Supplementary Materials for

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

Beata Turoňová*, Mateusz Sikora*, Christoph Schürmann*, Wim J. H. Hagen,
Sonja Welsch, Florian E. C. Blanc, Sören von Bülow, Michael Gecht, Katrin Bagola,
Cindy Hörner, Ger van Zandbergen, Jonathan Landry, Nayara Trevisan Doimo
de Azevedo, Shyamal Mosalaganti, Andre Schwarz, Roberto Covino,
Michael D. Mühlebach, Gerhard Hummer†, Jacomine Krijnse Locker†, Martin Beck†

*These authors contributed equally to this work.
†Corresponding author. Email: gerhard.hummer@biophys.mpg.de (G.H.); jacomina.krijnselocker@pei.de
(J.K.L.); martin.beck@biophys.mpg.de (M.B.)

This PDF file includes:

Materials and Methods
Figs. S1 to S7
Tables S1 and S2
Caption for Movie S1
References

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/cgi/content/full/science.abd5223/DC1)

MDAR Reproducibility Checklist (.pdf)
Movie S1 (.mov)

Materials and Methods

SARS-CoV-2 propagation

SARS-CoV-2 (isolate MUC-1, kind gift of G. Dobler, Bundeswehr Institute for Microbiology) was used for cell infections. It was propagated on Vero E6 cells (ATCC® CRL-1586) with an adsorption of 1 h at 37°C and shaking every 10-15 min. Infected cells were controlled for cytopathic effect every 24h and culture supernatant was usually collected 48h-72h post infection, clarified by centrifugation and was stored in aliquots at -80°C. Virus titer was titrated via TCID50 method according to Kaerber and Spaerman (26).

Virus preparation for cryoEM

SARS-CoV-2 identity was verified by Sanger sequencing of a SARS-CoV S gene fragment spanning the furin cleavage site. The RNA genomes of SARS-CoV-2 were isolated from infected Vero cells supernatant using TRIzol LS Reagent (Ambion, Thermo Fisher Scientific, Dreieich Germany) and Direct-zol RNA MiniPrep kit (Zymo research, Freiburg im Breisgau, Germany) according to manufacturers´ instructions and resuspended in 50 µL RNase-free water. Viral cDNA was reverse transcribed using Superscript II RT kit (Invitrogen) with 2 µL vRNA suspension as template and random hexamer primers, according to manufacturer´s instructions. For specific amplification F9newF and F9newR (5´- TAAGGTTGGTGGTAATTATAATTACCTG-3´ and 5´- AAAATAGTTGGCATCATAAAGTAATGGG-3´) primer were used and the respective genomic regions of SARS-CoV-2 S was amplified by PCR and cDNA as template (27). PCR products were directly sequenced using the following Primers: Wu_24_L and Wu_24_R (5´-TTGAACTTCTACATGCACCAGC-3´ and 5´-CCAGAAGTGATTGTACCCGC-3´) (Eurofins Genomics, Ebersberg, Germany). The furin cleavage site was sequenced at each virus passage. The virus used in this study, corresponding to passage 5 in Vero E6 cells, displayed deletion of residue 2035 to 2055 resulting in the inactivation of the furin cleavage site (11). For virus purification Vero cells were infected at a MOI of 0.1, the supernatant harvested at 6 hours post-infection and fixed overnight by addition of a final concentration of 4% paraformaldehyde, EM grade (EM sciences). The supernatant was clarified by two times 30 minutes of low-speed centrifugation. The clarified supernatant was placed on top of a 20% sucrose cushion in PBS (w/v) and centrifuged for 90 minutes at 30.000 rpm in a SW41 rotor. The pellet was resuspended in ~100 µl PBS and processed immediately for cryoEM.

Immunoblot Analysis

Cells were lysed and immunoblotted as previously described (28). Rabbit anti-SARS-CoV S antibody (1:3,000; ab252690; Abcam) and a rabbit anti-SARS-CoV N antibody (1:1,000; ABIN2214569, Antibody-online) were used for detection of SARS-CoV-2 S and -N respectively. A donkey anti-rabbit IgG-HRP (H&L) polyclonal antibody (1:10,000; Rockland, Gilbertsville, PA) served as secondary antibodies. Peroxidase activity was visualized with an enhanced chemiluminescence detection kit (Thermo Scientific, Bremen, Germany) on ChemiDoc MP Imaging System (Biorad, Dreieich, Germany).

Sequencing

Due to the limited amount of available RNA, the library was prepared using the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (New England Biolabs), which is able to generate libraries from low input samples. The protocol starts with the isolation of mRNA from total RNA by a polyA selection with oligodT beads, followed by fragmentation and priming of mRNA to generate double-stranded cDNA fragments. Subsequently, adaptors were ligated to cDNA fragments, which were then amplified and size-selected with SPRIselect beads (Beckman Coulter). Fragmentation time and size selections were optimized for the selection of larger fragment sizes (>200nt). The fragment size distribution of the final library was assessed using a Bioanalyzer with a DNA High Sensitivity kit (Agilent Technologies), and the concentration was measured with a Qubit® DNA High Sensitivity kit in Qubit® 2.0 Fluorometer (Life Technologies). The library was loaded on the Illumina MiSeq platform and sequenced bi-directionally, generating ~20 million reads, each 150 bases long.

The sequencing reads, which were mapped to the Vero cell genome (accessionID: DRA002256 from https://ddbj.nig.ac.jp/) using star (version 2.7.1a) were filtered out from subsequent analysis. The reads, longer than 40 bp, after trimming using cutadapt (version 2.3), of universal illumina adapter (AGATCGGAAGAG), were used as input in the guided-assembly galaxy pipeline VirAmp (29). The pipeline was run with default parameters, digital normalization option and the reference genome of the virus BetaCoV/Germany/BavPat1/2020 (accessionID: EPI_ISL_406862 from www.gisaid.org).

The genomic analysis confirmed the furin site deletion and the D614G allele. A discrepancy identified between automated genome assembly, a reference sequence and structural data has been clarified by forward and reverse Sanger sequencing. No additional differences in the sequence of the spike gene to the original MUC-1 isolate, which already contained the D614G allele, were detected.

Cryo electron tomography sample preparation and data acquisition

Purified virus suspension was mixed 4:1 (virus:gold) with 10 nm protein A-coated gold fiducials (UMC Utrecht). Per sample, 3 μl of the mix was applied onto a glow-discharged 200 mesh copper grid with a Quantifoil™ R 2/2 holey carbon support film of 2 μm diameter holes and an interspace of 2 μm. Quantifoil grids were glow discharged with a Pelco easiGlow™ glow discharge set for 45 sec at negative glow discharge head polarity and 15mA plasma current in 0.38 mbar residual air pressure. Max. 15 mins after glow discharging, samples were vitrified in liquid ethane after a blot time of 7 sec with Whatman 595 blot paper using a Vitrobot MarkIV (Thermo Scientific) at 100% humidity and 6°C.

The cryo-EM data was collected from three grids in three microscopy sessions at EMBL Heidelberg using a Thermo Scientific Titan Krios G3i electron microscope with a Gatan Bioquantum energy filter and K3 detector. SerialEM (30) was used to acquire the data. Over the three sessions, two grids were mapped in 6x6 montage with a bin 2 pixel size of 1,406 Å. Then grid square maps were collected with a bin 1 pixel size of 39 Å, -100 μm defocus, a 70 μm objective aperture and a 30 eV energy slit. During grid square mapping the already finished maps where inspected in a dummy instance of SerialEM, to

select the rare positions that had clusters of virus particles. A total of 61 grid square maps were manually searched to find the ~340 positions where tilt series were acquired.

