In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges

Beata Turoňová[1,2,*], Mateusz Sikora[3,*], Christoph Schürmann[4,*], Wim J. H. Hagen[1], Sonja Welsch[5], Florian E. C. Blanc[3], Sören von Bülow[3], Michael Gecht[3], Katrin Bagola[7], Cindy Hörner[4,8], Ger van Zandbergen[7,9,10], Shyamal Mosalaganti[1,2], Andre Schwarz[1], Roberto Covino[3,11], Michael D. Mühlebach[4,8], Gerhard Hummer[3,6], Jacomine Krijnse Locker[12], Martin Beck[1,2]

* contributed equally

Correspondence to gerhard.hummer@biophys.mpg.de, jacomina.krijnselocker@pei.de, martin.beck@biophys.mpg.de

1 European Molecular Biology Laboratory, Structural and Computational Biology Unit, Meyerhofstr. 1, Germany.

2 Department of Molecular Sociology, Max Planck Institute of Biophysics, Max-von-Laue Str. 3, 60438 Frankfurt am Main, Germany.

3 Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Max-von-Laue Str. 3, 60438 Frankfurt am Main, Germany.

4 Division of Veterinary Medicine, Paul Ehrlich Institute, Paul Ehrlich Strasse 51-59, 63225 Langen, Germany.

5 Central electron microscopy facility, Max Planck Institute of Biophysics, Max-von-Laue Str. 3, 60438 Frankfurt am Main, Germany.

6 Institute of Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany.

7 Division of Immunology, Paul Ehrlich Institute, Paul Ehrlich Strasse 51-59, 63225 Langen, Germany.

8 German Center for Infection Research, Gießen-Marburg-Langen, Germany

9 Institute for Immunology, University Medical Center, Johannes Gutenberg University Mainz, Mainz, Germany.

10 Research Center for Immunotherapy (FZI), University Medical Center, Johannes Gutenberg-University Mainz, Mainz, Germany.
11 Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany.
12 Electron Microscopy of Pathogens Unit, Paul Ehrlich Institute, Paul Ehrlich Strasse 51-59, 63225 Langen, Germany.

Abstract

The spike (S) protein of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is required for cell entry and is the major focus for vaccine development. We combine cryo electron tomography, subtomogram averaging and molecular dynamics simulations to structurally analyze S *in situ*. Compared to recombinant S, the viral S is more heavily glycosylated and occurs predominantly in a closed pre-fusion conformation. We show that the stalk domain of S contains three hinges that give the globular domain unexpected orientational freedom. We propose that the hinges allow S to scan the host cell surface, shielded from antibodies by an extensive glycan coat. The structure of native S contributes to our understanding of SARS-CoV-2 infection and the development of safe vaccines. The large scale tomography data set of SARS-CoV-2 used for this study is therefore sufficient to resolve structural features to below 5 Ångstrom, and is publicly available at EMPIAR-10453.

Introduction

As of June 2020, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic has reached all five continents with over >9 million confirmed cases and >460.000 SARS-related deaths. Compared to its close relatives SARS-CoV and MERS-CoV, SARS-CoV-2 appears to spread more efficiently. Without a vaccine, human distancing is currently the only effective measure to contain the pandemic, causing dramatic impact on the socio-economical landscape. The coronaviral spike surface protein (S) is the primary vaccine target, being exposed on the virus, mediating binding of the virus to the cell surface and fusion with cellular membranes required to initiate infection (*de Haan and Rottier, 2005*). S also determines tissue- and cell tropism and mutations in S may alter the host range of the virus and cross species barriers *(Hulswit et al., 2016; Li, 2016*). Vaccine efforts therefore focus on inducing cellular and humoral immunity with neutralizing antibodies that block viral binding, entry, fusion and infection by binding to S.

S is a trimer that adopts a typical club-like shape of ~20 nm in length on the surface of the virus. The ectodomain consists of a globular top and a slender stalk that connects the former to the membrane. Because S is essential for infection and the major target of vaccine development, understanding its molecular structure is of utmost importance and truncated or detergent extracted S over-expressed in mammalian cells has been subjected to various structural investigations *in vitro* by single particle cryo-EM (Wrapp et al., 2020, Walls et al., 2020, Cai et al., bioRxiv 2020, Henderson et al., bioRxiv 2020, Xiong et al., bioRxiv 2020). SARS-CoV-2 S is structurally similar to SARS-CoV S and binds to the same receptor, the angiotensin-converting enzyme 2 (ACE2) (Walls et al., 2020, Wrapp et al., 2020). The globular domain contains the receptor binding domains (RBDs) that reside on top and are shielded by the N-terminal domains (NTDs). Two conformational states have been observed for the truncated versions of S investigated *in vitro*. A closed conformation in which all three RBDs are positioned in-between the NTD and the RBD of the neighbouring monomer, thereby hiding the receptor binding site; and an open conformation in which one of the RBDs is exposed upwards, while the other two remain in the closed conformation (Wrapp et al., 2020, Walls et al., 2020). Although it has been suggested that this conformational variability might be related to receptor binding (Wrapp et al., Science 2020), *in vitro* structural analysis of detergent extracted, full-length S exclusively observed the fully closed conformation (Cai et al. bioRxiv 2020). The latter study also reported a spontaneous transition into the post-fusion state

during purification. The extent to which such spontaneous transitions occur on the viral surface remains uncertain. Further, the most C-terminal part of the protein, in particular large parts of the stalk and the entire transmembrane domain remained poorly resolved in any of the previous studies.

In some coronavirus strains, S may be cleaved at the interface between the S1 and S2 subunits by furin, a cellular protease located in late or post-Golgi compartments; the N-terminal S1 harbors the RBD, while the C-terminal S2 contains the fusion machinery and the transmembrane anchor. Although not necessary for infection, furin cleavage of S is linked to virulence *in vivo (Li 2016; Belouzard et al., 2012)*. Considerable effort has been undertaken to stabilize S for structural investigations *in vitro (Walls et al., 2017b)*. To lock S in a pre-fusion conformation the furin cleavage site was deleted and two proline residues introduced at position 986 and 987 *(Kirchdorfer et al. 2018)*. Further, the truncated ectodomain of S was fused to an artificial C-terminal trimerization tag *(Wrapp et al., Science 2020; Walls et al., 2020; Xiong et al.,* bioRxiv *2020)*. To potentially facilitate the generation of conformation specific antibodies, dedicated constructs of S have been engineered that lock the RBDs into either the open or closed conformation *(Henderson et al., bioRxiv 2020, Xiong etal., bioRxiv 2020)*. To what extent these conformational states are representative of the *in situ* scenario on the viral surface remains unknown.

