date. Here we report the architecture and assembly of the authentic SARS-CoV-2 virus.

RESULTS AND DISCUSSION

Molecular landscape of the SARS-CoV-2 virus.

SARS-CoV-2 virions (ID: ZJU_5) were collected on January 22nd, 2020, from a patient with severe symptoms, and were propagated in Vero cells. The patient was infected during a conference with attendees from Wuhan (Yao et al., 2020). For cryo-EM analysis, the viral sample was fixed by paraformaldehyde, which has minor effects on protein structure at 7-20 Å resolution (Li et al., 2016; Wan et al., 2017). Intact and unconcentrated virions were directly visualized from the supernatant by cryo-EM, showing both ellipsoidal and spherical enveloped particles (Figure S1A), consistent with the observation on virions concentrated by ultracentrifugation through sucrose cushion (Figure 1B). We modelled 2,294 virions as ellipsoids by meshing their lipid envelopes, measuring average diameters of 64.8±11.8 nm, 85.9±9.4 nm and 96.6±11.8 nm for the short, medium and long axis of the envelope respectively. S proteins and RNPs (Figures 1A, S1B, S5A and Video S1) represented the most distinctive features of the virus. A comprehensive analysis has been carried out on the data. In total, 56,832 spikes were manually identified from the virions, approximately 97% of which are in the prefusion conformation, 3% in the postfusion conformation (Method details). An average of 26 prefusion S were found randomly distributed on each virion (Figures 1B-C). The spike copy number per virion is comparable to HIV (Liu et al., 2008), but ~5 times less than the Lassa virus (LASV) (Li et al., 2016) or ~10 times less than the Influenza virus (Harris et al., 2006). 18,500 RNPs were manually identified in the viral lumen (Table S1), giving an

average of 26 RNPs per virion. However, since the viral lumen is tightly packed with RNPs and electron opaque, the actual number of RNPs per virion was estimated to be 20-30% more, i.e. 30-35 RNPs per virion. Regularly ordered RNP ultrastructures were occasionally observed (Figure S5A), indicating the RNPs could form local assemblies.

Two conformations of the prefusion S, namely the RBD down and one RBD up conformations from inactivated SARS-CoV-2 virions were classified and reconstructed to 8.7 Å and 10.9 Å resolution by STA, with local resolution reaching 7.8 Å (Figures S3A-S3C). The heptad repeat 1 (HR1) and central helix (CH) domains of the S2 subunits represent the best resolved domains (Figure S3D). The proportion of RBD down conformation among all prefusion S was estimated to be 54% per virion (Figure 1C). The membrane proximal stalk of S represented the poorest resolved region with a local resolution of ~20 Å, showing no trace of the TM or membrane in the structure (Figure 1D). When the tomogram slices were scrutinized, spike populations that either stand perpendicular to or lean towards the envelope were observed (Video S1), suggesting that the TM was averaged out in the map.

Refined orientations of the prefusion S showed they rotate around their stalks almost freely outside the envelope, with the highest possibility of leaning 40° relative to the normal axis of the envelope (Figures 1B and 1D). The rotational freedom of spikes is allowed by its low population density, which is prominently distinct from other enveloped viruses possessing class-I fusion proteins (Harris et al., 2006; Li et al., 2016; Liu et al., 2008). Interestingly, a minor population of Y-shaped spikes pair having two heads and one combined stem were observed (Figures S2A and S2C), which possibly represent spikes intertwined with their stems. These observations suggest that the SARS-CoV-2 spikes possess unusual freedom on the viral envelope. Such unique features may facilitate the virus in exploring the surrounding environment and better engaging with the cellular receptor ACE2, allowing multiple spikes to bind with one ACE2 or one spike with multiple ACE2 simultaneously. However, the sparsely packed spikes on the viral envelope are also more vulnerable to neutralizing antibodies that bind the otherwise less accessible domains (Chi et al., 2020) or glycan holes (Walls et al., 2019). Our observations on the structures and landscape of the intact SARS-CoV-2 are consistent with two other cryo-ET studies appeared at the same period (Turoňová, 2020; Ke et al., 2020).

Native structures of S in the prefusion conformation.

The native structures of S in the RBD down and one RBD up conformations were similar to the rigidly fitted recombinant protein structures (PDB: 6XR8 and 6VYB) (Cai et al., 2020; Walls et al., 2020), except for the N terminus domain (NTD). Comparison between the rigidly and flexibly fitted PDB: 6XR8 suggested that the NTD on the native spike structure shifted 9 Å (centroid distance) away from S2 (Figure S3E). The slight dilation and lower local resolution (Figure S3A and S3B) of the NTD on the native spike against recombinant spike structures was also observed on the other cryo-ET structures (Ke et al., 2020). It is known that NTD exhibits certain mobility as a rigid body (Cai et al., 2020; Walls et al., 2020; Wrapp et al., 2020; Xiong et al., 2020). It is possible that through large date set and classification, the near-atomic resolution cryo-EM reconstruction of the spike represents its metastable conformation, while the cryo-ET reconstructions represent an average of various dynamic states of the NTD.

Ten N-linked glycans are visible on the RBD down and seven on the one RBD up conformations, of which N61, N282, N801, N1098 and N1134 were best resolved. Interestingly, densities for glycans N1158 and N1173/N1194 are visible on the stem of the spike (Figures 2B and S4C). In general, the glycan densities observed on the native spike fit well with the full-length recombinant structure (Cai et al., 2020), however they are bulkier than observed in the TM truncated recombinant spikes (Walls et al., 2020; Wrapp et al., 2020). Similar observations were reported by two recent studies of the native spike structures (Turoňová, 2020; Ke et al., 2020).