SerialEM low-dose mode was setup to achieve 20 e$^-$/pixel/sec camera dose-rate with the 70 μm objective aperture and a 10 eV energy slit inserted, which gave spot size 7 with a 50 μm condenser aperture. SerialEM AutoComa and AutoStigmate were run, followed by a full Tune GIF and gain reference targeting the default 15 e$^-$/pixel/sec camera dose-rate. Positions were manually setup by taking a SerialEM Preview image (bin 4) that was saved as map, followed by a View image (bin 4) saved as anchor map. Before each tilt series a script was run to set eucentric height, realign to item, and set a target defocus as a persistent variable that changed each tilt series by 0.25 μm defocus in a range of -1.5 to - 4.5 μm defocus. After each stage tilt a script was run to recall the persistent variable with the target defocus value, check the liquid nitrogen autofill status of the system (and wait when active). This script further calls a 3 second wait time to reduce the drift. In contrast, the wait time parameter in the tilt controller was set to 0 to speed up the eucentric height procedure.

After each tilt-series acquisition, the energy filter zero loss peak was refined, and a new dark reference was acquired. The SerialEM tilt controller was set to autofocus and track each tilt. Dose-symmetric tilt scheme, +/-60 degree tilt range with a 3 degree tilt step in groups of 2 tilts was used to acquire the tilt series. Tilt images were acquired in counting mode with a calibrated physical pixel size of 1.329 Å and a total dose of 3.5 e$^-$/Å^2 over ten frames. Tilt series were saved 4x binned to perform quality checks and to store the metadata, while tilt images were saved as individual LZW compressed uncorrected TIF files to be used for later processing in IMOD. The average tilt series acquisition time was 14 minutes. The data collection parameters are summarized in Table S1.

Tilt-series preprocessing and tomogram reconstruction

The data processing followed the protocol from (31) unless stated otherwise. CTFFind4 (32) was used for CTF estimation and the tilt-series were corrected for dose-exposure (33) using Matlab scripts adapted for tomographic data (34). The tilt-series were aligned using fiducial-based alignment in eTomo (35). Gold beads were manually selected and automatically tracked. The fiducial model was corrected in all cases where the automatic tracking failed. The final alignment was computed without solving for any distortions. Two sets of tomograms were reconstructed. The first set consisted of 8x binned tomograms reconstructed using SIRT-like filter option in eTomo. These tomograms were used only for manual selection of virions' centers. A second set of tomograms was reconstructed with 3D-CTF correction using novaCTF (36) with phase-flip correction, astigmatism correction and 15 nm slab. Tomograms were binned 2x, 4x and 8x using Fourier3D (37). The second set was used for subtomogram averaging (SA). SA on binned data was performed with novaSTA (15) while the final refinement on unbinned data was done with STOPGAP (16).

Subtomogram averaging of the spike

In total 1094 virion centers were manually picked and their radii estimated. A spherical shape was used to generate initial positions and orientations on the lattice (38). The sampling distance among the positions on the sphere was 4 nm. Ten iterations of

alignment on 8x binned tomograms were run to allow the initial positions to lock on the membrane bilayer. The misaligned positions were removed using the geometry-based procedure described in (31). The remaining positions were used for ellipsoid fitting, which gave the parameters defining the algebraic equation for conics as well as more precise virion centers and radii. The mean radius of virions obtained in this way was ~40 nm. Five tomograms with lower defocus were used to manually pick the positions of 402 spikes. The initial orientations were generated in the same way as for the positions on the membrane. Ten iterations of alignment on 8x binned tomograms were run to obtain an initial reference of the spike. A new set of starting positions and orientations was generated for each virion by oversampling a sphere using the experimentally determined center and radius which was extended by 15 nm, i.e. the new positions were created roughly in the center of spikes rather than on the membrane. The sampling distance was 4 nm (which roughly corresponds to oversampling of 5-fold). The subtomograms that were too close to a carbon edge were automatically removed. Ten iterations of alignment were run on 8x binned tomograms using the initial reference of the spike. The dataset was subsequently cleaned to remove the subtomograms that shifted to the same positions (distance cleaning), to carbon edge, to the edge of tomograms and/or were overlapping with neighboring virions. Lastly, subtomograms with low cross-correlation (CC) values were removed from the dataset. As the CC values differ with respect to the tomogram defocus, the CC-based cleaning was done per tomogram. The CC values followed a bimodal distribution and thus the Otsu threshold (39) was used to find the optimal threshold for cleaning. To assure the best possible alignment, i.e. not driven by false-positive particles or the carbon edge, ten iterations of alignment were run on 8x binned tomograms but this time without any starting reference. The positions of the remaining subtomograms were kept but their orientations were set to the initial ones. The subtomograms were split in two halves and processed independently. Four iterations of alignment we run on 4x binned tomograms followed by 3 iterations of alignment on 2x binned data.

Classification of closed and open conformers

Classification was performed on 2x binned subtomograms using simulated annealing in combination with stochastic hill climbing (SHC) as implemented in STOPGAP. Eight starting references were created, each from 1000 randomly chosen particles. Alignment focused only the three possible in-plane orientations of the head. In total, 10 iterations of alignment using simulated annealing with starting temperature of 10 and SHC were run, followed by 20 iterations of SHC alignment. Four independent alignments with the same parameter setup but a different starting set of references were run. In each set of the final 8 classes the same 4 distinct particle populations were identified. These were used as starting references for 20 iterations of multi-reference alignment with SHC. To check the consistency of assignment of particles to different classes, four independent alignments were run with the same setup and the same starting references. The results from each alignment were cross-checked and only particles that were assigned to the same class in all 4 alignments were considered further. The class #1 with all three RBDs closed had 8,273 particles, the class #2 with 1 fully open RBD had 4,321 particles. Class #3 had 4,885 and class #4 had 5,556 particles. The remaining 3,198 particles were discarded as they could not be reliably assigned to any class. The class #3

has 2 closed RBDs, while the third RBD was not well resolved. This class most probably represented intermediated states between classes #1 and #2. Class #4 could not be clearly attributed to any known conformer, possibly due to technical limitations.

Subtomogram averaging of the asymmetric unit

The positions and orientations from the alignment on 4x binned data were used as a starting point for the asymmetric unit alignment. The spike was split into its asymmetric units which were superimposed, effectively triplicating the number of subtomograms. The classification results were used to identify asymmetric units containing closed RBDs. In total, 44,581 subtomograms with closed RBD (from classes #1, #2 and #3) were extracted. The one-open RBD conformer was not subjected to this procedure because it is asymmetric from the first place. For the closed RBD asymmetric unit four iterations of alignment were run on 2x binned subtomograms. Prior the alignment of unbinned data subtomograms with outlying density values were discarded. Fifteen iterations of alignment on unbinned data were run, followed by CC cleaning and final 9 iterations of alignment. The final average was computed from 35,997 particles yielding a 4.9 Å structure.

