Title: Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein

Authors: Christine Toelzer[1,2#], Kapil Gupta[1,2#], Sathish K.N. Yadav[1,2#], Ufuk Borucu[1,2#], Andrew D. Davidson[3], Maia Kavanagh Williamson[3], Deborah K. Shoemark[1,2], Frederic Garzoni[4], Oskar Staufer[5,6,7,8], Rachel Milligan[3], Julien Capin[1,2], Adrian J. Mulholland[9], Joachim Spatz[5,6,7,8], Daniel Fitzgerald[10], Imre Berger[1,2,8,9*] and Christiane Schaffitzel[1,2*]

Affiliations:

[1]School of Biochemistry, University of Bristol, 1 Tankard's Close, Bristol BS8 1TD, UK.

[2]Bristol Synthetic Biology Centre BrisSynBio, 24 Tyndall Ave, Bristol BS8 1TQ, UK.

[3]School of Cellular and Molecular Medicine, University of Bristol, University Walk, Bristol, BS8 1TD, UK.

[4]Imophoron Ltd, St. Philips Central, Albert Rd, St. Philips, Bristol, BS2 0XJ, UK.

[5]Department for Cellular Biophysics, Max Planck Institute for Medical Research, Jahnstraße 29, 69120 Heidelberg, Germany.

[6]Institute for Physical Chemistry, Department for Biophysical Chemistry, University of Heidelberg, Im Neuenheimer Feld 253, 69120 Heidelberg, Germany.

[7]Max Planck School Matter to Life, Jahnstraße 29, D-69120 Heidelberg, Germany.

[8]Max Planck Bristol Centre for Minimal Biology, Cantock's Close, Bristol BS8 1TS, UK.

[9]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK.

[10]Geneva Biotech Sàrl, Avenue de la Roseraie 64, 1205, Geneva, Switzerland.

*Correspondence to: cb14941@bristol.ac.uk, imre.berger@bristol.ac.uk

[#]These authors contributed equally

Abstract:

COVID-19, caused by severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2), represents a global crisis. Key to SARS-CoV-2 therapeutic development is unraveling the mechanisms driving high infectivity, broad tissue tropism and severe pathology. Our 2.85Å cryo-EM structure of SARS-CoV-2 spike (S) glycoprotein reveals that the receptor binding domains (RBDs) tightly bind the essential free fatty acid (FFA) linoleic acid (LA) in three composite binding pockets. The pocket also appears to be present in the highly pathogenic coronaviruses SARS-CoV and MERS-CoV. LA binding stabilizes a locked S conformation giving rise to reduced ACE2 interaction *in vitro.* In human cells, LA supplementation synergizes with the COVID-19 drug remdesivir, suppressing SARS-CoV-2 replication. Our structure directly links LA and S, setting the stage for intervention strategies targeting LA binding by SARS-CoV-2.

One Sentence Summary:

The SARS-CoV-2 spike binds linoleic acid, a key molecule in inflammation, immune modulation and membrane fluidity**.**

Main Text:

Seven coronaviruses are known to infect humans. The four endemic human coronaviruses OC43, 229E, HKU1 and NL63 cause mild upper respiratory tract infections while pandemic virus SARS-CoV-2, and earlier SARS-CoV and MERS-CoV, can cause severe pneumonia with acute respiratory distress syndrome, multi-organ failure, and death (*1*,*2*).

SARS-CoV-2 has acquired functions that promote its harsh disease phenotype. SARS-CoV-2 causes severe inflammation and damage to endothelial cells in the heart, kidneys, liver and intestines, suggestive of a vascular infection rather than a purely respiratory disease (*3*, *4*). The attachment of SARS-CoV-2 to a host cell is initiated by the spike protein trimer (S), which decorates the outer surface of the virus, binding to its cognate receptor angiotensin-converting enzyme 2 (ACE2), with higher affinity than SARS-CoV (*5*-*7*). A S1/S2 polybasic furin protease cleavage site distinguishes SARS-CoV-2 from SARS-CoV or other closely related bat coronaviruses and serves to stimulate entry into host cells and cell-cell fusion (*5*, *8*, *9*). Inside the host cell, human coronaviruses remodel the lipid metabolism to facilitate virus replication (*10*). Infection by SARS-CoV-2 triggers an unusually impaired and dysregulated immune response (*11*) and a heightened inflammatory response (*12*) working in synergy with interferon production in the vicinity of infected cells to drive a feed-forward loop to upregulate ACE2 and further escalate infection (*13*).

In the search for additional functions that contribute to the pathology of infection, we determined the structure of the SARS-CoV-2 S glycoprotein by cryo-EM (Fig. 1). We produced SARS-CoV-2 S as a secreted trimer (*14*) in MultiBac (*15*) baculovirus-infected Hi5 insect cells (fig. S1, (*16*)). Highly purified protein was used for cryo-EM data collection (fig. S2, table S1). After 3D classification and refinement without applying symmetry (C1) we obtained a 3.0 Å closed

conformation from 136,405 particles and a 3.5 Å open conformation with one receptor-binding domain (RBD) in the up position from 57,990 particles (figs. S2, S3). C3 symmetry was applied to the closed conformation particle pool yielding a 2.85 Å map (Fig.1A, figs. S2, S3).

The structure of S displays the characteristic overall shape observed for coronavirus S proteins in the closed and open conformations (*17-19*) with the closed form (~70%) predominating in our data set (Fig.1A, figs. S2-S4). Model building of the closed form revealed additional density in the RBDs in our structure (Fig.1B). The tube-like shape of this density was consistent with a fatty acid, with size and shape similar to LA bound to other proteins (Fig.1B, fig. S5) (*20, 21*). Liquid chromatography coupled ESI-TOF mass spectrometry (LS-MS) analysis confirmed the presence of a compound with the molecular weight of LA in our purified sample (Fig.1C).