We further determined the native glycan identity by analyzing the virus sample using mass spectrometry (MS). The viral particles, with or without PNGase F digestion, were resolved on SDS-PAGE. After PNGase F treatment, the S1 and S2 subunits were reduced by ~30 kDa and ~20 kDa, respectively, in weight (Figure 2C). The bands corresponding to S1 and S2 before PNGase F treatment were analyzed by MS to reveal the glycan compositions at each of the 22 glycosylation sites (Figures 2D and S4B). The overall processing states of the native glycans are highly similar to that of the recombinant glycoprotein glycans (Figures S4A and S4B) (Watanabe et al., 2020a), a feature shared with MERS and SARS-CoV (Walls et al., 2019). Populations of under-processed oligomannose-type glycans are found at the same sites as seen in the recombinant material, including at N234 where the glycan is suggested to have a structural role (Casalino et al., 2020). However, as is observed at many sites on HIV (Cao et al., 2018; Struwe et al., 2018), the virus exhibits somewhat lower levels of oligomannose-type glycosylation compared to the recombinant, soluble mimetic. Overall, the presence of

substantial population of complex-type glycosylation suggests that the budding route of SARS-CoV-2 into the lumen of endoplasmic reticulum-Golgi intermediate compartments (ERGIC) is not an impediment to glycan maturation and is consistent with both analysis of SARS-CoV glycans (Ritchie et al., 2010) and the identification of neutralizing antibodies targeting the fucose at the N343 glycan on SARS-CoV-2 (Pinto et al., 2020). Furthermore, the lower levels of oligomannose-type glycans compared to HIV and LASV are also consistent with lower glycan density (Watanabe et al., 2020b; Watanabe et al., 2018). Comparing the structures between our native prefusion S to the recombinant ones, we conclude that 1) the N-linked glycans present on the native spike are bulkier and contain elevated levels of complex-type glycans; 2) the recapitulation of the main features of native viral glycosylation by soluble, trimeric recombinant S glycoprotein is encouraging for vaccine strategies utilizing recombinant S protein.

Native structure of S in the postfusion conformation.

Apart from the triangular prefusion S, needle-like densities were occasionally observed on the viral envelope (Figure S2B). A 15.3 Å resolution structure was solved and well fitted with a postfusion form of S (PDB: 6XRA), suggesting a postfusion conformation of the S protein. The ectodomain measures 22.5 nm in length, 6 nm in width and stands perpendicular to the viral envelope. Densities of five N-linked glycans (N1098, N1074, N1158, N1173 and N1194) are displayed on the spike (Figure 3D), reminiscent to the recombinant postfusion structures (Cai et al., 2020; Fan et al., 2020). In comparison to the prefusion conformation, the fixation in orientation to the envelope suggests a dramatic conformational reordering of the stem region to achieve the postfusion conformation. The postfusion S were found only on a small fraction of the viruses (18% of all viruses, each carried an average of 5 postfusion S), hindering us from reaching higher resolutions. Interestingly, among postfusion S-carrying virions, the virions possessing less spikes in total tend to have higher proportion of postfusion S (Figure 3C). Statistics on the refined coordinates (Figures 3A-3C) suggested a tendency of the postfusion S to bundle on the viral surface. One such bundle was reconstructed, showing four postfusion S spaced by ~10 nm on the viral surface (Figure 3B), compared to ~15 nm average distance between the nearest prefusion S.

Distinguished from the SARS-CoV virus, which was estimated to possess an average of ~50-100 spikes per virion (Neuman et al., 2011), the SARS-CoV-2 virus possesses approximately half of the prefusion S and occasionally some postfusion S. The postfusion S observed on the

SARS-CoV-2 virus may come from 1) products of occasional, spontaneous dissociation of S1 (Cai et al., 2020), which was cleaved by host proteinases; 2) syncytium naturally formed on infected cells (Xia et al., 2020), when budding progeny virions carried a few residual postfusion S from the cell surface; 3) sample preparation procedure, as cryo-EM images of ß-propiolactone fixed viruses showed most spikes present on the virus are postfusion-like (Liu et al., 2020; Gao et al., 2020). Such instability of the prefusion S was reported on the other β-coronaviruses (Pallesen et al., 2017). In addition, the distribution graph (Figure 3C) implies that the kinetically trapped prefusion S is more fragile than the postfusion S, and could even dissociate from the virus. The speculation is based on the fact that intracellular virions possess more spikes on average (Klein et al., 2020) than the extracellular virions reported by us and the others (Turoňová, 2020; Ke et al., 2020), and the occasional observation of spike-less "bald" virus in our data.

In summary, we speculate that the SARS-CoV-2 prefusion S are unstable, indicating that the distribution of solvent exposed epitopes on the virions is more complicated than the observations on the recombinant proteins. Our observation has implications for efficient vaccine design and neutralizing antibody development, which prefer a sufficient number of stable antigens.

Architecture and assembly of the RNPs in intact virions.

It remains enigmatic how coronaviruses pack the ~30 kb RNA within the ~80 nm diameter viral lumen; if the RNPs are ordered relative to each other to avoid RNA entangling, knotting or even damaging; or if the RNPs are involved in virus assembly. When raw tomogram slices were inspected, densely packed, bucket-like densities were discernible throughout the virus lumen, some of which appeared to be locally ordered (Figure S5A). Combining previous cryo-ET observation on coronaviruses (Barcena et al., 2009) and SDS-PAGE/MS analysis (Figure 2C), the densities most likely belonged to the RNPs.

In total 18,500 RNPs were picked in the viral lumen, and initially aligned using a sphere as the template and a large spherical mask. A bucket-like conformation with little structural feature emerged adjacent to the density for lipid bilayer (Figure S5D), suggesting that a significant number of RNPs were membrane proximal. Alignment using a small spherical mask revealed a 13.1 Å resolution reverse G-shaped architecture of the RNP, measuring 15 nm in diameter and 16 nm in height (Figure 4B). Its shape is in good comparison to a recently reported SARS-

CoV-2 RNP conformation (Klein et al., 2020), as well as the *in situ* conformation of the chikungunya viral RNP, which is also positive stranded (Jin et al., 2018), but different from the MHV RNP released using detergent (Gui et al., 2017).