Subtomogram averaging of the lower leg

The average obtained after the alignment on 4x binned data was used to place the mask at the lower leg region and the corresponding positions and orientations were used as starting point for the alignment. In the first iteration on 4x binned data an angular range up to 90 degrees was chosen to allow locking on the lower leg followed by 2 iterations of 30 degrees angular range. The center of rotation was in the knee region of the leg. Three iterations on 2x binned data were run using both novaSTA and STOPGAP with the center of rotation being the lower leg. Simulated annealing in combination with stochastic hill climbing from STOPGAP was used to obtain the structure of the lower leg (Fig. S5B). The structure was resolved to a resolution of 11 Å using 8,885 particles and imposing C3 symmetry.

Distance-based classification

The orientations and positions found during the alignment on 8x binned data were used to measure the distance of the spike to the membrane. For each spike an intersection of its normal (with inverted direction) and the experimentally determined ellipsoid describing its corresponding virion was found. The distance was measured as the length between the spike center and the intersection with the membrane. Since the lengths of the upper and lower leg are similar this method gives sufficiently precise results also in cases where the spike is bend. However, in some cases the intersection could not be determined. This occurred mostly due to particle misalignment and in rare cases due to strong bending of a spike at one of its hinges. The particles without the intersection point were removed from further classification. Particles with distance less than 9 nm were also discarded as they consisted mostly of misaligned particles that shifted to the membrane during the alignment. The mean distance value was ~19 nm with standard deviation of ~4 nm. Four classes were created based on this: class 1 with distances 9-13 nm, class 2 (13-17 nm), class 3 (17-21 nm), and class 4 with distances larger than 21 nm. The class occupancy was 1,560, 4,366, 7,273, and 5,252 particles, respectively. Four iterations of

alignment on 8x binned data were run to obtain averages for each class (see Fig. 2F). The spike orientations towards the membrane were computed as an angular deviation between the spike normal and an expected membrane normal at the intersection point. The particles with a consistent distance from the membrane (up to the deviation of 2 nm from the mean) and with angular deviation of less than 35 degrees from the membrane normal were chosen for alignment of the whole spike including the leg. The final structure was resolved from 2x binned data to 15 Å (see Fig. S5A) using 3,193 particles and without imposing any symmetry.

S protein distance distribution and cluster analysis

We analyzed the organization of S proteins on the viral surfaces in terms of distance and cluster-size distributions. For this analysis, we used the (x,y,z) positions of S proteins on 418 virions. To avoid possible bias in top views, we included S proteins with a polar angle between $\pi/3$ and $2\pi/3$ with respect to the z axis, giving us ~8,150 spike positions. We corrected for the geometric restriction in this selection, for variations in virion size, and for the curved shape of the viral surface by repeatedly drawing an equal number of points randomly distributed within a ring around the equator of spheres matching in size. For reference, we also sampled the distribution of equal numbers of hard disks 10 nm in diameter on these spherical waists. For actual, random, and hard-disk S-protein positions, we calculated the distribution of pair distances r. The probability ratio for actual and random positions defines a pair-distance distribution function g(r), where a value of one indicates a random distribution of the S proteins on the viral surface. For actual and hard-disk S-protein positions, we also performed a clustering analysis, assigning two proteins (or disks) to the same cluster if their centers are within 20 nm. By comparing the cluster-size distributions for S protein on virions and for hard disks, we assessed the tendency of S proteins to cluster beyond what is expected to occur at random.

Construction of the S model

We built a model of the full-length S based on a single particle cryoEM structure (PDB ID: 6VSB, (3)). We constructed the missing loops and the N-terminal tail using MODELLER (40). We decorated all 22 N-glycosylation sites per chain of S, using at each site the most abundant species as detected by Watanabe et al. (41), including heavily branched and sialylated glycan species on the stalk.

Following bioinformatic predictions (42), we used Ccbuilder (43) to construct independent models of the S upper leg (residues 1137-1163) and lower leg (residues 1161-1214) as coiled coils. We constructed the transmembrane domain (TMD) as a trimer of transmembrane helices (residues 1212-1237) followed by an unstructured C-terminal region. We palmityolated all cysteines in the membrane region and embedded the transmembrane domain in an ER-like lipid composition (44) (Table S2). The lipid composition was chosen based on the point of viral assembly, lacking detailed lipidomics data on purified viral membranes.

Separate models of the S head, upper leg, lower leg and TMD domain were processed and solvated using CHARMM-GUI (45), equilibrated and evolved for up to 1 μs of unrestrained simulation time. We assembled final snapshots by first manually matching the separate structural domains using PyMOL (46), and generating connecting residues with MODELLER. A full-length model of S was then placed in a triclinic box

(a=17.4 nm, b=17.4 nm, c=38.4 nm with angles 90 deg, 90 deg, 120 deg), solvated and supplemented with 0.15M NaCl (Fig. 3A). We used periodic copies of the system to generate four S proteins in a simulation box containing ~4.1 million atoms in total.

We note that the simulation model described here had been generated independently from the tomographic data and before they were available. Although the simulation model and the EM maps are overall remarkably consistent, in particular concerning the positions of the hinges, there are some minor deviations. Notably, the short coiled coil forming the upper leg was modeled as left handed. Even though this does not affect the conclusions here, it will be interesting to explore the effect on stability of changing the handedness. We note further that the relatively high effective density of S proteins in the simulation system at ~15 nm spacing sterically restricted the mobility, thus preventing large bending angles.

Details of the simulations

We used the CHARMM36m protein and glycan force field (47), with the TIP3P water model (48) and ion parameters by Luo and Roux (49). Throughout this study, we used GROMACS 2019.6 (50) to perform MD simulations. We performed the initial equilibration of each of the component systems as well as the final one in the following way: first, we performed energy minimization using the steepest descent algorithm for 55000 steps, followed by equilibration in the NVT ensemble for 375 ps with a time step of 1 fs, and additional 1500 ps with a time step of 2 fs. The systems were coupled to a thermal bath using the Berendsen thermostat (51). After an initial 250 ps relaxation at constant volume, semiisotropic pressure coupling was added (Berendsen barostat tau=5ps, compressibility=4.5e-5 bar^{-1}). We constrained all bonds involving hydrogen atoms using the LINCS algorithm (52). We restrained atom positions and dihedral angles during equilibration, gradually decreasing force constants from 1000 kJ/mol/nm^2 to 0.

In the constant pressure (1 atm, Parinello-Rahman barostat (53)) and temperature (310 K, V-rescale thermostat (54)) production runs we used an integration time step of 4 fs and cutoffs of non-bonded interactions of 1 nm. To permit this relatively large time step, the system was fully deuterated (i.e., hydrogen masses were doubled). Electrostatic interactions were calculated with the particle-mesh Ewald algorithm. Structures were saved at intervals of 10 nanoseconds and used to identify the flexible hinges in the structure and to calculate density maps for as a basis for quantitative comparisons with cryoEM.

Polar angle – distance calculation

For each S and each frame of the trajectory, the TMD was translated so that its center of geometry coincided with the coordinate origin (0,0,0). A sphere of 40 nm radius representing a virion was then placed at (0,0,-40) nm; a ray aligned with the S head long axis was projected towards it; and the intersection point was marked. The distance was calculated between the intersection point and the S-head center of geometry (consistent with the choice of the S center in the analysis of the experiments). The polar angle was calculated between the S-head long axis and the virion normal at the intersection point. Polar angles were then plotted against the distance (Fig. S6D).