The SARS CoV2 S is a type I membrane glycoprotein with an N-terminal signal sequence for synthesis on, and translocation across the endoplasmic reticulum (ER), where it trimerizes and becomes glycosylated at its 22 predicted N-linked glycosylation sites. S protein transport through the Golgi apparatus and further processing of its glycosylation is thought to occur once incorporated into virions that bud at pre-Golgi membranes and exit infected cells by vesicular transport (*de Haan and Rottier, 2005*). Previous *in vitro* cryoEM studies resolved roughly two thirds of the predicted 22 N-linked glycans (Walls et al., 2020; Wrapp et al., 2020). Although an N-terminal signal peptide for secretion was fused to S, it is not clear if over-expression of the protein alone in HEK 293F cells (Walls et al., 2020) ultimately led to the same glycosylation pattern as compared to viral assembly that massively impacts onto various organellular systems (Gordon et al., 2020; Bojkova et al., 2020; Ulasli et al., 2010).

While these previous studies provided valuable data of isolated spike protein or its fragments, *in situ* structural analysis of SARS-CoV-2 remains preliminary. Structural information about S *in situ* is a prerequisite towards understanding the molecular mechanisms

of host cell entry. It would allow to explore the conformational variability within the virion and the accessibility of epitopes for neutralizing antibodies; to determine the glycosylation pattern of S generated during viral assembly; and to analyze the yet unresolved stalk domain that flexibly connects S to the viral surface. Here we combined cryo electron tomography of purified SARS-CoV-2 virions with subtomogram averaging, molecular dynamics simulations and integrative structural modeling to address these challenges.

Results

Large scale tomographic analysis of SARS-CoV-2 virions

To structurally analyze SARS-CoV-2 spike protein *in situ*, we passaged the virus through tissue culture cells and purified it from the inactivated supernatant by sucrose centrifugation (see methods for detail). The origin of the SARS-CoV-2 isolate used in this study is south Germany and described in detail in the supplement. We acquired a large scale cryo electron tomography data set of purified virions using acquisition parameters that are suitable for subtomogram averaging (see methods). Visual inspection of the tomographic reconstructions revealed a very high quality data set in which, individual protein domains are clearly visible (Figure 1A, Movie S1). The raw data of 266 tilt-series with >1000 viruses have been deposited EMPIAR-10453 and might be used in the future to explore various structural properties of SARS-CoV-2, in particular its pleiomorphic features. In this initial study we will exemplify the potential of our data as a resource with the *in situ* structural analysis of S.

The pre-fusion conformation of spike dominates the surface of virions purified from tissue culture cells

Visual inspection of viral surface revealed that spike was mostly present in the pre-fusion conformation, while some notable exceptions appeared strikingly similar to previous analyses of an alternative conformation (Figure 1B) thought to be the post-fusion conformation (Cai et al bioRxiv 2020; Duquerroy *et al.*, 2005). We hypothesized that the pre-fusion conformation might have been positively selected during passage through tissue culture cells, in the virions that we purified from the supernatant. We sequenced the viral genome at passages 1 to 5, the latter of which was used for the preparation of cryo-EM samples. The furin proteolytic cleavage site was gradually lost during passage through tissue

culture as shown by Sanger sequencing and western blot analysis on the protein level confirming previous studies (Figure 1C, Figure S1; Lau et al., bioRxiv 2020; Ogando et al., bioRxiv 2020). It has been previously noted that due to the location of the furin site within a loop, its cleavage may have little impact on the S trimer in the pre-fusion conformation, other than ultimately allowing for the dissociation of the S1 fragment (Cai et al., bioRxiv 2020). A study that appeared on bioRxiv during the preparation of this manuscript suggests that in virions produced in Vero E6 cells the pre-fusion is predominant while this may be different for other cell types (Klein et al., bioRxiv 2020), thus underscoring that our large scale data set is representative of the native state.

We manually picked the virions and oversampled their exterior to extract subtomograms (see methods for detail). After a few iterations of alignment, they converged into a low-resolution structure of the globular domain of spike in its pre-fusion conformation. Statistical analysis of the resulting positions revealed that on average 40 copies of S trimer reside on the surface, although their density varies across virions. Individual spike complexes appeared randomly distributed on the viral surface and did not show any significant tendency to cluster but rather had slightly repellent properties (Figure 2A).

We refined the globular domain to 7.6 Ångstrom using procedures implemented in NovaSTA (https://github.com/turonova/novaSTA) and STOPGAP packages (https://github.com/williamnwan/STOPGAP, see methods for detail). At this resolution, secondary structure elements and individual glycosylation sites are clearly discernible (Figure 2B, C). Although there was no symmetry imposed during the refinement, the average was highly symmetric, suggesting that conformational variability across the three different monomers is minimal. Structural analysis of the asymmetric unit yielded an average map from 65780 subtomograms at an overall 4.9 Ångstrom resolution. In particular the cluster of parallel helices in the center of the globular domain was nicely resolved (Figure 2D) and flexible fitting revealed that larger side chains are clearly discernible (Figure S3). Simulated annealing-based classification as implemented in STOPGAP of both, the S-trimer but also the asymmetric unit did not yield any distinct classes. However, separate cross-correlation against the single particle structures with either the closed or open RBD identified a fraction of S-trimers in the open conformation (Figure S4).

An unusual coiled-coil resides next to a hinge in the stalk domain of spike

While the globular domain was fully contained in the tomographic map, only the directly adjacent part of the stalk domain was resolved, emphasizing the flexibility of the latter. Remarkably, a right-handed coiled coil emerges from the neck of the spike head (~P1140-L1159), forming the upper part of the stalk domain (subsequently referred to as the "upper leg" of the stalk). Right-handed trimeric coiled coils were long thought to be absent from the structural proteome (Harbury *et al.*, 1998), but can be seen in the post-fusion structure of the related mouse hepatitis virus (MHV) spike (PDB ID 6B30; Walls *et al., 2017a*). The density in the single-particle structures of SARS-CoV-2 spike (Wrapp *et al.*, 2020), albeit only moderately resolved in this region next to the trimerization tag, is consistent with a right-handed coiled-coil. The assignment as right-handed coiled-coil is also consistent with the observation that this sequence stretch is not part of the post-fusion left-handed coiled-coil, as resolved for SARS-CoV spike (PDB ID 1WYY; Duquerroy *et al.*, 2005), which indicates a certain incompatibility with left-handedness.