Hallmarks of FFA-binding pockets in proteins are an extended 'greasy' tube lined by hydrophobic amino acids which accommodates the hydrocarbon tail, and a hydrophilic, often positively charged anchor for the acidic headgroup of the FFA. In our structure, a hydrophobic pocket mostly shaped by phenylalanines forms a bent tube into which the LA fits well (Fig. 1D,E). The anchor for the headgroup carboxyl is provided by an arginine (R408) and a glutamine (Q409) from the adjacent RBD in the trimer, giving rise to a composite LA-binding site (Fig. 1E). We confirmed the presence of LA in all three binding pockets in the S trimer in the unsymmetrized (C1) closed structure (fig. S6). Similarly, masked 3D classification focusing on the RBD domains could not identify any unoccupied pockets (fig. S7).

Our S construct contains alterations as compared to native SARS-CoV-2 S namely addition of a trimerization domain and deletion of the polybasic cleavage site, neither of which alter S conformation appreciably (*14, 17*) (fig. S8). Glycosylation sites are located away from the LA-binding pocket and largely native in our structure (*7, 17*) (table S2). Thus, neither mutations nor

glycosylation are likely to impact the LA-binding pocket. We compared S and RBD produced in insect cells with mammalian produced S to identify any potential influence of differences in glycosylation on ACE2 binding by competition enzyme-linked immunosorbent assay (ELISA) (Fig. 2A). All three reagents bound ACE2 efficiently. We further confirmed ACE2 binding by S using SEC with purified proteins (Fig. 2B). The LA-binding pocket and the receptor binding motif (RBM) are distal and non-overlapping (Fig. 2C). Notably, in the LA-bound S the RBM is ordered and buried at the interface between RBDs while it was disordered in previously described SARS-CoV-2 S cryo-EM structures (*7*, *17*).

SARS-CoV-2 S can also adopt an open conformation (fig. S4) which is compatible with binding ACE2. In previous apo S cryo-EM structures about 60-75% of the S trimers were in the open conformation (*7*, *17*), contrasting our observation of ~70% in the closed conformation. This could be due to LA stabilizing the closed conformation, and if so LA would be expected to reduce ACE2 binding. We performed surface plasmon resonance (SPR) experiments with biotinylated ACE2 immobilized on a streptavidin-coated chip (Fig. 2D, fig. S9). We first determined the K_D of the RBD / ACE2 interaction to validate our assay. Our value (26 nM, fig. S9C) is in good agreement with previous studies (44 nM, (*22*)) obtained by SPR with the RBD immobilized and ACE2 as analyte. Apo S was prepared by applying Lipidex, the established method for removing lipids from lipid-binding proteins (*23*). A K_D of 0.7 nM was obtained for the apo S / ACE2 interaction (fig. S9A). For the LA-bound S / ACE2 interaction we obtained a K_D of 1.4 nM (fig. S9B). We consistently obtained a markedly reduced resonance unit (RU) signal for LA-bound S as compared to apo S at the same concentrations (Fig. 2D, fig. S9A,B). This correlates with the apo state having a higher percentage of S trimers in the open, ACE2-accessible conformation.

We characterized the affinity of the LA interaction both experimentally and computationally. Our SPR assays utilizing immobilized RBD yielded a binding constant of ~41 nM exhibiting a slow off-rate, consistent with tight binding of LA (fig. S10). Repeated molecular dynamics simulations of the entire locked LA-bound spike trimer (3 x 100 ns) using GROMACS-2019 (*24*) corroborated the persistence of stable interactions between LA and the spike trimer (Movies S1,S2). The affinity of LA binding to the spike trimer will likely be higher than to the RBD alone, taking into account polar headgroup interactions with R408 and Q409 of the adjacent RBD (Fig. 1E). The resolution of the RBDs in our open S cryo-EM structure was insufficient to either assign, or rule out, a ligand-bound pocket (fig. S3). However, the slow off-rate observed with the RBD monomer (fig. S10) suggests that LA binding could be maintained when the S trimer transiently converts into the open conformation. This is supported by our observation that LA was retained during S purification in spite of S trimers adopting the open form ~30% of the time (fig. S2) and by our MD simulations with a modelled ligand-bound open spike trimer (Movie S3) in which all three LAs remained bound over 500 ns.

Next, we investigated the effect of LA in experiments using live SARS-CoV-2 virus to infect human epithelial cells. Remdesivir is an RNA-dependent RNA polymerase inhibitor and the first anti-viral drug showing benefit in the treatment of COVID-19 in clinical trials, albeit with considerable side effects at the doses required (*25*). LA supplementation at 50-100 μM concentrations was previously shown to affect coronavirus entry and replication (*10*). We administered remdesivir at 20, 64 and 200 nM concentration, supplementing with 50 μM LA (Fig. 2E). Our results revealed synergy, with the dose of remdesivir required to suppress SARS-CoV-2 replication markedly reduced by adding LA (Fig. 2E,F).

We superimposed our LA-bound structure on previous SARS-CoV-2 apo S structures in the closed conformation (*7*, *17*) and identified a gating helix located directly at the entrance of the binding pocket (Fig. 3A-C). This gating helix, comprising Ty365 and Tyr369, is displaced by about 6 Å when LA is bound, thus opening the pocket (Fig.3A,B). In the apo SARS-CoV-2 S trimer (*7*, *17*), a gap between adjacent RBDs places the hydrophilic anchor residues ~10 Å away from the position of the LA headgroup (Fig. 3C). Upon LA binding, the adjacent RBD in the trimer moves towards its neighbor, and the anchor residues Arg408 and Gln409 lock down on the headgroup of LA (Fig. 3C,D). Overall, this results in a compaction of trimer architecture in the region formed by the three RBDs giving rise to a locked S structure (Fig. 3D, Movie S4).