Sampling of high polar angles was limited in the simulations due to the relatively high density of S in the MD simulation system. To increase the sampling, we re-

assembled the hinges of the S stalk. First, coordinates of the sections of S surrounding each joint were extracted from each frame. Second, we rigid-body superimposed these fragments, using the upper leg to re-sample the hip joint, and the lower leg to re-sample the knee joint. By matching these fragments at random, we generated 28000 models of isolated S proteins on the membrane unencumbered by spike-spike interactions. Clashes of re-sampled S structures with the virion were removed (45 particles) as well as cases where the ray did not intersect the virion (3,333 particles). We then analyzed these re-sampled models of S in terms of their polar angle and distance as described above. Removing the constraints from spike-spike interactions broadened in particular the angle distribution (Fig. S6D).

Generation of density maps

Trajectories of all four S proteins were aligned on the head domain. An electron density map was subsequently calculated using Gromaps (55) with the same grid spacing as used in the experiment (0.1329 nm) and spread parameter sigma adjusted to reproduce the experimental resolution.

Density-guided simulations

In an initial step maps were superimposed onto one of the S proteins in a simulation box using ChimeraX (56). Density guided simulations were performed using a dedicated tool in GROMACS 2020.2 (57). Atoms belonging to S (excluding the TMD and lower leg domain and including all glycosylation sites) were used for fitting. Residues not participating in the fitting were position restrained with a force constant of 1000 kJ/mol/nm^2. Simulation-derived density maps were calculated with atoms spread on a 3D grid by a Gaussian transform (sigma of 0.2 nm and contribution vanishing above 4*sigma). All atoms were treated uniformly, i.e. gaussian spreading was independent on atom mass or charge. The density cross-correlation was used to quantify agreement between the (normalized) experimental and calculated density maps, with a scaling factor of 2 x 1010 kJ/mol. Forces acting on atoms were applied every 100 steps and adaptive scaling was used (time constant of 4 ps).

Rigid body fitting to raw tomogram data

S proteins were segmented manually using ImageJ (58). Snaphots of S proteins taken from the MD simulation were selected to best match the geometry of the hinges, centered in the density and rigid-body fitted using ChimeraX.

Rigid body fitting to subtomogram averaged maps

Rigid body fitting of the single particle model (PDB ID 6VXX) was performed in ChimeraX using the "Fit in Map" tool.

Visualization

All subtomogram averaged maps shown were post-processed in RELION using automatic B-factor sharpening (59). The raw tomograms visualized in the Figures were deconvolved using tom_deconv (60).
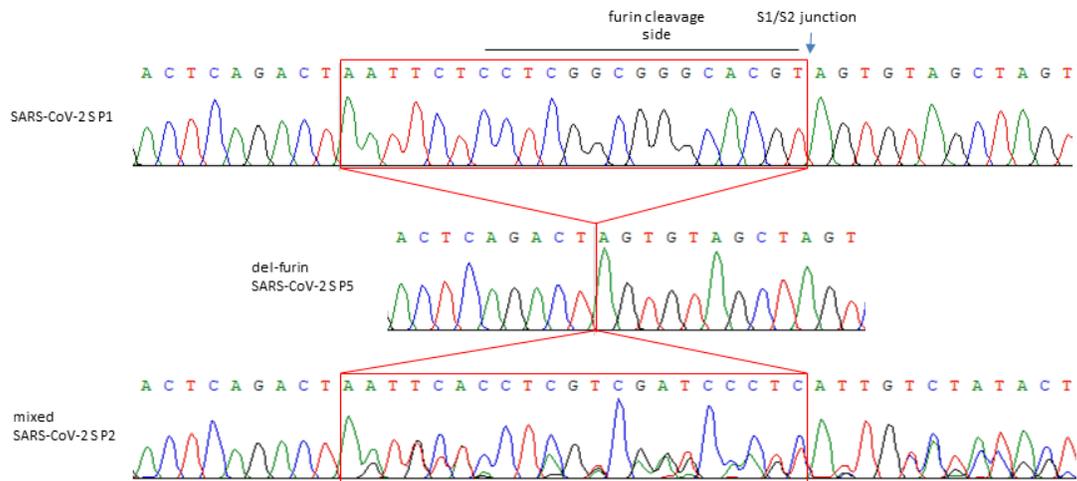
Fig. S1.

Deletion of the furin cleavage site within the viral mRNA in the SARS-CoV-2 S gene. Sanger sequencing of the SARS-CoV-2 (MUC-1) at passage 1 (P1), 2 (P2) and 5 (P5) propagated on Vero E6 cells as indicated. SARS-CoV-2 S P1 shows the nucleotide sequence spanning the S1/S2 junction region with furin cleavage site of the S glycoprotein. The chromatogram of SARS-CoV-2 passage 5 (P5) shows a 21bp deletion of the furine cleavage side (del-furin). The chromatogram of passage 2 SARS-CoV-2 (P2) shows sequences with and without the 21bp deletion (mixed). Sanger sequencing data from passage 3-4 show exclusively the 21bp deletion (data not shown).
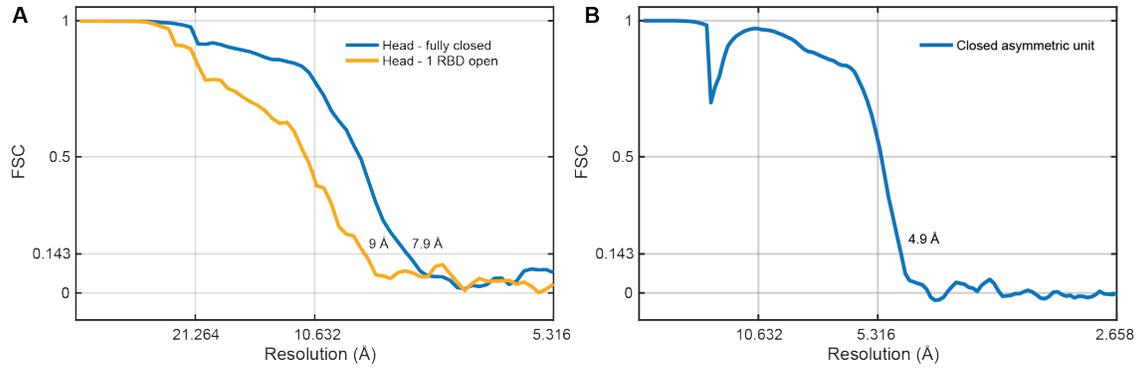
Fig. S2