The observed electron density sharply declined at a defined position within the trimeric coiled-coil, suggesting at least one flexible hinge (Figure 2B). We analyzed the conformational variability of the stalk domains based on the position of the globular domain and its orientation with respect to the membrane as defined by subtomogram averaging. The globular domain exhibited large positional and orientational freedom, tilting up to ~90 degrees with respect to the normal at distances of 5-35 nm from the membrane without discernible correlation (Figure 2E). We grouped our subtomograms into four classes according to their distance from the bilayer and averaged them separately. In most classes the stalk remained invisible and the density of the bilayer was blurred, suggesting a highly kinked conformation of the stalk (Figure 2F). At an intermediate distance, however, parts of the stalk and bilayer were resolved, suggesting a more defined conformation. We sub-selected ~3200 particles in which the globular domain was oriented roughly perpendicular to the membrane and averaged them separately (see methods). In the resulting average, the stalk domain was resolved (Figure S5A). Visual inspection of the respective subtomograms, in which the stalk domains are clearly observed, further corroborated the idea of a kinked stalk with potentially several hinges (Figure 2F). We used local masking to focus the subtomogram averaging on the lower part of the stalk domain (subsequently referred to as the lower leg). We obtained a

moderately resolved structure that would be consistent with the continuation of the coiled coil below a flexible hinge (subsequently referred to as the "knee", Figure S5B).

Molecular dynamics simulations of spike in the context of the bilayer

We performed a 2.8 microseconds long all-atom molecular dynamics (MD) simulations of a 4.1-million atom system containing four glycosylated S proteins anchored into a patch of viral membrane and embedded in aqueous solvent (Figure 3A). The simulations helped us pinpoint the molecular origins of the flexibility seen in the tomograms. Whereas the S head remained stable, the stalk exhibited pronounced hinging motions at the junctions between S head and upper leg ("hip"), between upper and lower leg ("knee") and between lower leg and transmembrane domain ("ankle"), consistent with pronounced leg segments seen in raw tomograms (Figure 3B-C). The hip joint flexed the least (16.5 +/-8.8 deg), followed by the ankle (23.0 +/- 11.7 deg) and knee (28.4 +/- 10.2 deg) (Figure 3D).

Hinges predicted by MD simulations are consistent with the tomographic data

The 3D tomographic density of S proteins protruding from the viral surface is well described by selected MD snapshots even without flexible fitting (Figure 4A). Strikingly, hip, knee and ankle can be pinpointed in individual copies of S seen in the tomographic reconstructions (Figure 4A). We flexibly fitted suitable snapshots of the MD simulations into subtomogram averages classified according to the distance of the globular domain from the membrane (compare Figure 2F to Figure 4B). In the stalk region, hinge bending gives S the required flexibility. The hinges facilitate a lateral displacement of the upper and lower leg with respect to the globular domain, if the latter is positioned closer to the membrane. As a result the stalk is diluted out in subtomogram averages focused on the globular domain (Figure 2B-C, F, 4B) but can be retrieved if the globular domains are aligned with the membrane normal (Figure S5A) or if the stalk itself is averaged separately (Figure S5B). We calculated electron density averaged over the whole MD trajectory and filtered it to a resolution comparable to the classified subtomograms. We obtained a highly similar 3D map that underscores this interpretation (Figure 4B). In rare cases, the coiled-coil near the membrane appears to be unfolded in the original tomograms (Figure S5C) and continuous with the disordered loops of the MD model.

N-glycosylation sites are very pronounced in tomographic data

The predicted N-glycosylation sites, many already annotated in single particle EM maps (Walls et al., 2020), are generally very pronounced in the subtomogram averages. Aided by the MD model, we systematically examined glycosylation of S. The electron density of N-glycans averaged over the MD trajectory was highly consistent with the tomographic map (Figure 5A). Clustered glycosylation sites are visible in the raw electron density before averaging, e.g. protruding from the lower part of the S body (Figure 5B). Analysis of individual sites in subtomogram averages further supports the decoration of spikes with rather large glycan chains (Figure 5C). Notably, a number of sequons were resolved with more pronounced branching than previously reported in the single particle maps (Walls et al., 2016). By contrast, the density map in the region of two previously reported O-glycosylation sites (Shajahan et al., 2020) does not allow for unambiguous assignment (Figure S6A), suggestive of low glycan occupancy at these sites *in situ*. Sequon N17LT, due to its location on the unstructured N-terminus, was not localised in the density (Figure S6B). Elongated features protruding from the tip of the N-terminal domain (Figure S6B) may suggest presence of sequons N74GT and N149KS, but are less pronounced as compared to the N-glycosylation sites discussed above, likely because they are located on highly flexible loops. All these sites were recently shown to be in a close proximity to an antibody bound to the N-terminal domain (Chi et al., 2020).

Integrative analysis suggests that the hinges are heavily N-glycosylated

N-glycosylation is predicted also on the knee (N1158HT, N1173AS) and the ankle (N1194ES) outside the region resolved by single particle EM (Figure 3A). We observed that these positions generally appear more bulky in tomographic reconstructions than one might expect if they were not glycosylated (Figures 1B and 5D). In subtomogram averages of the globular domain and stalk, additional density is very clearly observed at the respective positions (Figure 5E, F and Figure S5A, B), strongly suggesting that they are modified. Consistent features are present in the electron density calculated from the MD trajectory aligned on the lower leg (Figure S6C), where we can unequivocally attribute them to glycans. N-glycosylation in this position might protect the functionally important hinges from antibody binding and help to keep them flexible.

Discussion

The two major structural analysis techniques combined in this study are complementary. While cryo electron tomography facilitates the experimental analysis of cryo fixed specimen at moderate resolution in their native context, molecular dynamics simulations reveal protein motions at atomic resolution as a function of time. In the particular case of the SARS-CoV-2 spike protein, the tomographic data reveal its conformational freedom and glycosylation sites on the viral surface but they cannot resolve the flexible stalk domain to high detail. The molecular dynamics simulations integrate various experimental parameters, in particular single-particle structures and independently determined glycosylation sites, and allow fully atomic models to evolve in time.

Our MD simulations revealed three highly flexible hinges within the stalk, coined hip, knee and ankle, which are consistent with the experimental data. One might speculate that the high degree of conformational freedom of the spike on the viral surface is important for the mechanical robustness of the virus. It might also allow the spike to engage the relatively flat surface of host cells with better avidity prior to fusion or find access to its receptor more efficiently in the crowded environment of tissue *in vivo* (Figure 5G-H). This would allow the virus to explore the surface of host cells more effectively and potentially to bind to multiple copies of ACE2. Tomographic studies of actual infection events might further address this in the future.

In contrast to the pre-fusion conformation of spike, the post-fusion conformation previously observed *in vitro* and *in situ* (Cai et al., bioRxiv 2020 Klein et al., bioRxiv 2020) and in this study (Figure 1B) is entirely straight and apparently inflexible. A particularly unusual feature masked at the edge of the resolved density of the single-particle structures but well resolved in the subtomogram averages is the short right-handed coiled-coil at the top of the pre-fusion stalk. Being lost in the post-fusion structure as resolved for SARS-CoV (PDB ID 1WYY; Duquerroy et al. 2005), we speculate that the right-handed coiled-coil is only marginally stable, priming the protein for a large structural reorganization in a spring-loaded viral fusion mechanism. These findings underline that all three hinges become irrelevant during the transition to the post-fusion conformation, which places them outside the structural core (Cai et al., bioRxiv 2020; Duquerroy et al., 2005).