The map was segmented into five head-to-tail reverse L-shaped densities, each fitted with a pair of N proteins (N terminus domain (N_NTD): 6WKP, C terminus domain (N_CTD): 6WJI) dimerized by the N_CTD (Chen et al., 2007) (Figure S6A and S6C). We analyzed the electrostatic potential distribution on the surface of the decamer, and suggested a tentative structural model of RNA winded RNP (Figures S6B and S6D). Interestingly, an early observation on the MHV RNPs showed ~15 nm diameter helices with five subunits per turn (Caul and Eggleston, 1979). Due to the limited resolution and little previous structural knowledge about the (+) RNA virus RNPs, our model shall be interpreted with caution.

Further 2D classification of the RNPs revealed three classes: 1) closely packed against the envelope, 2) hexagonally and 3) triangularly packed RNPs (Figure 4A). Following 3D refinement, a membrane proximal, "eggs-in-a-nest" shaped RNP assembly (referred to as the "hexon"), and a membrane-free, "pyramid" shaped RNP assembly (referred to as the "tetrahedron") emerged (Figures 4C and S5E-S5G). Projection of the two class averages back onto their refined coordinates revealed that the majority of hexons came from spherical virions, while more tetrahedrons from ellipsoidal virions (Figure 4D). This was quantified by statistics: ellipsoidal virions tend to pack more RNP tetrahedrons (Figure 4E). Furthermore, the spacing between two neighboring RNPs (~18 nm) is the same for both the tetrahedrons and hexons, and some tetrahedrons could assemble into hexons when projected onto their *in situ* coordinates (Figure S5B), suggesting that the RNP triangle is a key and basic packing unit throughout the virus. We further propose that the RNPs are involved in coronavirus assembly and help strengthen the virus against environmental and physical challenges, as purified virions remained intact after five cycles of freeze-and-thaw treatment (Figure S7). Such involvement of RNPs in viral assembly was also reported by (Neuman et al., 2006), who showed RNPs form a lattice underneath the envelope; as well as seen in intracellular virions (Steffen Klein et al., 2020). However, it remains unanswered if the ultrastructures of RNPs are assembled by RNA, the transmembrane M or E proteins, the RNP itself, or multiples of the above.

Solving RNPs to subnanometer resolution was hindered by the crowding of RNPs against each other (Figures S5E-S5G). Furthermore, structural features of the RNPs on higher order assemblies smeared, possibly due to the symmetry mis-match between individual RNPs and

the assembly (Figures 4C and 4D). No virus with strictly ordered RNPs throughout the lumen was found by projections. We conclude that the native RNPs are highly heterogeneous, densely packed yet locally ordered in the virus, possibly interacting with the RNA in a beads-on-a-string stoichiometry.

From ~2,300 intact virions and over 300 tilt series, we provide molecular insights into the structures of spikes in the pre- and postfusion conformations, the RNPs and how they assemble on the authentic virus. We also analyzed the detailed glycan compositions of the native spikes. The reconstructed virus map (Figure 1B) consisting the RBD down S, one RBD up S, lipid envelope and RNP components of the authentic SARS-CoV-2 has been deposited in the Electron Microscopy Data Bank under accession codes EMD-30430, which provide a model for full virus molecular dynamic simulation, 3D printing, education or public media in the future.

ACKNOWLEDGMENTS

S.L. thank Tsinghua University for providing a Start-up fund, the Tsinghua University Branch of China National Center for Protein Sciences (Beijing) for the cryo-EM facility and the computational facility support, and Dr. Fan Yang, Jie Wen, Danyang Li, Yakun Wang, Anbao Jia and Tao Yang for technical support. We thank Dr. Haiteng Deng and Chongchong Zhao in the Proteinomics Facility at Technology Center for Protein Sciences, Tsinghua University, for protein MS analysis. We thank the computational facility support on the cluster of Bio-Computing Platform (Tsinghua University Branch of China National Center for Protein Sciences Beijing). We are in debt to Dr. Hongwei Wang, Dr. Nieng Yan, Dr. Xinquan Wang, Dr. Lingqi Zhang, Dr. Qiang Ding and Dr. Xueming Li for providing critical advices. **Funding:** This work was supported in part by funds from Major Project of Zhejiang Provincial Science and Technology Department #2020C03123-1, and National Science and Technology Major Project for the Control and Prevention of Major Infectious Diseases in China (#2018ZX10711001, #2018ZX10102001).

AUTHOR CONTRIBUTIONS

S.L. conceived and supervised the project. H.Y., N.W., T.W., Z.W., L.C., D.S. and X.L. propagated, fixed and verified the fixation of the virus sample. Y.S. and S.L. isolated the viruses, performed biochemical analysis and prepared the cryo-sample for cryo-EM. Y.S., Y.C., S.L., J.Z. and J.L. collected the EM data. Y.S., Y.C., J.Z., J.X., C.S., Z.Z., and S.L. processed the EM data. S.L. and J.X. performed the subtomogram averaging and classification. Y.C., J.X., S.L. and Y.S. analyzed the structures. C.S., and J.X. performed the statistical analysis. M.C., Y.S. and Z.Z. analyzed the glycan data. S.L., Y.Shi., M.C. and L.L. wrote the manuscript. All authors critically revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Barcena, M., Oostergetel, G.T., Bartelink, W., Faas, F.G., Verkleij, A., Rottier, P.J., Koster, A.J., and Bosch, B.J. (2009). Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. Proc. Natl. Acad. Sci. U S A *106*, 582-587.

Beata Turoňová, M.S., Christoph Schürmann, Wim J. H. Hagen, Sonja Welsch, Florian E. C. Blanc, Sören von Bülow, Michael Gecht, Katrin Bagola, Cindy Hörner, Ger van Zandbergen, Shyamal Mosalaganti, Andre Schwarz, Roberto Covino, Michael D. Mühlebach, Gerhard Hummer, Jacomine Krijnse Locker, Martin Beck (2020). In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. bioRxiv, https://doi.org/10.1101/2020.1106.1126.173476.