We investigated whether the LA-binding pocket is conserved in the seven coronaviruses that infect humans (Fig. 4A, table S3). Sequence alignment shows that all residues lining the hydrophobic pocket and the anchor residues (Arg408/Gln409) in SARS-CoV-2 are fully conserved in SARS-CoV (Fig. 4A). Structural alignment of LA-bound RBDs within the trimer of SARS-CoV-2 and 'apo' SARS-CoV RBDs (*19*) reveals that the LA-binding pocket is present in SARS-CoV. The greasy tube is flanked by a gating helix as in SARS-CoV-2, with Arg395/Gln396 of SARS-CoV positioned 10Å and 11Å from the entrance, respectively, virtually identical to apo SARS-CoV-2 (Fig. 3C,4B). In MERS-CoV, the gating helix and hydrophobic residues lining the pocket are also present. Tyr365, Tyr369 and Phe374 are substituted by likewise hydrophobic leucines and a valine, respectively (Fig. 4A,C) (*19*). The Arg408/Gln409 pair is not conserved, however, we identify Asn501/Lys502 and Gln466 as potential anchor residues, located on a β-sheet and an α-helix within the adjacent RBD, up to 11Å away from the entrance (Fig. 4C). Thus, the greasy tube and hydrophilic anchor appear to be present in MERS-CoV, suggesting convergent evolution. In HCoV OC43, gating helix and hydrophobic residues lining the pocket are largely conserved, while

Tyr365, Tyr369 and Phe374 are replaced by methionines and alanine, respectively (Fig. 4A) (*18*). Arg413 is located on the same helix as Arg408/Gln409 in SARS-CoV-2 and could serve as a hydrophilic anchor (Fig. 4D). No gap exists in this presumed 'apo' form structure between the RBDs which appear already in the locked conformation (Fig. 4D, fig. S11 (*18*)). In HCoV HKU1, the hydrophobic residues are again largely conserved, but a charged residue (Glu375) is positioned directly in front of the entrance, obstructing access for a putative ligand (Fig. 4E) (*26*). The RBDs of HCoVs 229E and NL63 adopt a very different fold (fig. S13) (*27*, *28*), and many of the LA-binding residues are not present (Fig. 4A), hampering predictions of a binding site for fatty acids. In summary, we find four molecular features mediating LA binding to SARS-CoV-2, and potentially also SARS-CoV and MERS-CoV S proteins: a conserved hydrophobic pocket, a gating helix, amino acid residues pre-positioned to interact with the LA carboxy headgroup, and loosely packed RBDs in the 'apo' form. In contrast, in each of the four common circulating HCoVs, it appears that one or more of these four architectural prerequisites are lacking in the S protein structures (Fig. 4, figs. S11,S12). LA binding to SARS-CoV-2 S triggers a locking down of the hydrophilic anchor and a compaction of the RBD trimer (Fig. 3C,D). In addition to stabilizing the closed conformation, this could also help stabilize the S1 region comprising the N-terminal domain and the RBD. The RBM, central to ACE2 binding, appears to be conformationally preorganized in our structure (Fig. 2C) indicating a generally more rigid RBD trimer when LA is bound. While direct crosstalk in between the LA-binding pocket and the RBM is not apparent from our structure (Fig. 2C), the conformational changes in the RBD trimer (Fig. 3) could impact ACE2 docking and infectivity as indicated by our SPR assays showing reduced levels of S binding in the presence of LA (Fig. 2D). The S protein's tight binding of LA originates from a well-defined size and shape complementarity afforded by the pocket (Fig. 1B,D). The LA-binding pocket thus presents a

promising target for future development of small molecule inhibitors that, for example, could irreversibly lock S in the closed conformation and interfere with receptor interactions. It is noteworthy in this context that a fatty acid binding pocket was exploited previously to develop potent small molecule anti-viral drugs against rhinovirus, locking viral surface proteins in a conformation incompatible with receptor binding (*29*, *30*). These anti-virals were successful in human clinical trials (*31*, *32*).

A recent proteomic and metabolomic study of COVID-19 patient sera showed continuous decrease of FFAs including LA (*33*). Lipid metabolome remodeling is a common element of viral infection (*34*, *35*). For coronaviruses, the LA to arachidonic acid metabolic pathway was identified as central to lipid remodeling (*10*). We hypothesize that LA sequestration by SARS-CoV-2 could confer a tissue-independent mechanism by which pathogenic coronavirus infection may drive immune dysregulation and inflammation (*35*-*37*). Our findings provide a direct structural link between LA, COVID-19 pathology and the virus itself and suggest that both the LA-binding pocket within the S protein and the multi-nodal LA signaling axis, represent excellent therapeutic intervention points against SARS-CoV-2 infections.

References and Notes:

1. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
2. A. Zumla, J. F. Chan, E. I. Azhar, D. S. Hui, K. Y. Yuen, Coronaviruses - drug discovery and therapeutic options. *Nat Rev Drug Discov* **15**, 327-347 (2016).
3. Z. Varga *et al.*, Endothelial cell infection and endotheliitis in COVID-19. *Lancet* **395**, 1417-1418 (2020).
4. V. G. Puelles *et al.*, Multiorgan and Renal Tropism of SARS-CoV-2. *N Engl J Med* **383**, 590-592 (2020).
5. M. Hoffmann *et al.*, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278 (2020).
6. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* **5**, 562-569 (2020).
7. A. C. Walls *et al.*, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292 e286 (2020).
8. M. Hoffmann, H. Kleine-Weber, S. Pohlmann, A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol Cell* **78**, 779-784 e775 (2020).
9. S. Matsuyama *et al.*, Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc Natl Acad Sci U S A* **117**, 7001-7003 (2020).
10. B. Yan *et al.*, Characterization of the Lipidomic Profile of Human Coronavirus-Infected Cells: Implications for Lipid Metabolism Remodeling upon Coronavirus Replication. *Viruses* **11**, 73 (2019).