A remarkable difference between SARS-CoV and SARS-CoV-2 spike is the presence of a furin cleavage site in the latter (Wrapp et al., 2020). This site is rapidly lost by passaging in cell culture as shown before (Lau et al., bioRxiv 2020; Ogando et al., bioRxiv 2020) likely because its loss favours virions with stable spikes. In Vero cells this results in the release of SARS-CoV-2 virions with S in a closed formation (this study; Klein et al., bioRxiv 2020). In the latter recent study the authors proposed that S released on SARS-CoV-2 particles is readily transformed into a conformation resembling the post-fusion protein when exposed to cells that express high levels of ACE2. Whether this conformational change relates to receptor binding, to cleavage at the S2' site by the serine protease TMPRSS2 (Belouzard et al., 2012) or to both was not addressed in this study. In addition, it was not clear if the furin cleavage site is intact in the unpassaged isolate used in this study and, if so, to what extent S was cleaved during passage of newly assembled virus through the secretory pathway. In Vero cells we found for instance cleavage of S containing the furin cleavage site to be partial, with the majority of S remained unprocessed, suggesting that efficiency of furin cleavage may be cell type dependent. Nonetheless, the observations made in this study are interesting and open the possibility to compare the pre-fusion with the presumed post-fusion structure of S *in situ*.

The predominant *in situ* conformation of S observed in our study is the closed pre-fusion conformation. This finding emphasizes that the highly-engineered, recombinant versions of S in which this conformation is locked (Henderson et al., bioRxiv 2020, Xiong bioRxiv 2020) may indeed be valuable tools for vaccine development, although there are also differences to the *in situ* structure. N-glycosylation sites appear very bulky in the tomographic map as compared to previous single particle analysis, suggesting that decoration with sugars may indeed be more extensive during viral assembly as compared to the recombinant ectodomain modified during default vesicular transport. Our map is suggestive of additional N-glycosylation at the hinges of the stalk domain and possibly on the very tips of S NTDs. This might be important for the accessibility of epitopes on the crowded viral surface were the NTD and stalk domains appear occluded by neighbouring spikes (Figure 5G). Indeed, glycan shielding of epitopes on the coronavirus spike protein has been described before (Walls et al., 2016) emphasizing the importance of identifying the glycosylation pattern on the native spike

protein. We did not detect any density for the predicted O-glycosylation sites emphasizing that it must be less extensive as compared to N-glycosylation.

Here we demonstrate the potential of our large scale tomographic data set for resolving structural features of SARS-CoV-2 particles to high resolution in their native context. The *in situ* structure of several key viral components, including e.g. the M protein that is highly enriched in the membrane and the nucleocapsid remain enigmatic. Our data might thus be explored to resolve the structures of such features in the future. We further demonstrate that high resolution structural models can be fitted directly into the tomographic reconstructions, underlining the remarkable quality of the data. This strategy might be further explored in the future to build structural models of entire virions.

Depositions

The original tilt series have been deposited (EMPIAR-10453). Subtomogram averages were deposited into the electron microscopy data base under accession numbers EMD-11222 (S-trimer), EMD-11223 (asymmetric unit with closed RBD), and EMD-11347 (asymmetric unit with open RBD).

Author contributions

B.T. experimental design, tomographic reconstruction, particle picking, subtomogram averaging, structural analysis, paper writing. M.S. molecular dynamics simulations, structural analysis, paper writing. C.S. experimental design, virus purification and biochemical analysis. W.H. experimental design, cryo-EM data acquisition, tomographic reconstruction. S.W. experimental design, sample preparation and screening, data analysis. F.E.C.B., S.v.B., M.G. and R.C. molecular dynamics simulations, structural analysis. K.B. experimental design and virus purification. C.H. experimental design and virus growth. GvZ experimental design, supervision. S.M. subtomogram averaging. A.S. tomographic reconstruction, particle picking. M.D.M. experimental design, supervision. G.H. data analysis, supervision, paper writing. J.K.L. experimental design, supervision, paper writing. M.B. experimental design, supervision, paper writing.

Acknowledgments

The cryo electron tomography data was collected at the EMBL Heidelberg Cryo Electron Microscopy Service Platform. We thank EMBL (BT, WH, SM, AS, MB), the Max Planck Society (BT, MS, SW, FECB, SvB, MG, SM, RC, GH, MB), the Max Planck Computing and Data Facility for providing computational resources. We also acknowledge a generous SuperMUC-NG computing allocation at the Leibniz Supercomputing Centre. (MS,SvB,MG,FECB,RC,GH), the Human Frontier Science Program (RGP0026/2017; SvB, GH), the German Ministry of Health (CS), the German Center for Infection Research (CH,MDM) and the Loewe centre DRUID from the Justus Liebig university Giessen (JKL) for funding. BT acknowledges William Wan (Vanderbilt University) for helpful discussions. JKL acknowledges excellent support by Regina Eberle (PEI). RC acknowledges the support of the Frankfurt Institute for Advanced Studies. MS acknowledges support from the Austrian Science Fund *FWF* (*Schroedinger* Fellowship, J4332-B28). The authors are indebted to G. Dobler and R. Wölfel, Bundeswehr Institute for Microbiology, for providing SARS-CoV-2 strain MUC-IMB1. The authors declare no competing interests.