Resolution measurement of subtomogram averages. Fourier Shell Correlation (FSC) curves are shown for the subtomogram averages of the two conformers of the spike head (A) and its asymmetric unit (B).
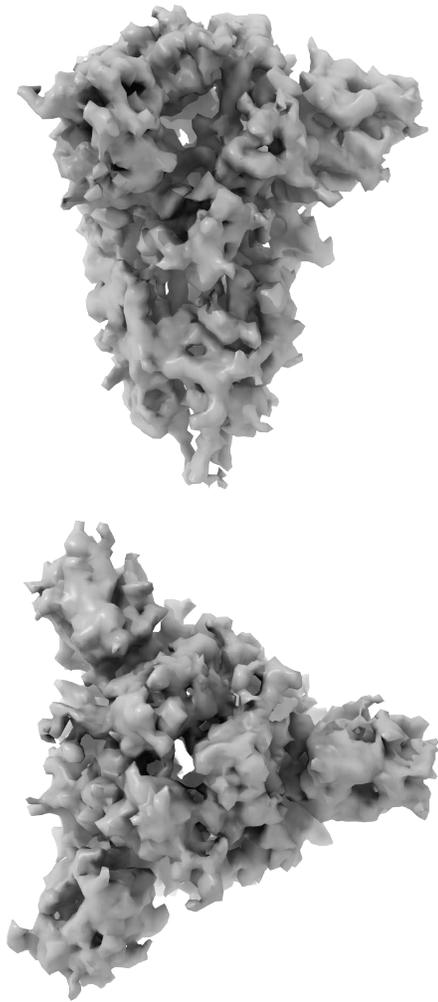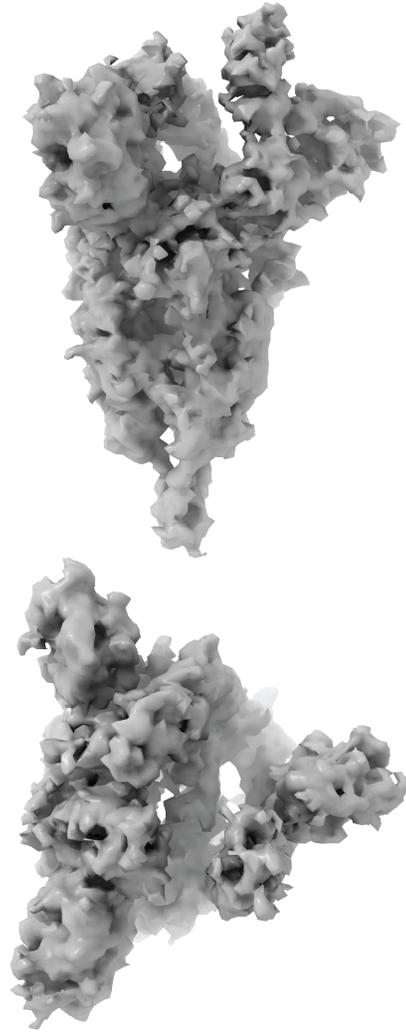
Fig. S3

Closed and open state conformers. (A) Side and top view of the subtomogram averaged structure of the fully closed conformation of the S trimer. The depicted structure was obtained from 2x binned data and has a resolution of 7.9 Å. (B) Side and top view of the subtomogram averaged structure of S trimer with 1 RBD open. The depicted structure was obtained from 2x binned data and has resolution of 9 Å.
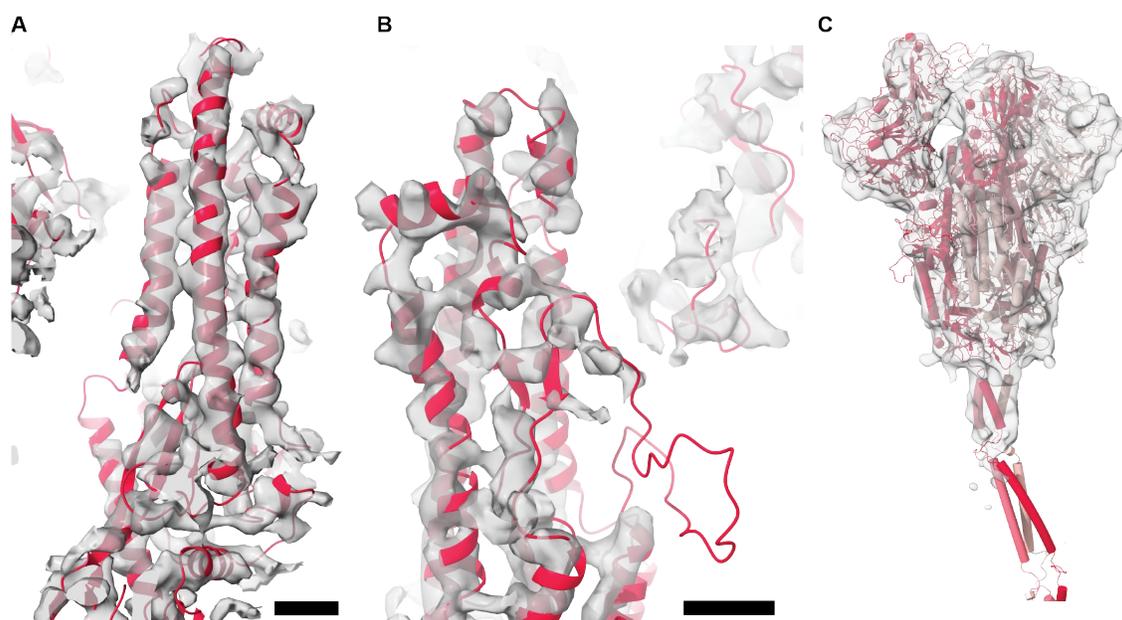
Fig. S4

Flexible fitting into the cryoEM map of S. (A) Flexible fitting into the asymmetric unit average of S in the region of the helical bundle with in the head. Scale bar 10 Å, contour level 0.6. (B) Same as (A) but highlighting a detail at the top in a rotated view. Scale bar 10 Å, contour level 0.6. (C) Flexible fit into the one-open RBD conformation. Averaged electron optical density is shown isosurface rendered in transparent gray; the flexibly fitted MD model is shown as red ribbons. Contour level 0.48.