11. C. Qin *et al.*, Dysregulation of immune response in patients with COVID-19 in Wuhan, China. *Clin Infect Dis* **71**, 762-768 (2020).

12. M. Z. Tay, C. M. Poh, L. Renia, P. A. MacAry, L. F. P. Ng, The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol* **20,** 363-374 (2020).

13. C. G. K. Ziegler *et al.*, SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. *Cell* **181**, 1016-1035 e1019 (2020).

14. F. Amanat *et al.*, A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat Med* **26**, 1033-1036 (2020).

15. D. J. Fitzgerald *et al.*, Protein complex expression by using multigene baculoviral vectors. *Nat Methods* **3**, 1021-1032 (2006).

16. To produce spike glycoprotein, we used here baculovirus-infected insect cells (Hi5) cultured in ESF921 media which contains cod liver oil as a nutrient supplement (Corey Jacklin, Expression Systems, personal communication). Cod liver oil comprises hundreds of FFAs including LA (*38*) which could be a possible source of the fatty acid. In parallel to our work, cryo-EM structures were determined of spike expressed in a mammalian system (HEK293) in serum and protein-free media (*39, 40*). The tube-shaped density we identified in our study is also present in those structures (but was not assigned or interpreted). We conclude that LA binding in the SARS-CoV-2 S pocket is thus not dependent on the expression system or media used.

17. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).

18. M. A. Tortorici *et al.*, Structural basis for human coronavirus attachment to sialic acid receptors. *Nat Struct Mol Biol* **26**, 481-489 (2019).

19. Y. Yuan *et al.*, Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun* **8**, 15092 (2017).

20. J. M. Gullett, M. G. Cuypers, M. W. Frank, S. W. White, C. O. Rock, A fatty acid-binding protein of *Streptococcus pneumoniae* facilitates the acquisition of host polyunsaturated fatty acids. *J Biol Chem* **294**, 16416-16428 (2019).

21. J. Wang, E. J. Murphy, J. C. Nix, D. N. M. Jones, *Aedes aegypti* Odorant Binding Protein 22 selectively binds fatty acids through a conformational change in its C-terminal tail. *Sci Rep* **10**, 3300 (2020).

22. J. Shang *et al.*, Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224 (2020).

23. J. F. Glatz, J. H. Veerkamp, Removal of fatty acids from serum albumin by Lipidex 1000 chromatography. *J Biochem Biophys Methods* **8**, 57-61 (1983).

24. H. J. C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43-56 (1995).

25. J. H. Beigel *et al.*, Remdesivir for the Treatment of Covid-19 - Preliminary Report. *N Engl J Med,* NEJMoa2007764 (2020).

26. X. Ou *et al.*, Crystal structure of the receptor binding domain of the spike glycoprotein of human betacoronavirus HKU1. *Nat Commun* **8**, 15216 (2017).

27. Z. Li *et al.*, The human coronavirus HCoV-229E S-protein structure and receptor binding. *Elife* **8**, e51230 (2019).

28. A. C. Walls *et al.*, Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol* **23**, 899-905 (2016).

29. J. Badger *et al.*, Structural analysis of a series of antiviral agents complexed with human rhinovirus 14. *Proc Natl Acad Sci U S A* **85**, 3304-3308 (1988).

30. M. A. Oliveira *et al.*, The structure of human rhinovirus 16. *Structure* **1**, 51-68 (1993).

31. V. Casanova, F. H. Sousa, C. Stevens, P. G. Barlow, Antiviral therapeutic approaches for human rhinovirus infections. *Future Virol* **13**, 505-518 (2018).

32. K. K. W. To, C. C. Y. Yip, K. Y. Yuen, Rhinovirus - From bench to bedside. *J Formos Med Assoc* **116**, 496-504 (2017).

33. B. Shen *et al.*, Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **182**, 59-72 (2020).

34. C. M. Goodwin, S. Xu, J. Munger, Stealing the Keys to the Kitchen: Viral Manipulation of the Host Cell Metabolic Network. *Trends Microbiol* **23**, 789-798 (2015).

35. J. W. Schoggins, G. Randall, Lipids in innate antiviral defense. *Cell Host Microbe* **14**, 379-385 (2013).

36. M. M. Zaman *et al.*, Linoleic acid supplementation results in increased arachidonic acid and eicosanoid production in CF airway cells and in cftr-/- transgenic mice. *Am J Physiol Lung Cell Mol Physiol* **299**, L599-606 (2010).

37. I. Kimura, A. Ichimura, R. Ohue-Kitano, M. Igarashi, Free Fatty Acid Receptors in Health and Disease. *Physiol Rev* **100**, 171-210 (2020).

38. S. Hauff, W. Vetter, Quantitation of cis- and trans-monounsaturated fatty acids in dairy products and cod liver oil by mass spectrometry in the selected ion monitoring mode. *J Agric Food Chem* **57**, 3423-3430 (2009).

39. Y. Cai *et al.*, Distinct conformational states of SARS-CoV-2 spike protein. *Science* (2020).

40. A. G. Wrobel *et al.*, SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat Struct Mol Biol* **27**, 763-767 (2020).

41. I. Berger, D. J. Fitzgerald, T. J. Richmond, Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol* **22**, 1583-1587 (2004).

42. M. M. Vork, J. F. Glatz, D. A. Surtel, G. J. van der Vusse, Assay of the binding of fatty acids by proteins: evaluation of the Lipidex 1000 procedure. *Mol Cell Biochem* **98**, 111-117 (1990).