Figures

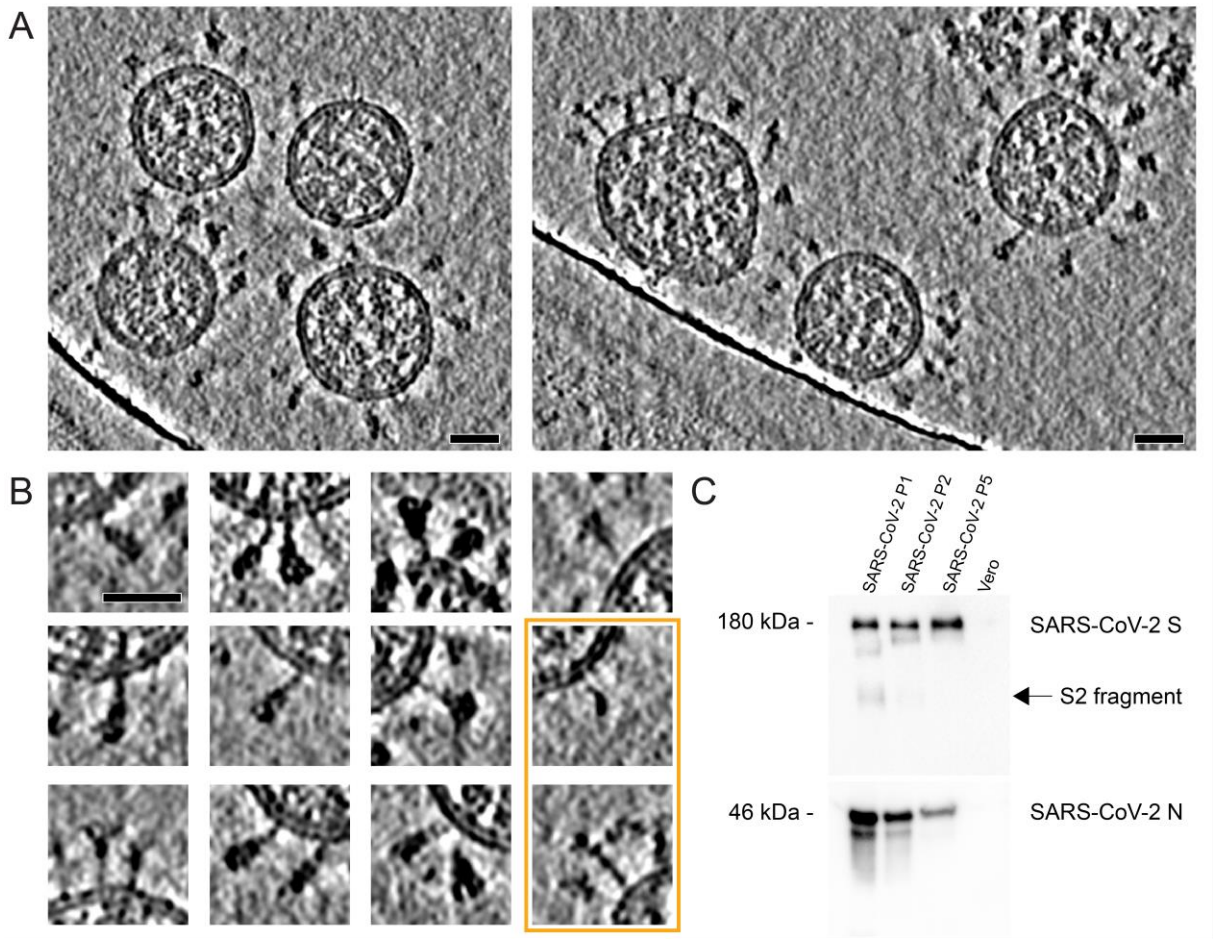

Figure 1 - Cryo electron tomography of SARS-CoV-2 virions. (A) Slices through tomographic reconstructions of highly pleiomorphic SARS-CoV-2 virions. Scale bar 30 nm. **(B)** Same as (A), but as a gallery highlighting individual copies of S. All domains, including the transmembrane part, are clearly resolved. Although the vast majority of S is reminiscent of the pre-fusion conformation, the particles framed in orange are strikingly similar to the post-fusion conformation as proposed by (Cai et al, bioRxiv 2020; compare to their Figure 4C). **(C)** Western blot showing the loss of furin cleavage products of SARS-CoV-2 S and the increase of uncleaved S (180 kDa) within 5 passages through tissue culture (loading control using anti-N antibody).

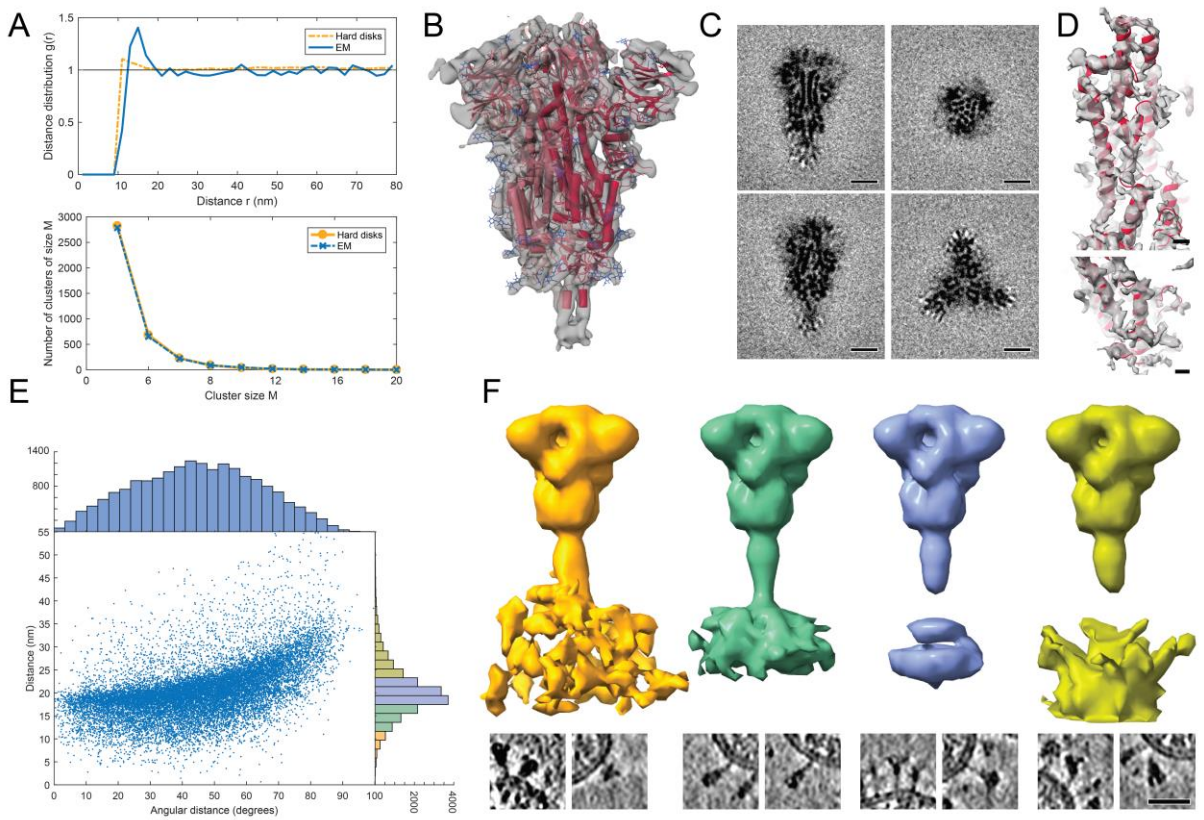

Figure 2 - Subtomogram analysis of SARS-CoV-2 S protein. (A) Distance (top) and cluster-size distributions (bottom) of S on the viral surface, with non-overlapping hard disks of 10 nm diameter as reference. **(B)** Subtomogram average of the ectodomain of S shown isosurface rendered and fitted with the previously published atomic model as determined by single particle EM (PDB ID 6VXX). Subtomogram average in transparent grey, secondary structure elements in red and glycosylation sites in blue. **(C)** Same subtomogram average but shown as slices through the reconstruction. Scale bar 5 nm. **(D)** Detail of the average of the symmetric unit of S. Scale bar 5 Å. **(E)** Distribution of the angular orientation and distance of the ectodomain with respect to the bilayer. **(F)** Averages of subtomograms classified according to the distance of the ectodomain (increasing from left to right) from the bilayer. Examples of individual particles are shown as slices at the bottom panel (scale bar 30 nm). At an optimal distance, the stalk domain stretches out and is resolved.