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

Bharat, T.A.M., and Scheres, S.H.W. (2016). Resolving macromolecular structures from electron cryo-tomography data using subtomogram averaging in RELION. Nat. Protoc. *11*, 9-20.

Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S.M., Walsh, R.M., Jr., Rawson, S., Rits-Volloch, S., and Chen, B. (2020). Distinct conformational states of SARS-CoV-2 spike protein. Science, eabd4251.

Cao, L., Pauthner, M., Andrabi, R., Rantalainen, K., Berndsen, Z., Diedrich, J.K., Menis, S., Sok, D., Bastidas, R., Park, S.R.*, et al.* (2018). Differential processing of HIV envelope glycans on the virus and soluble recombinant trimer. Nat. Commun. *9*, 3693.

Castano-Diez, D., Kudryashev, M., Arheit, M., and Stahlberg, H. (2012). Dynamo: a flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. J. Struct. Biol. *178*, 139-151.

Castano-Diez, D., Kudryashev, M., and Stahlberg, H. (2016). Dynamo Catalogue: Geometrical tools and data management for particle picking in subtomogram averaging of cryo-electron tomograms. J. Struct. Biol. *197*, 135-144.

Caul, E.O., and Eggleston, S.I. (1979). Coronavirus-like particles present in simian faeces. Vet. Rec. *104*, 168-169.

Chang, C.K., Hsu, Y.L., Chang, Y.H., Chao, F.A., Wu, M.C., Huang, Y.S., Hu, C.K., and Huang, T.H. (2009). Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: implications for ribonucleocapsid protein packaging. J. Virol. *83*, 2255-2264.

Chen, C.Y., Chang, C.K., Chang, Y.W., Sue, S.C., Bai, H.I., Riang, L., Hsiao, C.D., and Huang, T.H. (2007). Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J. Mol. Biol. *368*, 1075-1086.

Chi, X., Yan, R., Zhang, J., Zhang, G., Zhang, Y., Hao, M., Zhang, Z., Fan, P., Dong, Y., Yang, Y.*, et al.* (2020). A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. Science *369*, 650-655.

Chuang Liu, Y.Y., Yuanzhu Gao, Chenguang Shen, Bin Ju, Congcong Liu, Xian Tang, Jinli Wei, Xiaomin Ma, Weilong Liu, Shuman Xu, Yingxia Liu, Jing Yuan, Jing Wu, Zheng Liu, Zheng Zhang, Peiyi Wang, Lei Liu (2020). Viral Architecture of SARS-CoV-2 with Post-Fusion Spike Revealed by Cryo-EM. bioRxiv, https://doi.org/10.1101/2020.1103.1102.972927.

Fan, X., Cao, D., Kong, L., and Zhang, X. (2020). Cryo-EM analysis of the post-fusion structure of the SARS-CoV spike glycoprotein. Nat. Commun. *11*, 3618.

Gao, Q., Bao, L., Mao, H., Wang, L., Xu, K., Yang, M., Li, Y., Zhu, L., Wang, N., Lv, Z.*, et al.* (2020). Development of an inactivated vaccine candidate for SARS-CoV-2. Science *369*, 77-81.

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

Gui, M., Liu, X., Guo, D.Y., Zhang, Z., Yin, C.C., Chen, Y., and Xiang, Y. (2017). Electron microscopy studies of the coronavirus ribonucleoprotein complex. Protein Cell *8*, 219-224.

Hagen, W.J.H., Wan, W., and Briggs, J.A.G. (2017). Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. J. Struct. Biol. *197*, 191-198.

Hangping Yao, Xiangyun Lu, Qiong Chen, Kaijin Xu, Yu Chen, Linfang Cheng, Fumin Liu, Zhigang Wu, Haibo Wu, Changzhong Jin*, et al.* (2020). Patient-derived mutations impact pathogenicity of SARS-CoV-2. medRxiv, https://doi.org/10.1101/2020.1104.1114.20060160.

Harris, A., Cardone, G., Winkler, D.C., Heymann, J.B., Brecher, M., White, J.M., and Steven, A.C. (2006). Influenza virus pleiomorphy characterized by cryoelectron tomography. Proc. Natl. Acad. Sci. U S A *103*, 19123-19127.

Henderson, R., Edwards, R.J., Mansouri, K., Janowska, K., Stalls, V., Gobeil, S.M.C., Kopp, M., Li, D., Parks, R., Hsu, A.L.*, et al.* (2020). Controlling the SARS-CoV-2 spike glycoprotein conformation. Nat. Struct. Mol. Biol.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A.*, et al.* (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell *181*, 271-280 e278.

Huiskonen, J.T., Parsy, M.L., Li, S., Bitto, D., Renner, M., and Bowden, T.A. (2014). Averaging of Viral Envelope Glycoprotein Spikes from Electron Cryotomography Reconstructions using Jsubtomo. Jove-J. Vis. Exp., e51714.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33-38, 27-38.

Jin, J., Galaz-Montoya, J.G., Sherman, M.B., Sun, S.Y., Goldsmith, C.S., O'Toole, E.T., Ackerman, L., Carlson, L.A., Weaver, S.C., Chiu, W.*, et al.* (2018). Neutralizing Antibodies Inhibit Chikungunya Virus Budding at the Plasma Membrane. Cell Host Microbe *24*, 417.

Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N., and Chang, H. (2020). The Architecture of SARS-CoV-2 Transcriptome. Cell *181*, 914-921 e910.

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

Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. *116*, 71-76.

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L.*, et al.* (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature *581*, 215-220.

Li, S., Sun, Z., Pryce, R., Parsy, M.L., Fehling, S.K., Schlie, K., Siebert, C.A., Garten, W., Bowden, T.A., Strecker, T.*, et al.* (2016). Acidic pH-Induced Conformations and LAMP1 Binding of the Lassa Virus Glycoprotein Spike. PLoS Pathog. *12*, e1005418.

Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat. Methods *10*, 584-590.

Liu, J., Bartesaghi, A., Borgnia, M.J., Sapiro, G., and Subramaniam, S. (2008). Molecular architecture of native HIV-1 gp120 trimers. Nature *455*, 109-113.