43. Q. Wang *et al.*, Protocols and pitfalls in obtaining fatty acid-binding proteins for biophysical studies of ligand-protein and protein-protein interactions. *Biochem Biophys Rep* **10**, 318-324 (2017).

44. M. Fairhead, M. Howarth, Site-specific biotinylation of purified proteins using BirA. *Methods Mol Biol* **1266**, 171-184 (2015).

45. S. Q. Zheng *et al.*, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).

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

47. S. H. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530 (2012).

48. T. D. Goddard, C. C. Huang, T. E. Ferrin, Visualizing density maps with UCSF Chimera. *J Struct Biol* **157**, 281-287 (2007).

49. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

50. T. C. Terwilliger, O. V. Sobolev, P. V. Afonine, P. D. Adams, Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr D Struct Biol* **74**, 545-559 (2018).

51. N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* **65**, 1074-1080 (2009).

52. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).

53. V. B. Chen *et al.*, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).

54. B. A. Barad *et al.*, EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat Methods* **12**, 943-946 (2015).

55. J. L. Daly *et al.*, Neuropilin-1 is a host factor for SARS-CoV-2 infection. *bioRxiv, doi: https://doi.org/10.1101/2020.06.05.134114*, (2020).

56. A. D. Davidson *et al.*, 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).

57. V. M. Corman *et al.*, Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* **25**, 2000045 (2020).

58. R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. P. Thornton, PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.* **26**, 283-291 (1993).

59. A. W. Sousa da Silva, W. F. Vranken, ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res Notes* **5**, 367 (2012).

60. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-1174 (2004).

61. D. A. Case *et al.*, The Amber biomolecular simulation programs. *J Comput Chem* **26**, 1668-1688 (2005).

62. K. Lindorff-Larsen *et al.*, Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950-1958 (2010).

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

Acknowledgments: We thank all members of the Berger and Schaffitzel teams as well as Gunjita Singh, Yohei Yamauchi and David Matthews (University of Bristol, UK) for their assistance. We thank Florian Krammer (Icahn School of Medicine, USA) for kindly sharing expression plasmids. We thank Veronica Chang and Radu Aricescu (MRC-LMB, UK) for kind gift of S expressed in HEK293. We are indebted to Adam Finn (Bristol UNCOVER Group and Children's Vaccine Centre, Bristol Medical School), Jeremy Tavaré (School of Biochemistry, Bristol), Kathleen Gillespie (Diabetes and Metabolism Unit, Southmead Hospital, Univ. of Bristol) and Donald Fitzgerald MD (Quest Imaging Medical Associates, USA) for helpful discussions and careful

reading of the manuscript. We thank Simon Burbidge, Thomas Batstone and Matt Williams for computation infrastructure support. We would like to thank the Advanced Computing Research Centre (ACRC) at the University of Bristol for access to BlueCryo, BlueCrystal Phase 4 and BlueGEM, and the UK HECBioSim for access to the UK supercomputer, ARCHER. We are particularly grateful to Thiru Thangarajah (Genscript) for early access to Genscript's cPass™ SARS-CoV-2 Neutralization Antibody Detection/Surrogate Virus Neutralization Test Kit (L00847). We thank Sebastian Fabritz and the Core Facility for Mass Spectrometry at the Max Planck Institute for Medical Research for their support on MS measurements.

Funding: This research received support from the Elizabeth Blackwell Institute for Health Research and the EPSRC Impact Acceleration Account EP/R511663/1, University of Bristol, from BrisSynBio a BBSRC/EPSRC Research Centre for synthetic biology at the University of Bristol (BB/L01386X/1) (to I.B., A.J.M., D.K.S.) and from the BBSRC (BB/P000940/1) (to C.S. and I.B.). This work received generous support from the Oracle Higher Education and Research program to enable cryo-EM data processing using Oracle's high-performance public cloud infrastructure (https://cloud.oracle.com/en_US/cloud-infrastructure) and the EPSRC through a COVID-19 project award via HECBioSim to access ARCHER (A.J.M.). We acknowledge support and assistance by the Wolfson Bioimaging Facility and the GW4 Facility for High-Resolution Electron Cryo-Microscopy funded by the Wellcome Trust (202904/Z/16/Z and 206181/Z/17/Z) and BBSRC (BB/R000484/1). The authors would like to acknowledge support of the University of Bristol's Alumni and Friends, which funded the ImageXpress Pico Imaging System. O.S. acknowledges support from the Elisabeth Muerer Foundation, the Max Planck School Matter to Life and the Heidelberg Biosciences International Graduate School. A.D.D. is supported by the United States Food and Drug Administration (HHSF223201510104C) and UK Research and

Innovation / Medical Research Council (MRC) (MR/V027506/1). M.K.W is supported by MRC grants MR/R020566/1 and MR/V027506/1 (awarded to A.D.D). A.J.M. is supported the British Society for Antimicrobial Chemotherapy (BSAC-COVID-30) and the EPSRC (EP/M022609/1, CCP-BioSim). I.B. acknowledges support from the EPSRC Future Vaccine Manufacturing and Research Hub (EP/R013764/1). C.S. and I.B. are Investigators of the Wellcome Trust (210701/Z/18/Z; 106115/Z/14/Z).

Author Contributions: C.S. and I.B. conceived and guided the study. F.G., K.G. and J.C. produced, purified and analyzed sample, K.G. carried out biochemical experiments, S.K.N.Y. and U.B. prepared grids and collected EM data, S.K.N.Y., U.B., K.G and C.T. carried out image analysis and model building. A.D.D., M.K.W. and R.M. performed all live virus CL3 work and analyzed data. D.K.S. and A.J.M. performed all MD simulations. O.S. and J.S. performed and interpreted mass spectrometry. C.T., K.G., D.F., I.B. and C.S. interpreted results. D.F., I.B. and C.S. wrote the manuscript with input from all authors.