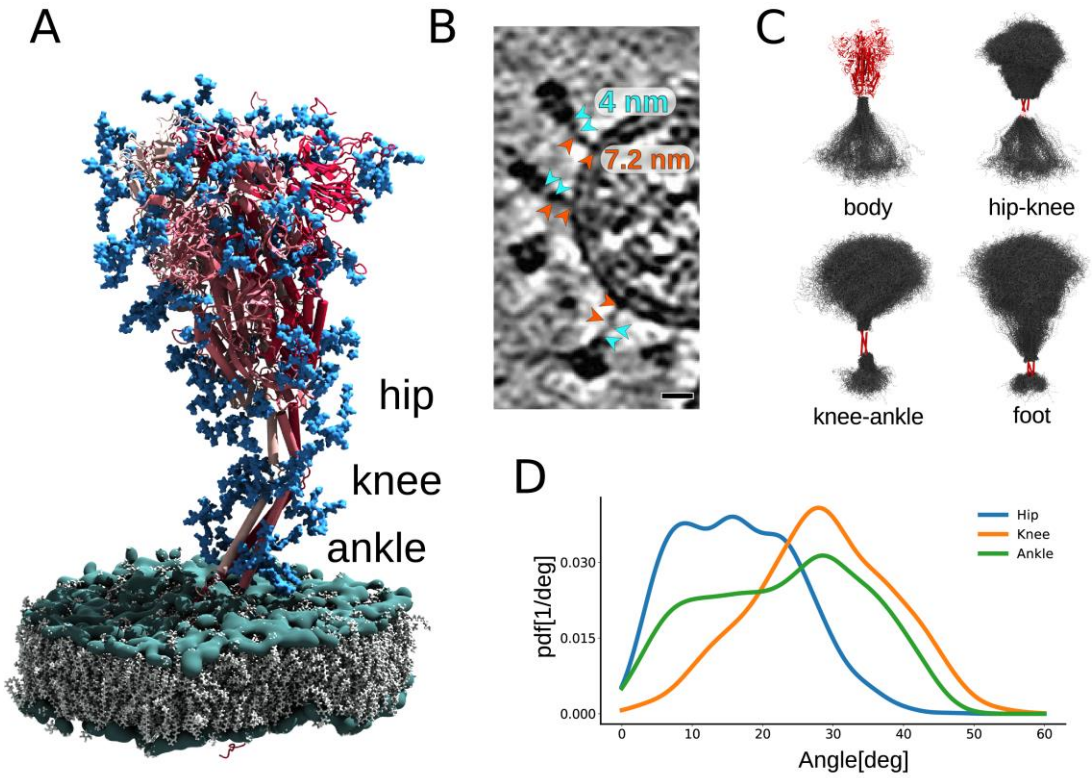

Figure 3 - MD simulations of SARS-CoV-2 S protein. (A) Model of the S protein. The three individual chains of S are shown in shades of red, N-glycosylation in blue, lipids of the ER-like membrane in grey with phosphates in green; "hip", "knee" and "ankle" mark positions of the three flexible hinges. **(B)** Examples of the hinges as seen in the de-convoluted tomograms. Cyan and orange arrowheads indicate upper and lower leg respectively, with their typical lengths indicated. Scale bar 10 nm. **(C)** Hinge flexibility in the MD simulation illustrated through backbone traces (grey) at 75 ns intervals with different parts of the S protein fixed (red). **(D)** Probability density functions for hinge bending angles at hip, knee and ankle.