13
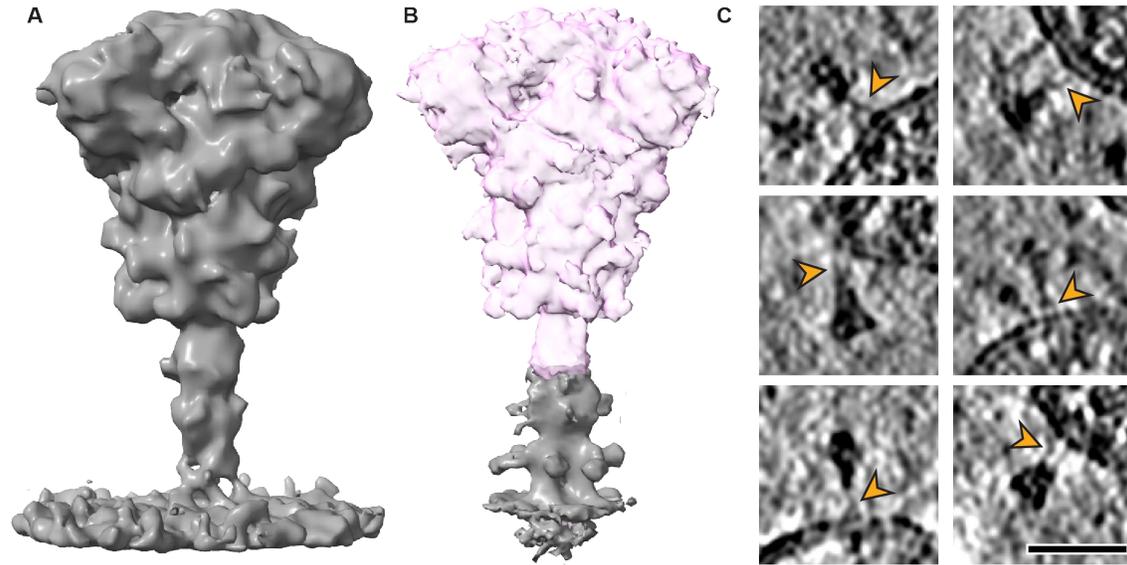
Fig. S5

Refinement of the lower leg. (A) Subtomogram averaged structure with 16 Å resolution of the whole S obtained on 2x binned data. (B) Structure of the lower leg resolved to 11 Å using focused masking (gray) and superimposed by a separate average of S (pink). (C) Examples of unfolded coiled-coils in tomographic slices. Red arrowheads indicated "forking" at hinge positions. Scale bar 30 nm.
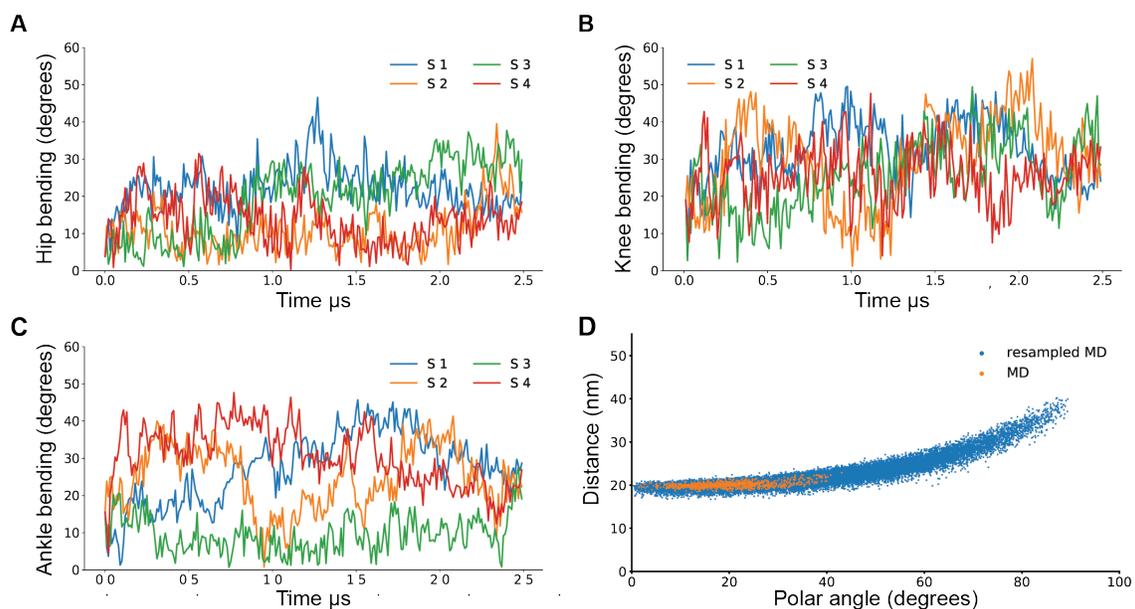
Fig. S6

Flexibility of S in MD simulations. (A-C) Time traces of the bending angles of the hinges at the hip (A), knee (B) and ankle (C) shown in different colors for the four spike proteins in the simulation system. (D) Polar angle vs. distance plot (compare to Fig. 2E) calculated from the MD trajectory (orange) and from a resampled trajectory (blue, see Methods for the details).
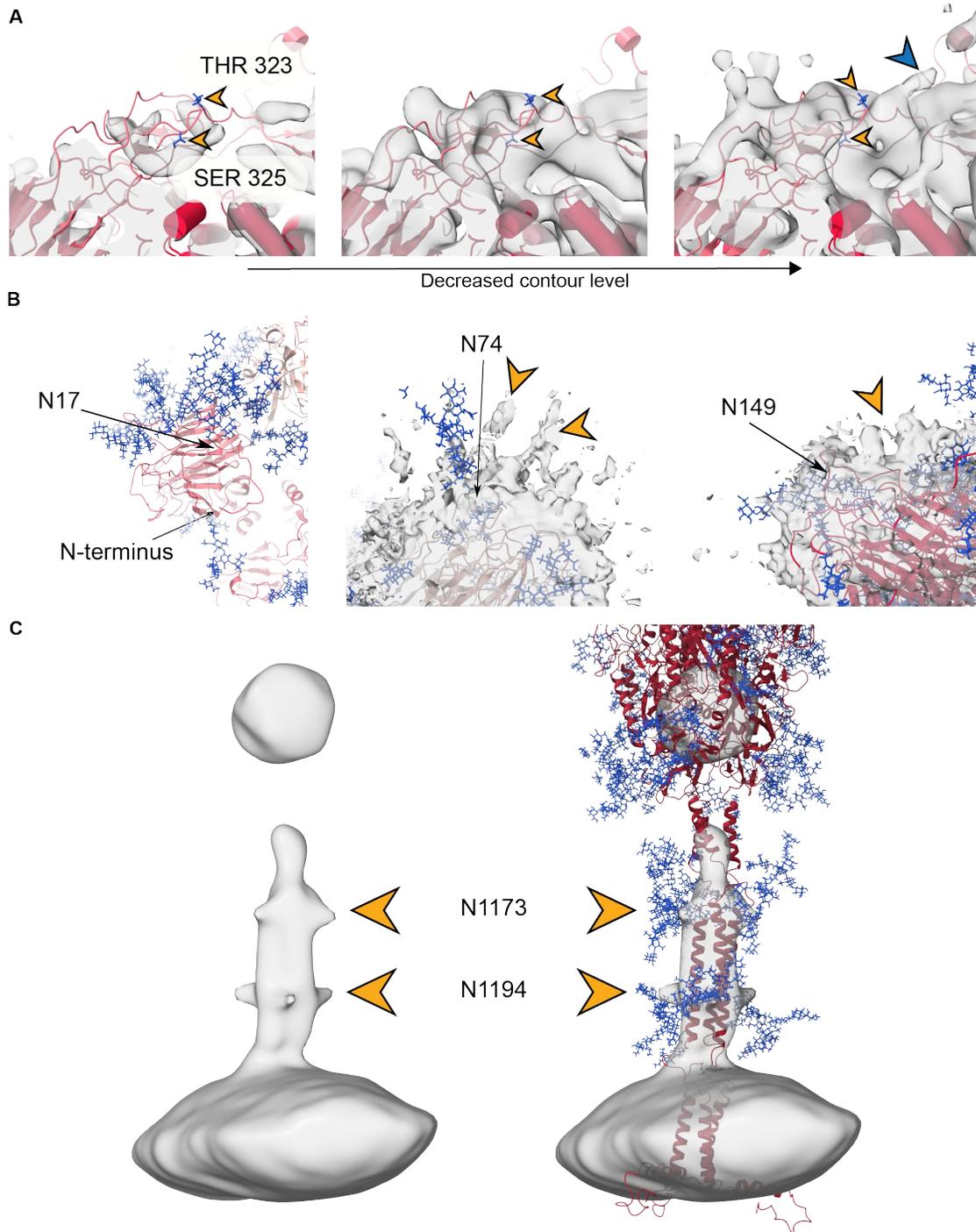
Fig. S7

O- and N-glycosylation of S. (A) Two potential O-glycosylation sites are marked with orange arrowheads. Maps shown at different contour levels. In comparison to the N-glycosylation sites, these sites lack significant extra density. At the lowest level of contouring an elongated feature becomes visible (blue arrowhead). (B) N-glycosylation of the N-terminal domain. The positions of sequons N17, N74 and N149 are indicated

with arrows. N17 is located on a flexible N-terminus, likely resulting in very diffuse electron optical density. Unassigned density that could correspond to a glycan is marked with orange arrowheads. (C) Average density calculated for the MD trajectory aligned on the lower leg and filtered to the experimental resolution. Left: Due to the flexibility at the hinges, the S head becomes invisible in the averaged density. The lowermost glycosylation sites can be seen protruding from the lower leg. Right: Atomistic model fitted to the density.