Competing interests: The authors declare competing interests. I.B and F.G. report shareholding in Imophoron Ltd. unrelated to this Correspondence. I.B. and D.F. report shareholding in Geneva Biotech Sàrl related to this Correspondence. Patent applications describing methods and material compositions based on the present observations have been filed.

Data and materials availability: Datasets generated during the current study have been deposited in the Electron Microscopy Data Bank (EMDB) under accession numbers EMD-11145 (C3 closed conformation), EMD-11144 (C1 closed conformation) and EMD-11146 (open conformation), and in the Protein Data Bank (PDB) under accession numbers: 6ZB5 (C3 closed conformation) and 6ZB4 (C1 closed conformation). Reagents are available from I.B. and C.S. under a material transfer agreement with the University of Bristol.

Supplementary Materials:

Materials and Methods

Figures S1-S12

Tables S1-S3

Movies S1-S4

References (*41-63*) are part of the Supplement and not cited in the Main Text.

Figure Legends

Figure 1. Cryo-EM structure of SARS-CoV-2 spike linoleic acid complex. (**A**) Cryo-EM density of spike trimer (left). Monomers in cyan, green and pink, respectively. The structure in a cartoon representation in a front (middle) and top view (right). Bound LA illustrated as spheres colored in orange. One LA-binding pocket is boxed in red. (**B**) Composite LA-binding pocket formed by adjacent RBDs. Tube-shaped EM density is shown. (**C**) LC-MS of purified S. Chemical structure and molecular weight of LA (top), C4 column elution profile (middle) and ESI-TOF of wash solution (grey) and C4 peak elution fraction (black) with peak molecular weight indicated (bottom). (**D**) Hydrophobic LA-binding pocket in a surface representation illustrating excellent fit of bound LA (orange, sticks and balls representation). Blue and red indicate positive and negative surface charge, respectively. (**E**) LA interactions with amino acids in the binding pocket. The acidic LA headgroup is in the vicinity of an arginine (408) and a glutamine (409).

Figure 2. Functional characterization of LA-bound SARS-CoV-2 spike. (**A**) Insect cell (Hi5) expressed spike (dark blue bars), insect cell expressed RBD (light blue bars) and mammalian (HEK293) expressed spike (white bars) in competition ELISAs utilizing immobilized ACE2. Error bars: standard deviations, three replicates. **(B)** LA-bound SARS-CoV-2 spike protein interaction with ACE2 was analyzed by size exclusion chromatography evidencing complex formation. SEC profiles (left) for ACE2 (yellow, III), LA-bound spike (green, II) and a mixture of ACE2 and LA-bound spike (orange, I) are shown. Peak fractions (I.-III.) were analyzed by SDS-PAGE evidencing the expected proteins (right). **(C)** LA-bound S glycoprotein trimer top view with RBDs shown in cyan, green and pink respectively (left). In each RBD subunit the motif responsible for ACE2 binding (RBM) is in red, and LA is shown as spheres colored in orange. A close-up view

into the cyan RBD (right) shows that the RBM is fully ordered, and LA and RBM are not in direct contact. **(D)** SPR analysis of the binding of LA-bound spike trimer (orange curves) and apo spike trimer (green curves) to immobilized ACE2. Apo spike and LA-bound spike were diluted to concentrations of 40 nM and 160 nM, respectively. Black lines correspond to a global fit using a 1:1 binding model. **(E)** Synergistic effect of LA and remdesivir on SARS-CoV-2 viral replication. Effects of varying doses of remdesivir +/- 50 µM LA on virus infection are shown. Human Caco-2 ACE2+ cells were infected with SARS-CoV-2 and then treated with varying doses of remdesivir +/- 50 µM LA. At 96h post-infection, cells were fixed and infected cells detected by immunofluorescence assay using an anti-N antibody (green). Cell nuclei were stained by DAPI (blue). Representative images corresponding to the remdesivir dose range 20 – 200 nM are shown. **(F)** The amount of extracellular virus present in wells (n=3) at the dose combinations shown was determined by qRT-PCR.

Figure 3. Comparison of LA-bound and apo S structures. (A) Superimposition of LA-bound SARS-CoV-2 (cyan RBD) and the ligand-free 'apo' RBD (grey) (PDBID 6VXX, (*7*)). The gating helix at the entrance of the hydrophobic pocket moves by 6 Å in the presence of LA. Tyr365 and Tyr369 swing away avoiding clashes with LA (orange). Black arrows indicate the rearrangements. **(B)** The same structure as panel A rotated by 90º as indicated, showing the entrance of the hydrophobic pocket. **(C)** Formation of a composite LA-binding pocket by two adjacent RBDs in LA-bound S involves a ~5 Å movement of RBD2 (green) towards RBD1 (cyan) compared to apo S (grey). **(D)** Superimposition of the RBD trimer of apo S (grey) and LA-bound S (RBD1 cyan, RBD2 green, RBD3 pink, LA orange) is shown (left). The individual RBD trimers are depicted

for LA-bound S (right, top) and apo S (right, bottom) with RBDs boxed in black, highlighting the compaction of RBDs in the LA-bound S structure.