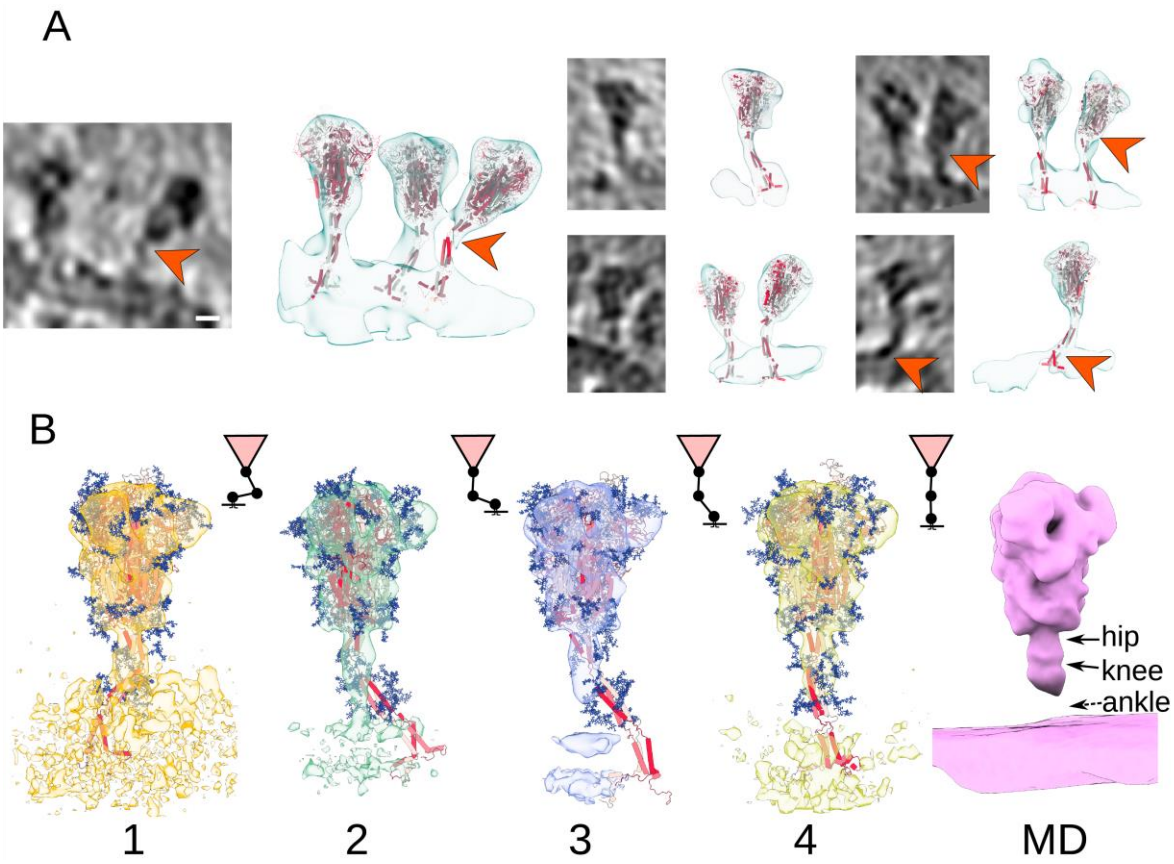

Figure 4 - Fitting of molecular simulations into cryo electron tomograms. (A) The hinges of the stalk domain predicted by structural modeling (orange arrowheads) are consistent with the tomographic data. Slices through tomograms (left) and snapshots of respective MD simulations superimposed with isosurface rendered tomograms are shown for multiple instances. Scale bar 5 nm. **(B)** Fit of snapshots of MD simulations into the classes obtained for different distances of the globular domain from the membrane (1-4) as presented in Figure 2F. Shorter distances are concomitant with a stronger bending of the hinges and a lateral displacement of the stalk. Average MD density filtered to a resolution comparable to the subtomogram averages is shown as isosurface render (right).

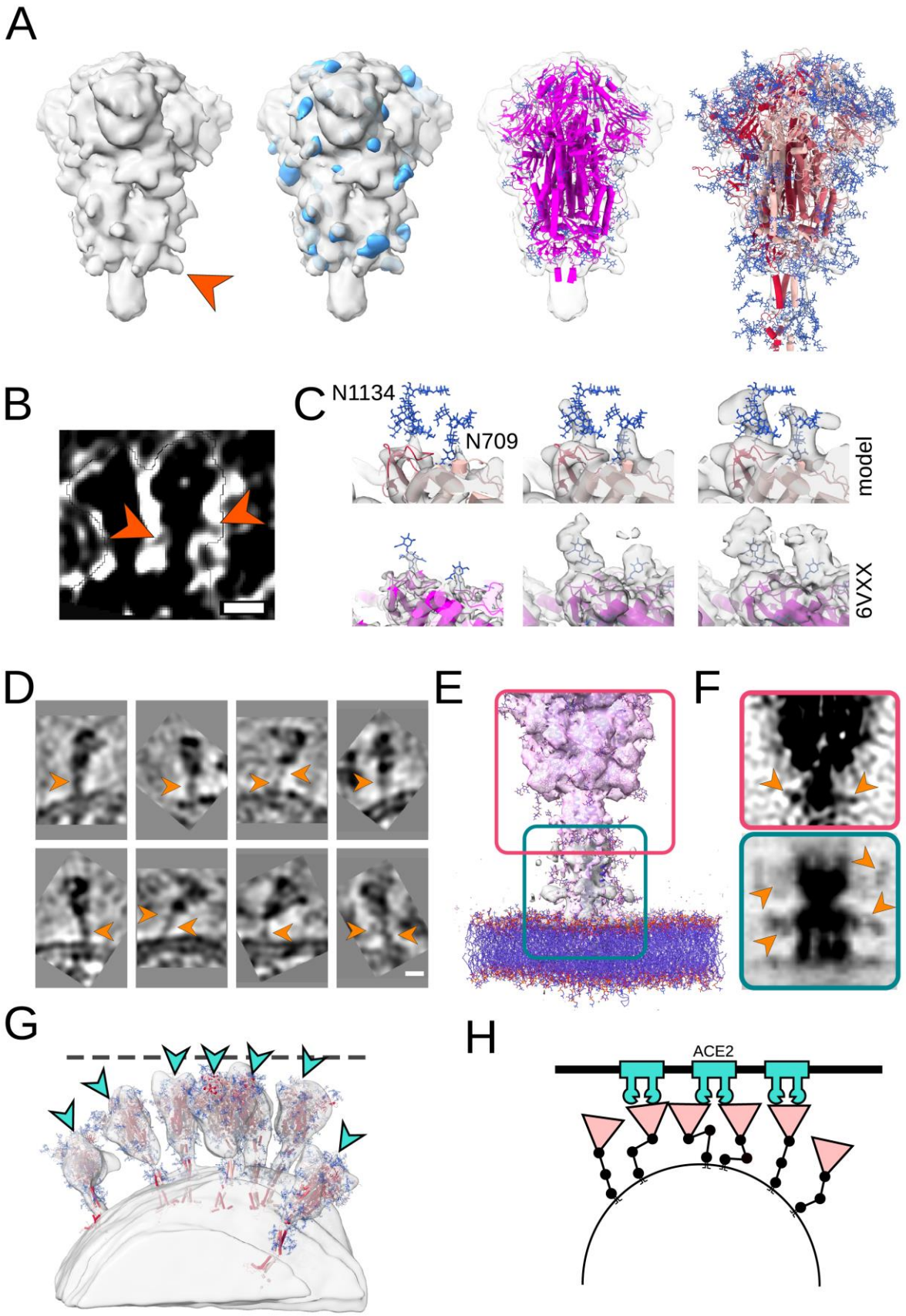

Figure 5 - Analysis of S protein glycosylation sites and epitopes. (A) N-glycosylation sites are clearly discernible in the subtomogram average of the globular domain. From left to right:

Isosurface rendering of subtomogram average with an individual N-glycosylation site indicated (orange arrowhead); same superimposed with the MD-calculated density for all annotated N-glycosylation sites; superimposed with previous structural model of the globular domain (PDB ID 6VXX); superimposed with a snapshot of the MD simulations. N-glycosylation sites are shown in blue. **(B)** Tomographic slice highlighting an N-glycosylation site (orange arrow heads) in the original data. Scale bar 5 nm. **(C)** Highlight of N-glycosylation positions 709 and 1134 of the MD simulations (top) and in a previous structural model (bottom; PDB ID 6VXX, EMD-21452). The subtomogram average is shown superimposed at different isosurface thresholds (transparent grey). Extensive additional density is visible. **(D-F)** The stalk domain is heavily glycosylated at the hinges. **(D)** Exemplifying tomographic slices with bulky density at the hinge positions (orange arrowheads). Scale bar 5 nm. **(E)** Superposition of the subtomogram averages (transparent grey isosurfaces) of the globular domain (framed red) and the stalk domain (framed green) with a respective snapshot of the MD simulations emphasizing the glycosylation at the hinges. **(F)** Same as (E) but shown as maximum intensity projection through the subtomogram averages. **(G)** Fits of snapshots from MD simulations into the surface of a virion; tomogram shown isosurface rendered in transparent grey. The position of epitopes for neutralizing antibodies at the RBDs are indicated with cyan arrowheads. **(H)** Cartoon illustrating a hypothetical docking event in which the hinges facilitate the engagement of multiple instances of S with their receptors.