Table S1.

Data collection parameters.

Microscope	Titan Krios G3i
Acceleration voltage [KeV]	300
Energy slit [eV]	10
Detector	Gatan K3
Mode	Counting mode
Pixel size [Å]	1.329
Defocus range [μm]	-1.5, -4.5
Tilt scheme	Dose-symmetric
Tilt range	-60/+60º
Tilt step	3º
Total dose [e$^-$/Å^2]	140
Number of frames per tilt	10
Tilt series acquired	~340

Table S2.

Composition of the membrane used in the simulations.

Lipid	Full name	Percentage
DOPC	1,2-dioleoyl-glycero-3-phosphocholine	25
POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine	25
POPE	1-palmitoyl-2-oleoyl-glycero-3-phosphoethanolamine	20
POP	1-palmitoyl-2-oleoyl-glycero-3-phosphoinositol	15
POPS	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine	5
CER160	N-palmitoyl-D-erythro-sphingosine	5
CHOL	Cholesterol	5

Movie S1

Cryo electron tomography of SarS-CoV-2. A representative cropped tomogram is shown as constitutive slices along the z-axis. For software used to deconvolve tomograms see (60).

References and Notes

1. C. A. M. de Haan, P. J. M. Rottier, Molecular interactions in the assembly of coronaviruses. *Adv. Virus Res.* **64**, 165–230 (2005). doi:10.1016/S0065-3527(05)64006-7 Medline

2. A. C. Walls, Y.-J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, D. Veesler, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281–292.e6 (2020). doi:10.1016/j.cell.2020.02.058 Medline

3. D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, J. S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020). doi:10.1126/science.abb2507 Medline

4. R. J. Hulswit, C. A. de Haan, B.-J. Bosch, Coronavirus Spike Protein and Tropism Changes. *Adv. Virus Res.* **96**, 29–57 (2016). doi:10.1016/bs.aivir.2016.08.004 Medline

5. F. Li, Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* **3**, 237–261 (2016). doi:10.1146/annurev-virology-110615-042301 Medline

6. S. C. Harrison, Viral membrane fusion. *Virology* **479–480**, 498–507 (2015). doi:10.1016/j.virol.2015.03.043 Medline

7. Y. Cai, J. Zhang, T. Xiao, H. Peng, S. M. Sterling, R. M. Walsh Jr., S. Rawson, S. Rits-Volloch, B. Chen, Distinct conformational states of SARS-CoV-2 spike protein. *Science* 10.1126/science.abd4251 (2020). doi:10.1126/science.abd4251 Medline

8. S. Duquerroy, A. Vigouroux, P. J. Rottier, F. A. Rey, B. J. Bosch, Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein. *Virology* **335**, 276–285 (2005). doi:10.1016/j.virol.2005.02.022 Medline

9. S. Klein, M. Cortese, S. L. Winter, M. Wachsmuth-Melm, C. J. Neufeldt, B. Cerikan, M. L. Stanifer, S. Boulant, R. Bartenschlager, P. Chlanda, SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. bioRxiv 167064 [Preprint]. 16 August 2020. https://doi.org/10.1101/2020.06.23.167064.

10. S. Belouzard, J. K. Millet, B. N. Licitra, G. R. Whittaker, Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses* **4**, 1011–1033 (2012). doi:10.3390/v4061011 Medline

11. S.-Y. Lau, P. Wang, B. W.-Y. Mok, A. J. Zhang, H. Chu, A. C.-Y. Lee, S. Deng, P. Chen, K.-H. Chan, W. Song, Z. Chen, K. K.-W. To, J. F.-W. Chan, K.-Y. Yuen, H. Chen, Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. *Emerg. Microbes Infect.* **9**, 837–842 (2020). doi:10.1080/22221751.2020.1756700 Medline

12. N. S. Ogando, T. J. Dalebout, J. C. Zevenhoven-Dobbe, R. W. A. L. Limpens, Y. van der Meer, L. Caly, J. Druce, J. J. C. de Vries, M. Kikkert, M. Bárcena, I. Sidorov, E. J. Snijder, SARS-coronavirus-2 replication in Vero E6 cells: Replication kinetics, rapid adaptation and cytopathology. *J. Gen. Virol.* 10.1099/jgv.0.001453 (2020). doi:10.1099/jgv.0.001453 Medline

13. C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, M. Seilmaier, C. Drosten, P. Vollmar, K. Zwirglmaier, S. Zange, R. Wölfel, M. Hoelscher, Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany. *N. Engl. J. Med.* **382**, 970–971 (2020). doi:10.1056/NEJMc2001468 Medline

14. L. Zhang, C. B Jackson, H. Mou, A. Ojha, E. S Rangarajan, T. Izard, M. Farzan, H. Choe, The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. bioRxiv 148726 [Preprint]. 12 June 2020. https://doi.org/10.1101/2020.06.12.148726.

15. B. Turoňová, turonova/novaSTA: NovaSTA, Version v1.0, Zenodo (2020); https://doi.org/10.5281/zenodo.3973623.

16. W. Wan, williamnwan/STOPGAP: STOPGAP 0.7.1, Version 0.7.1, Zenodo (2020); https://doi.org/10.5281/zenodo.3973664.

17. P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, High-resolution protein design with backbone freedom. *Science* **282**, 1462–1467 (1998). doi:10.1126/science.282.5393.1462 Medline

18. A. C. Walls, M. A. Tortorici, J. Snijder, X. Xiong, B.-J. Bosch, F. A. Rey, D. Veesler, Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11157–11162 (2017). doi:10.1073/pnas.1708727114 Medline

19. A. C. Walls, M. A. Tortorici, B. Frenz, J. Snijder, W. Li, F. A. Rey, F. DiMaio, B.-J. Bosch, D. Veesler, Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat. Struct. Mol. Biol.* **23**, 899–905 (2016). doi:10.1038/nsmb.3293 Medline

20. A. Shajahan, N. T. Supekar, A. S. Gleinich, P. Azadi, Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* 10.1093/glycob/cwaa042 (2020). doi:10.1093/glycob/cwaa042 Medline

21. D. J. Benton, A. Nans, L. J. Calder, J. Turner, U. Neu, Y. P. Lin, E. Ketelaars, N. L. Kallewaard, D. Corti, A. Lanzavecchia, S. J. Gamblin, P. B. Rosenthal, J. J. Skehel, Influenza hemagglutinin membrane anchor. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10112–10117 (2018). doi:10.1073/pnas.1810927115 Medline

22. Z. Ke, J. Oton, K. Qu, M. Cortese, V. Zila, L. McKeane, T. Nakane, J. Zivanov, C. J. Neufeldt, B. Cerikan, J. M. Lu, J. Peukes, X. Xiong, H. G. Kräusslich, S. H. W. Scheres, R. Bartenschlager, J. A. G. Briggs, Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* 10.1038/s41586-020-2665-2 (2020). doi:10.1038/s41586-020-2665-2 Medline

23. H. Yao, Y. Song, Y. Chen, N. Wu, J. Xu, C. Sun, J. Zhang, T. Weng, Z. Zhang, Z. Wu, L. Cheng, D. Shi, X. Lu, J. Lei, M. Crispin, Y. Shi, L. Li, S. Li, Molecular architecture of the SARS-CoV-2 virus. *Cell* 10.1016/j.cell.2020.09.018 (2020). doi:10.1016/j.cell.2020.09.018