Figure 4. Human coronavirus RBD architectures. (**A**) Alignments of the seven CoV strains that can infect humans, highlighting conserved residues. Residues lining the hydrophobic pocket are underlaid (cyan). Gating helix residues are marked (purple). Residues positioned to interact with the LA polar headgroup are underlaid in green. Glu375 in HKU1 is underlaid in red (cf. panel E). **(B)** Superimposition of RBD1 of LA-bound SARS-CoV-2 (RBD2 in green) with RBD1 of ligand free 'apo' SARS-CoV (RBD1 and RBD2 magenta, PDBID 5X58 (*19*)) indicates a conservation of the composite binding pocket. **(C)** Superimposition of RBD1 of LA-bound SARS-CoV-2 (RBD2 is omitted for clarity) with RBD1 of MERS-CoV (RBD1 and RBD2 forest green, PDBID 5X5F (*19*)). **(D)** Superimposition of RBD1 of LA-bound SARS-CoV-2 (RBD2 in green) with RBD1 of OC43 (RBD1 and RBD2 purple, PDBID 6NZK, (*18*)). **(E)** Superimposition of LA-bound SARS-CoV-2 RBD with HKU1 RBD (brown, PDBID 5GNB (*26*)). LA is omitted in SARS-CoV-2 RBD for clarity.

Science
AAAS

Supplementary Materials for

Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein

Christine Toelzer[1,2#], Kapil Gupta[1,2#], Sathish K.N. Yadav[1,2#], Ufuk Borucu[1,2#], Andrew D. Davidson[3], Maia Kavanagh Williamson[3], Deborah K. Shoemark[1,2], Frederic Garzoni[4], Oskar Staufer[5,6,7,8], Rachel Milligan[3], Julien Capin[1,2], Adrian J. Mulholland[9], Joachim Spatz[5,6,7,8], Daniel Fitzgerald[10], Imre Berger[1,2,8,9*] and Christiane Schaffitzel[1,2*]

[#]These authors contributed equally to this work.

[*]Correspondence to: cb14941@bristol.ac.uk, imre.berger@bristol.ac.uk

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Materials and Methods

Protein production.

SARS-CoV-2 spike protein. The pFastBacDual plasmid for SARS-CoV-2 S ectodomain for expression in insect cells (*14*) was kindly provided by Florian Krammer (Icahn School of Medicine, USA). In this construct, S comprises amino acids 1 to 1213 and is fused to a C-terminal thrombin cleavage site followed by a T4-foldon trimerization domain and a hexahistidine affinity purification tag. The polybasic cleavage site has been deleted (RRAR to A) in the construct (*14*). Protein was produced with the MultiBac baculovirus expression system (Geneva Biotech, Geneva, Switzerland) (*41*) in Hi5 cells using ESF921 media (Expression Systems Inc.). Supernatants from transfected cells were harvested 3 days post-transfection by centrifugation of the culture at 1,000g for 10 min followed by another centrifugation of supernatant at 5,000g for 30 min. The final supernatant was incubated with 7 ml HisPur Ni-NTA Superflow Agarose (Thermo Fisher Scientific) per 3 liters of culture for 1h at 4ºC. Subsequently, a gravity flow column was used to collect the resin bound with SARS-CoV-2 spike protein, the resin was washed extensively with wash buffer (65 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 7.5), and the protein was eluted using a step gradient of elution buffer (65 mM NaH_2PO_4, 300mM NaCl, 235mM imidazole, pH 7.5). Elution fractions were analyzed by reducing SDS-PAGE and fractions containing SARS-CoV-2 spike protein were pooled, concentrated using 50 kDa MWCO Amicon centrifugal filter units (EMD Millipore) and buffer-exchanged in phosphate-buffered saline (PBS) pH 7.5. Concentrated SARS-CoV-2 S was subjected to size exclusion chromatography (SEC) using a Superdex 200 increase 10/300 column (GE Healthcare) in PBS pH 7.5. Peak fractions from SEC were analyzed by reducing SDS-PAGE and negative stain electron microscopy (EM); fraction 7 was used for cryo-EM (fig.S1).

Lipidex-treated SARS-CoV-2 spike protein. Purified SARS-CoV-2 spike protein was treated with lipidex-1000 resin (Perkin Elmer; cat no. 6008301) according to published protocols (*23, 42, 43*). Briefly, S was incubated with lipidex-1000 resin pre-equilibrated in PBS pH 7.5 overnight at 4°C on a roller shaker. Subsequently, lipidex-treated S was separated from the resin using a gravity flow column. Integrity of the protein was confirmed by size exclusion chromatography (SEC) using a S200 10/300 increase column (GE Healthcare) and SDS-PAGE.

ACE2 protein. The gene encoding for the ACE2 ectodomain (amino acids 1 to 597) was codon optimized for insect cell expression, synthesized (Genscript Inc., New Jersey USA) and inserted into pACEBac1 plasmid (Geneva Biotech, Geneva, Switzerland). The resulting construct pACEBac1_MelACE2his construct contains an N-terminal melittin signal sequence for secretion and a C-terminal octa-histidine affinity purification tag. ACE2 protein was produced and purified following the same protocol as for spike protein.

Biotinylated ACE2 protein. ACE2 was modified by inserting a synthetic DNA between the *Stu*I and *Hind*III restriction sites in pACEBac1-MelACE2his giving rise to a modified C-terminus including a linker (amino acid sequence GGSGSG) followed by avi tag (amino acid sequence GLNDIFEAQKIEWHE), a second linker (amino acid sequence GGSGG) and an octa-histidine tag for purification. Avi-tagged ACE2 was produced using the same protocol as for ACE2 protein. Biotinylation was achieved by incubation with BirA in the presence of biotin according to established protocols (*44*). Remaining free biotin and BirA were removed by purifying biotinylated ACE2 using Ni-NTA Superflow Agarose followed by size exclusion chromatography (SEC) using a S200 10/300 increase column (GE Healthcare).