Lorenzo Casalino, Zied Gaieb, Abigail C. Dommer, Aoife M. Harbison, Carl A. Fogarty, Emilia P. Barros, Bryn C. Taylor, Elisa Fadda, and Amaro, R.E. (2020). Shielding and Beyond: The Roles of Glycans in SARS-CoV-2 Spike Protein. bioRxiv, https://doi.org/10.1101/2020.1106.1111.146522.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N.*, et al.* (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet *395*, 565-574.

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S.*, et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B *102*, 3586-3616.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36-51.

Neuman, B.W., Adair, B.D., Yeager, M., and Buchmeier, M.J. (2008). Purification and electron cryomicroscopy of coronavirus particles. Methods Mol. Biol. *454*, 129-136.

Neuman, B.W., Adair, B.D., Yoshioka, C., Quispe, J.D., Orca, G., Kuhn, P., Milligan, R.A., Yeager, M., and Buchmeier, M.J. (2006). Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. J. Virol. *80*, 7918-7928.

Neuman, B.W., Kiss, G., Kunding, A.H., Bhella, D., Baksh, M.F., Connelly, S., Droese, B., Klaus, J.P., Makino, S., Sawicki, S.G., *et al.* (2011). A structural analysis of M protein in coronavirus assembly and morphology. J. Struct. Biol. *174*, 11-22.

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

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF chimera - A visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605-1612.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J Comput Chem *26*, 1781-1802.

Pintilie, G.D., Zhang, J., Goddard, T.D., Chiu, W., and Gossard, D.C. (2010). Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. J. Struct. Biol. *170*, 427-438.

Pinto, D., Park, Y.J., Beltramello, M., Walls, A.C., Tortorici, M.A., Bianchi, S., Jaconi, S., Culap, K., Zatta, F., De Marco, A., *et al.* (2020). Structural and functional analysis of a potent sarbecovirus neutralizing antibody. bioRxiv, https://doi.org/10.1101/2020.1104.1107.023903.

Rey, F.A., and Lok, S.M. (2018). Common Features of Enveloped Viruses and Implications for Immunogen Design for Next-Generation Vaccines. Cell *172*, 1319-1334.

Ritchie, G., Harvey, D.J., Feldmann, F., Stroeher, U., Feldmann, H., Royle, L., Dwek, R.A., and Rudd, P.M. (2010). Identification of N-linked carbohydrates from severe acute respiratory syndrome (SARS) spike glycoprotein. Virology *399*, 257-269.

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., and Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. Nature *581*, 221-224.

Simmons, G., Reeves, J.D., Rennekamp, A.J., Amberg, S.M., Piefer, A.J., and Bates, P. (2004). Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. Proc. Natl. Acad. Sci. U S A *101*, 4240-4245.

Simmons, G., Zmora, P., Gierer, S., Heurich, A., and Pohlmann, S. (2013). Proteolytic activation of the SARS-coronavirus spike protein: cutting enzymes at the cutting edge of antiviral research. Antiviral Res. *100*, 605-614.

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

Steffen Klein, Mirko Cortese, Sophie L. Winter, Moritz Wachsmuth-Melm, Christopher J. Neufeldt, Berati Cerikan, Megan L. Stanifer, Steeve Boulant, Ralf Bartenschlager, and Chlanda, P. (2020). SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. bioRxiv, https://doi.org/10.1101/2020.1106.1123.167064.

Struwe, W.B., Chertova, E., Allen, J.D., Seabright, G.E., Watanabe, Y., Harvey, D.J., Medina-Ramirez, M., Roser, J.D., Smith, R., Westcott, D., *et al.* (2018). Site-Specific Glycosylation of Virion-Derived HIV-1 Env Is Mimicked by a Soluble Trimeric Immunogen. Cell Rep. *24*, 1958-1966 e1955.

Trabuco, L.G., Villa, E., Schreiner, E., Harrison, C.B., and Schulten, K. (2009). Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. Methods *49*, 174-180.

Turonova, B., Schur, F.K.M., Wan, W., and Briggs, J.A.G. (2017). Efficient 3D-CTF correction for cryo-electron tomography using NovaCTF improves subtomogram averaging resolution to 3.4A. J. Struct. Biol. *199*, 187-195.

Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell *181*, 281-292.

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

Wan, W., Kolesnikova, L., Clarke, M., Koehler, A., Noda, T., Becker, S., and Briggs, J.A.G. (2017). Structure and assembly of the Ebola virus nucleocapsid. Nature *551*, 394-397.

Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., Lu, G., Qiao, C., Hu, Y., Yuen, K.Y., *et al.* (2020). Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. Cell *181*, 894-904 e899.

Watanabe, Y., Allen, J.D., Wrapp, D., McLellan, J.S., and Crispin, M. (2020a). Site-specific glycan analysis of the SARS-CoV-2 spike. Science, eabb9983.

Watanabe, Y., Berndsen, Z.T., Raghwani, J., Seabright, G.E., Allen, J.D., Pybus, O.G., McLellan, J.S., Wilson, I.A., Bowden, T.A., Ward, A.B., *et al.* (2020b). Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. Nat. Commun. *11*, 2688.

Watanabe, Y., Raghwani, J., Allen, J.D., Seabright, G.E., Li, S., Moser, F., Huiskonen, J.T., Strecker, T., Bowden, T.A., and Crispin, M. (2018). Structure of the Lassa virus glycan shield provides a model for immunological resistance. Proc. Natl. Acad. Sci. U S A *115*, 7320-7325.

Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.L., Abiona, O., Graham, B.S., and McLellan, J.S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science *367*, 1260-1263.

Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., *et al.* (2020). A new coronavirus associated with human respiratory disease in China. Nature *579*, 265-269.

Xia, S., Liu, M.Q., Wang, C., Xu, W., Lan, Q.S., Feng, S.L., Qi, F.F., Bao, L.L., Du, L.Y., Liu, S.W., *et al.* (2020). 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.