24. R. Henderson, R. J. Edwards, K. Mansouri, K. Janowska, V. Stalls, S. M. C. Gobeil, M. Kopp, D. Li, R. Parks, A. L. Hsu, M. J. Borgnia, B. F. Haynes, P. Acharya, Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* 10.1038/s41594-020-0479-4 (2020). doi:10.1038/s41594-020-0479-4 Medline

25. X. Xiong, K. Qu, K. A. Ciazynska, M. Hosmillo, A. P. Carter, S. Ebrahimi, Z. Ke, S. H. W. Scheres, L. Bergamaschi, G. L. Grice, Y. Zhang, J. A. Nathan, S. Baker, L. C. James, H. E. Baxendale, I. Goodfellow, R. Doffinger, J. A. G. Briggs; CITIID-NIHR COVID-19 BioResource Collaboration, A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct. Mol. Biol.* 10.1038/s41594-020-0478-5 (2020). doi:10.1038/s41594-020-0478-5 Medline

26. J. Hubert, in *Bioassay* (Hunt Publishing, ed. 2, 1984), pp. 65–66.

27. A. D. Davidson, M. K. Williamson, S. Lewis, D. Shoemark, M. W. Carroll, K. J. Heesom, M. Zambon, J. Ellis, P. A. Lewis, J. A. Hiscox, D. A. Matthews, Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med.* **12**, 68 (2020). [doi:10.1186/s13073-020-00763-0](doi:10.1186/s13073-020-00763-0) Medline

28. S. Funke, A. Maisner, M. D. Mühlebach, U. Koehl, M. Grez, R. Cattaneo, K. Cichutek, C. J. Buchholz, Targeted cell entry of lentiviral vectors. *Mol. Ther.* **16**, 1427–1436 (2008). [doi:10.1038/mt.2008.128](doi:10.1038/mt.2008.128) Medline

29. Y. Wan, D. W. Renner, I. Albert, M. L. Szpara, VirAmp: A galaxy-based viral genome assembly pipeline. *Gigascience* **4**, 19 (2015). [doi:10.1186/s13742-015-0060-y](doi:10.1186/s13742-015-0060-y) Medline

30. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005). [doi:10.1016/j.jsb.2005.07.007](doi:10.1016/j.jsb.2005.07.007) Medline

31. B. Turoňová, W. J. H. Hagen, M. Obr, S. Mosalaganti, J. W. Beugelink, C. E. Zimmerli, H.-G. Kräusslich, M. Beck, Benchmarking tomographic acquisition schemes for high-resolution structural biology. *Nat. Commun.* **11**, 876 (2020). [doi:10.1038/s41467-020-14535-2](doi:10.1038/s41467-020-14535-2) Medline

32. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015). [doi:10.1016/j.jsb.2015.08.008](doi:10.1016/j.jsb.2015.08.008) Medline

33. T. Grant, N. Grigorieff, Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* **4**, e06980 (2015). doi:10.7554/eLife.06980 Medline

34. W. Wan, L. Kolesnikova, M. Clarke, A. Koehler, T. Noda, S. Becker, J. A. G. Briggs, Structure and assembly of the Ebola virus nucleocapsid. *Nature* **551**, 394–397 (2017). doi:10.1038/nature24490 Medline

35. J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996). doi:10.1006/jsbi.1996.0013 Medline

36. B. Turoňová, F. K. M. Schur, W. Wan, J. A. G. Briggs, Efficient 3D-CTF correction for cryo-electron tomography using NovaCTF improves subtomogram averaging resolution to 3.4Å. *J. Struct. Biol.* **199**, 187–195 (2017). [doi:10.1016/j.jsb.2017.07.007](doi:10.1016/j.jsb.2017.07.007) Medline

37. B. Turoňová, turonova/Fourier3D: Fourier3D, Version v1.0, Zenodo (2020); https://doi.org/10.5281/zenodo.3973621.

38. F. Förster, O. Medalia, N. Zauberman, W. Baumeister, D. Fass, Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4729–4734 (2005). doi:10.1073/pnas.0409178102 Medline

39. N. Otsu, A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979). doi:10.1109/TSMC.1979.4310076

40. B. Webb, A. Sali, Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **54**, 5.6.1–5.6.37 (2016). doi:10.1002/cpbi.3 Medline

41. Y. Watanabe, J. D. Allen, D. Wrapp, J. S. McLellan, M. Crispin, Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330–333 (2020). Medline

42. T. L. Vincent, P. J. Green, D. N. Woolfson, LOGICOIL—Multi-state prediction of coiled-coil oligomeric state. *Bioinformatics* **29**, 69–76 (2013). doi:10.1093/bioinformatics/bts648 Medline

43. C. W. Wood, D. N. Woolfson, CCBuilder 2.0: Powerful and accessible coiled-coil modeling. *Protein Sci.* **27**, 103–111 (2018). doi:10.1002/pro.3279 Medline

44. J. Jacquemyn, A. Cascalho, R. E. Goodchild, The ins and outs of endoplasmic reticulum-controlled lipid biosynthesis. *EMBO Rep.* **18**, 1905–1921 (2017). doi:10.15252/embr.201643426 Medline

45. J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks 3rd, A. D. MacKerell Jr., J. B. Klauda, W. Im, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413 (2016). doi:10.1021/acs.jctc.5b00935 Medline

46. The PyMOL Molecular Graphics System, Version 2.3.5, Schrödinger, LLC; pymol.org.

47. S.-J. Park, J. Lee, Y. Qi, N. R. Kern, H. S. Lee, S. Jo, I. Joung, K. Joo, J. Lee, W. Im, CHARMM-GUI *Glycan Modeler* for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* **29**, 320–331 (2019). doi:10.1093/glycob/cwz003 Medline

48. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983). doi:10.1063/1.445869

49. Y. Luo, B. Roux, Simulation of Osmotic Pressure in Concentrated Aqueous Salt Solutions. *J. Phys. Chem. Lett.* **1**, 183–189 (2010). doi:10.1021/jz900079w

50. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015). doi:10.1016/j.softx.2015.06.001

51. H. J. Berendsen, J. Postma, W. F. van Gunsteren, A. DiNola, J. Haak, Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984). doi:10.1063/1.448118

52. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997). doi:10.1002/(SICI)1096-987X(199709)18:12<1463:AID-JCC4>3.0.CO;2-H

53. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981). doi:10.1063/1.328693

54. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007). doi:10.1063/1.2408420 Medline

55. R. Briones, C. Blau, C. Kutzner, B. L. de Groot, C. Aponte-Santamaría, GROmaρs: A GROMACS-Based Toolset to Analyze Density Maps Derived from Molecular Dynamics Simulations. *Biophys. J.* **116**, 4–11 (2019). doi:10.1016/j.bpj.2018.11.3126 Medline

56. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018). doi:10.1002/pro.3235 Medline

57. M. Igaev, C. Kutzner, L. V. Bock, A. C. Vaiana, H. Grubmüller, Automated cryo-EM structure refinement using correlation-driven molecular dynamics. *eLife* **8**, e43542 (2019). doi:10.7554/eLife.43542 Medline

58. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi:10.1038/nmeth.2019 Medline

59. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018). doi:10.7554/eLife.42166 Medline

60. D. Tegunov, tom_deconv, GitHub (2019); https://github.com/dtegunov/tom_deconv.