Materials and Methods: see Supplementary Material - separate document

References

Belouzard, S., Millet, J.K., Licitra, B.N., and Whittaker, G.R. (2012). Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. Viruses *4*, 1011-1033.

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., Cinatl, J., and Münch, C. (2020). Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. Nature in press.

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. eabc6952.

Chuang Liu, Yang Yang, 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, L.L. (2020). Viral Architecture of SARS-CoV-2 with Post-Fusion Spike Revealed by Cryo-EM. bioRxiv.

de Haan, C.A.M., and Rottier, P.J.M. (2005). Molecular Interactions in the Assembly of Coronaviruses. In Advances in Virus Research, pp. 165-230.

Duquerroy, S., Vigouroux, A., Rottier, P.J.M., Rey, F.A., and Jan Bosch, B. (2005). Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein. Virology *335*, 276-285.

Forster, F., Medalia, O., Zauberman, N., Baumeister, W., and Fass, D. (2005). Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. Proc. Natl. Acad. Sci. *102*, 4729-4734.

Funke, S., Maisner, A., Mühlebach, M.D., Koehl, U., Grez, M., Cattaneo, R., Cichutek, K., and Buchholz, C.J. (2008). Targeted Cell Entry of Lentiviral Vectors. Mol. Ther. *16*, 1427-1436.

Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V. V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature, in press.

Grant, T., and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. Elife *4*.

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.

Harbury, P.B., Plecs, J.J., Tidor, B., Alber, T., Kim, P.S. 1998. High-Resolution Protein Design with Backbone Freedom. Science 282 (5393), 1462-1467. DOI: 10.1126/science.282.5393.146

Hubert, J. (1984). Spearman-Karber method. Bioassay *2nd editio*.

Hulswit, R.J.G., de Haan, C.A.M., and Bosch, B.-J. (2016). Coronavirus Spike Protein and Tropism Changes. In Advances in Virus Research, pp. 29-57.
Kirchdoerfer, R.N., Cottrell, C.A., Wang, N., Pallesen, J., Yassine, H.M., Turner, H.L., Corbett, K.S., Graham, B.S., McLellan, J.S., and Ward, A.B. (2016). Pre-fusion structure of a human coronavirus spike protein. Nature *531*, 118-121.

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.

Lau, S.-Y., Wang, P., Mok, B.W.-Y., Zhang, A.J., Chu, H., Lee, A.C.-Y., Deng, S., Chen, P., Chan, K.-H., Song, W., et al. (2020). Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. Emerg. Microbes Infect. *9*, 837-842.

Li, F. (2016). Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu. Rev. Virol. *3*, 237-261.

Loganathan, S.K., Schleicher, K., Malik, A., Quevedo, R., Langille, E., Teng, K., Oh, R.H., Rathod, B., Tsai, R., Samavarchi-Tehrani, P., et al. (2020). Rare driver mutations in head and neck squamous cell carcinomas converge on NOTCH signaling. Science (1264-1269). *367*, 1264-1269.

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

Natacha S. Ogando, Tim J. Dalebout, Jessika C. Zevenhoven-Dobbe, Ronald W. Limpens, Yvonne van der Meer, Leon Caly, Julian Druce, Jutte J. C. de Vries, Marjolein Kikkert, Montserrat Bárcena, Igor Sidorov, E.J.S. (2020). SARS-coronavirus-2 replication in Vero E6 cells: replication kinetics, rapid adaptation and cytopathology. bioRxiv.

Otsu, N.A. (1979). A threshold selection method from gray-level histograms. IEEE Trans. Sys. Man. Cyber *9*, 62-66.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216-221.

Rory Henderson, Robert J Edwards, Katayoun Mansouri, Katarzyna Janowska, Victoria Stalls, Sophie Gobeil, Megan Kopp, Allen Hsu, Mario Borgnia, Rob Parks, Barton F. Haynes, P.A. (2020). Controlling the SARS-CoV-2 Spike Glycoprotein Conformation. bioRxiv.

Sebastian Lewis, Deborah Andrew D. Davidson, M.K.W., Joanna Ellis, Phillip Shoemark, Miles W. Carroll, Kate Heesom, M.Z., and Hiscox, David A. Matthews A. Lewis, J.A. (2020). Characterisation of the transcriptome and proteome of SARS-CoV-2 using direct RNA sequencing and tandem mass spectrometry reveals evidence for a cell passage induced in-frame deletion in the spike glycoprotein that removes the furin-like cleavage site. bioRxiv.

Shajahan, A., Supekar, N.T., Gleinich, A.S., and Azadi, P. (2020). Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. Glycobiology. Steffen Klein, Mirko Cortese, Sophie L. Winter, Moritz Wachsmuth-Melm, C.J.N., and Berati Cerikan, Megan L. Stanifer, Steeve Boulant, Ralf Bartenschlager, P.C. (2020). SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. bioRxiv.

Turoňová, B., Hagen, W.J.H., Obr, M., Mosalaganti, S., Beugelink, J.W., Zimmerli, C.E., Kräusslich, H.-G., and Beck, M. (2020). Benchmarking tomographic acquisition schemes for high-resolution structural biology. Nat. Commun. *11*, 876.

Turoňová, 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 3.4 Å. J. Struct. Biol. *199*, 187-195.

Ulasli, M., Verheije, M.H., De Haan, C.A.M., and Reggiori, F. (2010). Qualitative and quantitative ultrastructural analysis of the membrane rearrangements induced by coronavirus. Cell. Microbiol. *12*, 844-861.

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.e6.
Walls, A.C., Tortorici, M.A., Frenz, B., Snijder, J., Li, W., Rey, F.A., DiMaio, F., Bosch, B.-J., and Veesler, D. (2016). Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. Nat. Struct. Mol. Biol. *23*, 899-905.

Walls, A.C., Tortorici, M.A., Snijder, J., Xiong, X., Bosch, B.-J., Rey, F.A., and Veesler, D. (2017a). Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. Proc. Natl. Acad. Sci. *114*, 11157-11162.

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

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.

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.

Xiaoli Xiong, Kun Qu, Katarzyna A. Ciazynska, Myra Hosmillo, Andrew P. Carter, Soraya Ebrahimi, Zunlong Ke, Sjors H.W. Scheres, Laura Bergamaschi, Guinevere L. Grice, Ying Zhang, The CITIID-NIHR COVID-19 BioResource Collaboration, James A. Nathan, Stephen, J.A.G.B. (2020). A thermostable, closed, SARS-CoV-2 spike protein trimer. bioRxiv.

Yongfei Cai, Jun Zhang, Tianshu Xiao, Hanqin Peng, Sarah M. Sterling, Richard M. Walsh Jr., Shaun Rawson, Sophia Rits-Volloch, B.C. (2020). Distinct conformational states of SARS-CoV-2 spike protein. bioRxiv.