Xiong, X., Qu, K., Ciazynska, K.A., Hosmillo, M., Carter, A.P., Ebrahimi, S., Ke, Z., Scheres, S.H.W., Bergamaschi, L., Grice, G.L., *et al.* (2020). A thermostable, closed SARS-CoV-2 spike protein trimer. Nat. Struct. Mol. Biol.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., and Zhou, Q. (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science *367*, 1444-1448.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1-12.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331-332.

Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., *et al.* (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature *579*, 270-273.

Zunlong Ke, Joaquin Oton, Kun Qu, Mirko Cortese, Vojtech Zila, Lesley McKeane, Takanori Nakane, Jasenko Zivanov, Christopher J. Neufeldt, John M. Lu, *et al.* (2020). Structures, conformations and distributions of SARS-CoV-2 spike protein trimers on intact virions. bioRxiv, https://doi.org/10.1101/2020.1106.1127.174979.

FIGURES

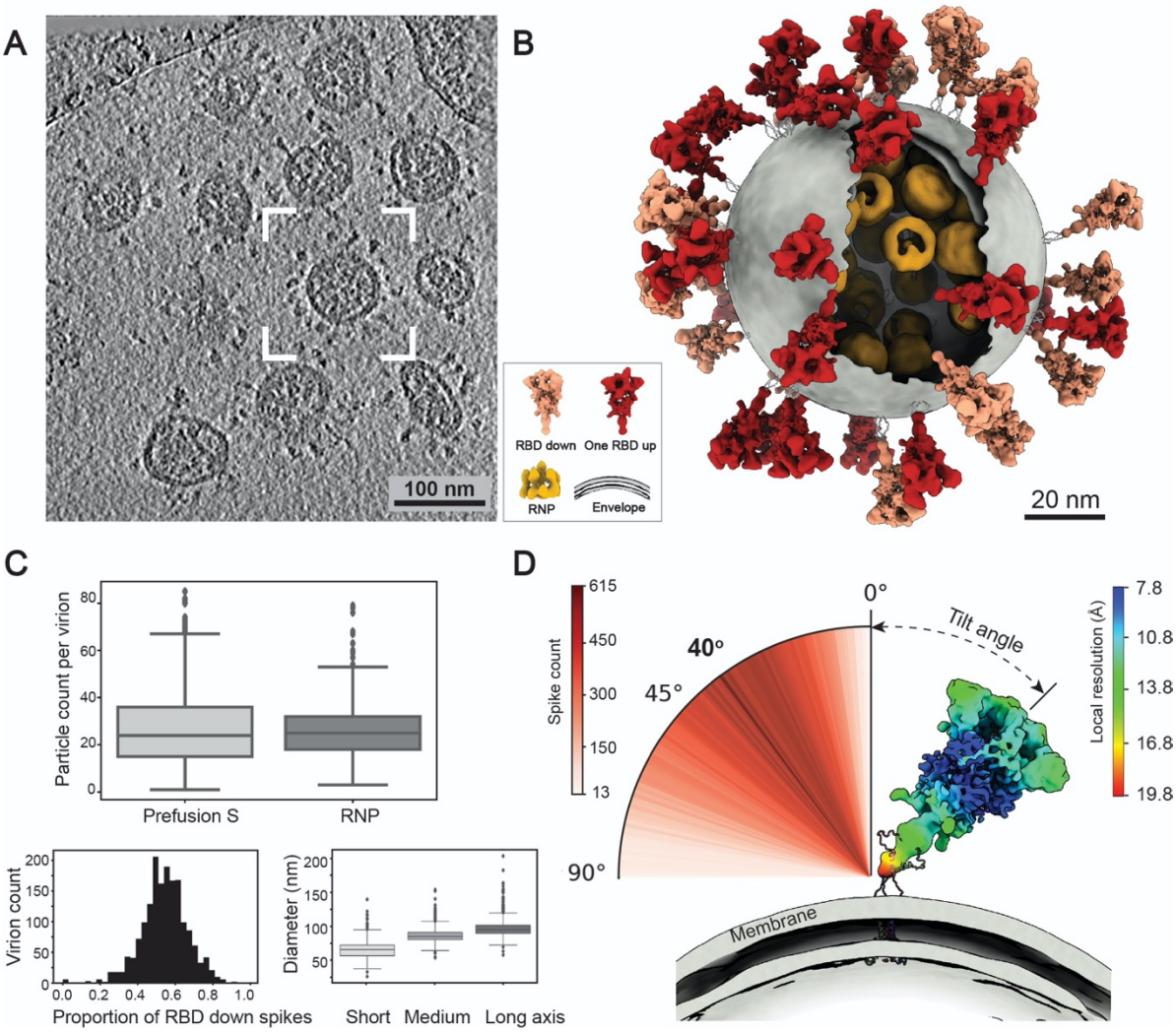

Figure 1. Molecular architecture of the SARS-CoV-2 virus. (A) A representative tomogram slice (5 Å thick) showing pleomorphic SARS-CoV-2 virions. (B) A representative virus (boxed in A) is reconstructed by projecting the prefusion S in RBD down conformation (salmon) and one RBD up conformation (red), lipid envelope (grey), and RNPs (yellow) onto their refined coordinates. RNPs away from the envelope are hidden for clarity. The unsolved stem regions of the spikes are sketched from a predicted model (https://zhanglab.ccmb.med.umich.edu/COVID-19). (C) Summary of structural observations. *Upper*: number of prefusion S and RNPs per virion; *Lower left*: ratio of the S proteins between RBD down and one RBD up conformations; *Lower right*: statistics of the dimension of SARS-CoV-2 viral envelopes. (D) Distribution of the spike tilt angle reveals a prevailing tilt of 40° relative to the normal axis of the envelope. Shown here is a representative RBD down spike in authentic orientation to the envelope (grey). The spike is colored by local resolution, and the predicted model of the stem is fitted for illustration purpose.

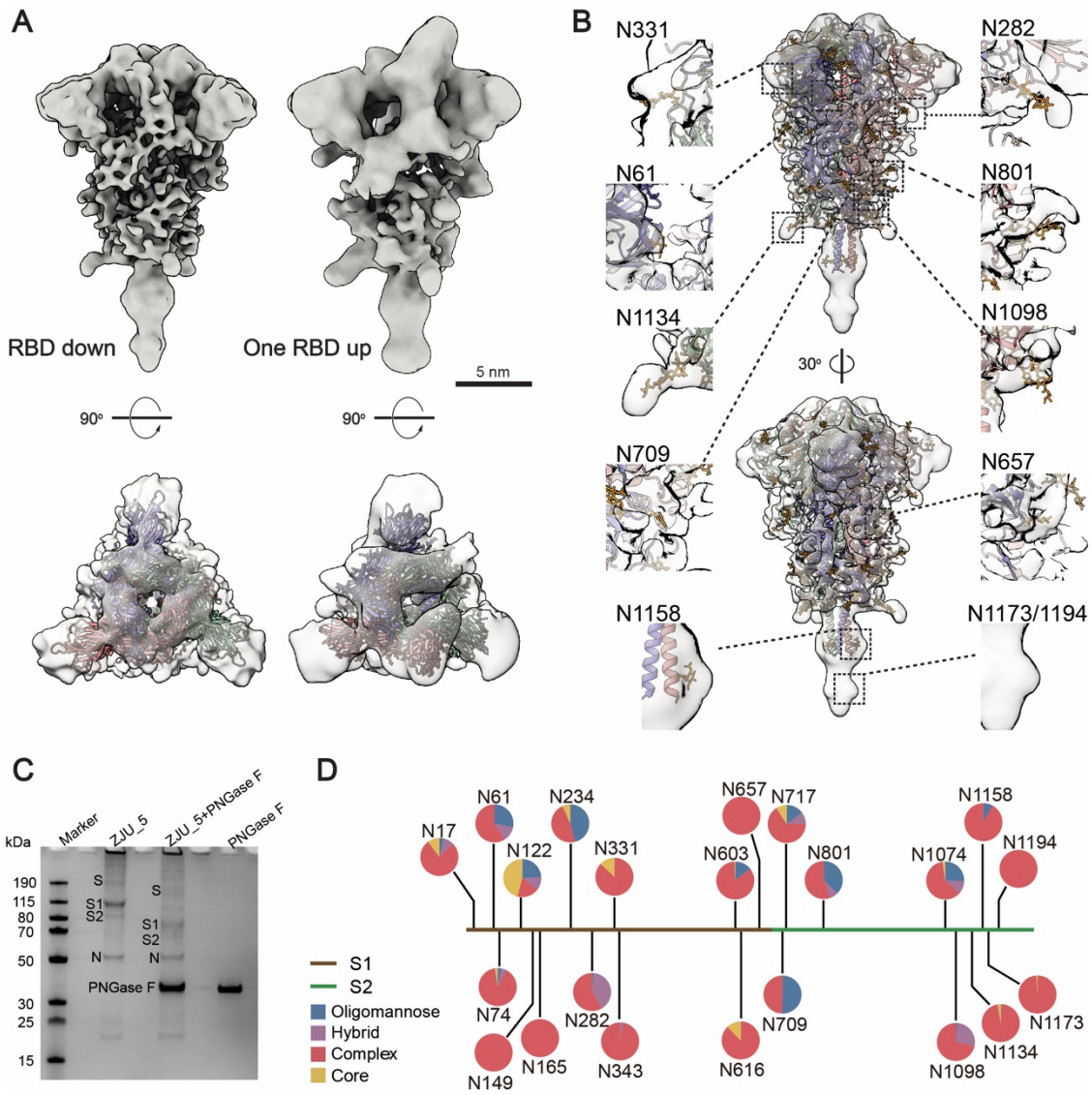

Figure 2. Native structure of the S protein in the prefusion conformation. (A) S in RBD down (resolution 8.7 Å, threshold: 1.2) and one RBD up (resolution 10.9 Å, threshold: 1.5) conformations. Sideview (upper) and topview (lower) of the maps respectively fitted with PDB entries 6XR8 and 6VYB are shown. (B) The RBD down S maps fitted with PDB entry 6XR8. Densities of ten glycans are highlighted in insets. (C) Compositional analysis of the viral sample. The treated viruses, designated ZJU_5, were resolved by SDS-PAGE and visualized by Coomassie blue staining. Lane (1) protein ladder; (2) purified ZJU_5; (3) ZJU_5 treated with PNGase F and (4) PNGase F as control. After PNGase F treatment, the molecular weight of the S1 subunit is reduced by ~30 kDa, and S2 by ~20 kDa in weight. (D) Compositional analysis of the surface glycans. The identity and proportion of 22 N-linked glycans from the native S glycans were analyzed by mass spectrometry and presented in pie charts.

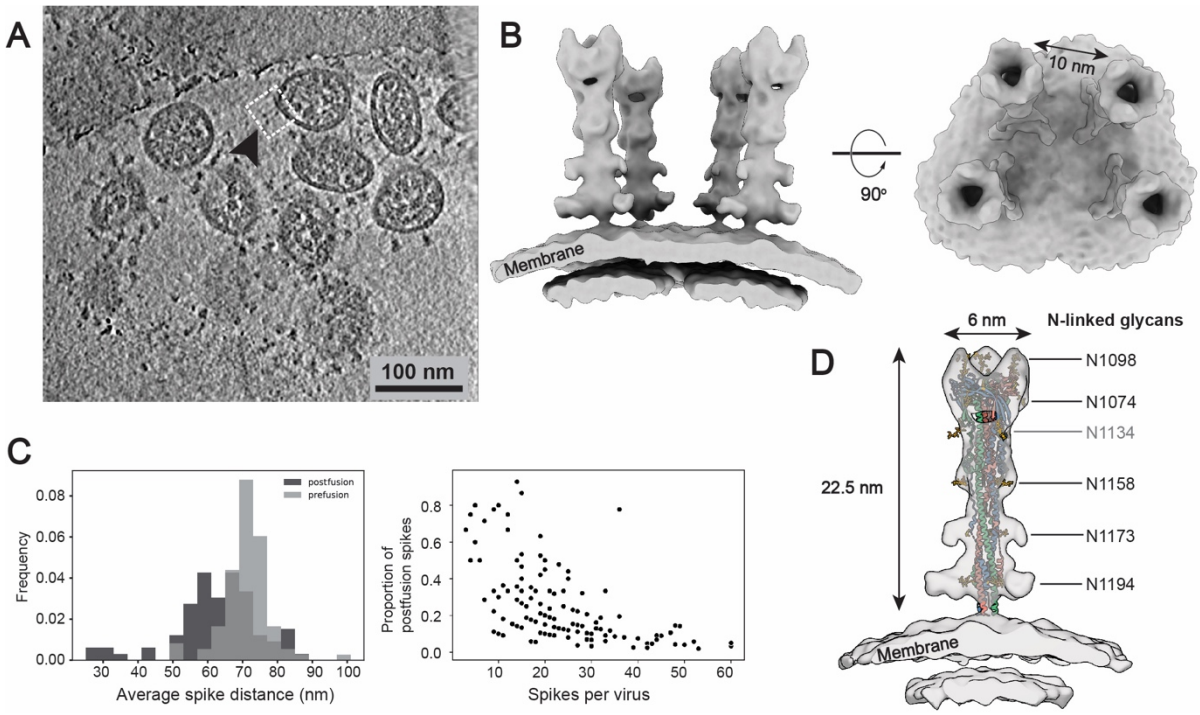

Figure 3. Native structure of S2 in the postfusion conformation. (A) A representative tomogram slice (5 nm thick) showing a cluster of spikes in postfusion conformation on a SARS-CoV-2 virion. (B) 3D reconstructions of S2 in postfusion state. The boxed region in (A) is reconstructed by projecting the solved postfusion S structure onto their refined coordinates. Resolution: 15.3 Å, threshold: 1.0. (C) Distribution of the postfusion S. *Left:* Statistics indicate that the postfusion S are closer to each other on the viral surface compared to the prefusion S; *Right:* viruses possessing less spikes in total tend to have more postfusion S (only virions possessing postfusion S were counted). (D) The postfusion S structure fitted with PDB: 6XRA, displaying densities of N1098, N1074, N1158, N1173 and N1194. Density of N1134 is not visible on the map and is colored grey.

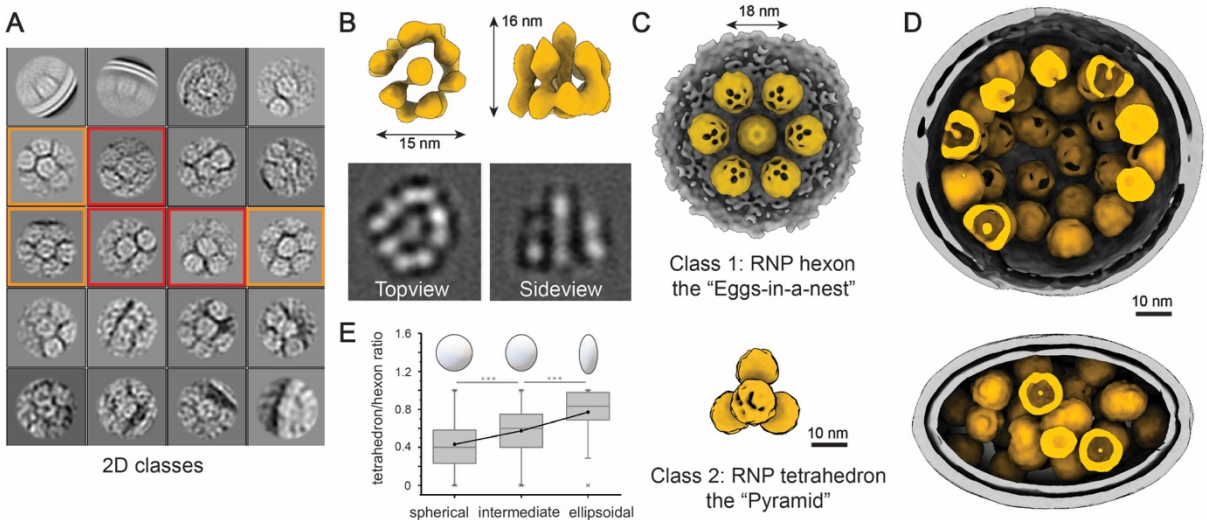

Figure 4. Native assembly of the ribonucleoproteins (RNPs). (A) 2D class averages of RNPs reveal two distinctive types of the RNP ultrastructure: hexameric (boxed in orange) and tetrahedral (boxed in red) assembly. (B) 3D reconstruction and 2D slices of the RNP. Resolution: 13.1 Å, threshold: 1. (C) Ultrastructure of the RNP hexon and tetrahedron assemblies. Seven RNPs are packed against the viral envelope (grey), forming an "eggs-in-a-nest" shaped hexagonal assembly (upper). Four RNPs are packed as a membrane-free tetrahedral assembly (lower), most of which were found inside the virus away from the envelope. Structural features of the RNPs on the assembly are smeared due to the symmetry mis-match between individual RNPs and the assembly. (D) Representative projection of the RNP hexons assembling into a spherical virus, and tetrahedrons into an ellipsoidal virus. (E) Statistics of the ratio of tetrahedron/hexon assembly reveals that the spherical and ellipsoidal virions are more likely to be packed with hexons and tetrahedrons, respectively. The ratio is estimated by sorting 382 virions that have over 5 RNP assemblies by their ratio of long/short axis and counting their ratio of tetrahedron to hexon RNPs.

Video S1: A representative raw tomogram (inverted density, lowpassed to 80 Å resolution) overlaid with the reconstruction of a SARS-CoV-2 virus. One of its RBD down S is fitted with PDB: 6XR8.