Receptor Binding domain (RBD) and biotinylated RBD. The gene encoding for the SARS-CoV-2 receptor binding domain (amino acids R319 to F541), fused at its N-terminus to the native spike signal sequence (amino acid sequence MFVFLVLLPLVSSQ), was codon optimized for insect cell expression, synthesized (Genscript Inc., New Jersey USA) and inserted into pACEBac1 plasmid (Geneva Biotech, Geneva, Switzerland). The resulting construct pACEBac1-RBDhis comprises a C-terminal octa-histidine affinity purification tag. Protein was produced and purified as described for spike trimer. Biotinylated RBD was generated by adding an avi tag for BirA mediated biotinylation as described above for biotinylated ACE2. Lipidex-treated biotinylated RBD was then prepared as described above for lipidex-treated S.

Negative stain sample preparation and microscopy

4 μL of 0.05 mg/mL SARS-CoV-2 spike protein was applied onto a freshly glow discharged (1 min at 10 mA) CF300-Cu-50 grid (Electron Microscopy Sciences), incubated for 1 min, and manually blotted. 4 μL of 3% Uranyl Acetate was applied onto the same grid and incubated for 1 min before the solution was blotted off. The grid was loaded onto a FEI Tecnai 12 120 kV BioTwin Spirit TEM. Images were acquired at a nominal magnification of 49,000x.

Cryo-EM sample preparation and data collection

4 μL of 1.25 mg/mL SARS-CoV-2 spike protein was loaded onto a freshly glow discharged (2 min at 4 mA) Quantifoil R1.2/1.3 carbon grid (Agar Scientific), blotted using a Vitrobot MarkIV (Thermo Fisher Scientific) at 100% humidity and 4°C for 2 s, and plunge frozen. Data were acquired on a FEI Talos Arctica transmission electron microscope operated at 200 kV and equipped with a Gatan K2 Summit direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV using the EPU software.

Data were collected in super-resolution at a nominal magnification of 130,000x with a virtual pixel size of 0.525 Å. The dose rate was adjusted to 6.1 counts/physical pixel/s. Each movie was fractionated in 55 frames of 200 ms. 3289 micrographs were collected in a single session with a defocus range comprised between -0.8 and -2 μm.

Cryo-EM data processing

The dose-fractionated movies were gain-normalized, aligned, and dose-weighted using MotionCor2 (*45*). Defocus values were estimated and corrected using CTFFIND4 (*46*). 611,879 particles were automatically picked using Relion 3.0 software (*47*). Reference-free 2D classification was performed to select well-defined particles, after four rounds of 2D classification a total of 386,510 good particles were selected for further 3D classification. An initial model was generated using 50,000 particle images in Relion 3.0, and the selected particles from 2D classification were subjected to 3D classification using 8 classes. Classes 4 and 6 (fig. S2), showing prominent features representing a total of 202,082 particles were combined and used for 3D refinement. The 3D refined particles were then subjected to a second round of 3D classification using 5 classes. Class 4 and 5 were combined for the closed conformation map, class 3 represented the open conformation, comprising 136,405 and 57,990 particles respectively. The selected maps were subjected to 3D refinement without applying any symmetry with respective 3D models. The

maps were subsequently subjected to local defocus correction and Bayesian particle polishing in Relion 3.0. Global resolution and B factor (-89 Å^2 and -116 Å^2 for closed and open maps respectively) of the map were estimated by applying a soft mask around the protein density, using the gold standard Fourier shell correlation (FSC) = 0.143 criterion, resulting in an overall resolution of 3.03 Å and 3.7 Å respectively. The open map was imported into Relion 3.1 software, where it was further corrected for higher-order aberrations and magnification anisotropy. This yielded a final resolution of 3.5 Å (B factor of -87 Å^2). C3 symmetry was applied to the closed conformation map using Relion 3.0, followed by 3D classification using 3 classes. Class 3 with 217,815 particles was selected for CTF refinement and Bayesian polishing, yielding a final resolution of 2.85 Å (B factor of -86.8 Å^2). Local resolution maps were generated using Relion 3.0 (fig. S3).

Cryo-EM model building and analysis

UCSF Chimera (*48*) and Coot (*49*) were used to fit atomic models (PDBID 6VXX, (*7*)) into the cryo-EM map. The model was subsequently manually rebuilt using Coot and the closed conformation map. This closed conformation model was used to build features specific of the open conformation map and to further improve the model by using the C3-symmetrized map. To guide the model building, sharpened (*50*) and unsharpened maps were used in different steps of the process. N-linked glycans were hand-built into the density where visible, and restraints for non-standard ligands were generated with eLBOW (*51*). The model for the closed conformation was real space refined with Phenix (*52*) and the quality was additionally analyzed using MolProbity (*53*) and EMRinger (*54*), to validate the stereochemistry of the components. The quasi-atomic model of the open conformation was generated by first fitting the atomic model of the closed conformation into the open cryo-EM map using UCSF Chimera. We then used COOT and UCSF Chimera to move one RBD into the open conformation, by aligning this RBD to the atomic model of a published open form of SARS-CoV-2 S (PDBID 6VSB, (*17*)). Finally, this open model was fitted into the cryo-EM open map with COOT and UCSF Chimera. Figures were prepared using UCSF Chimera and PyMOL (Schrödinger, Inc.).

Masked 3D classification in Relion

The refined C3 particles stack was expanded 3-fold according to C3 symmetry in Relion. The symmetry expanded particle stack was then used as input for the masked 3D classification with the focus mask corresponding to NTD-RBD region created by one NTD and one RBD from neighbouring subunits. The same procedure was performed with refined C1 particles without particle expansion, applying 3 masks, one for each NTD-RBD pair. Masked 3D classification was performed without alignment using 5 classes for C3 particle and 3 classes for C1 particle data sets. Visual inspection of the 3D classes showed ~80% of the RBD's clustered in class 3 in C3 and in class 1 in C1 with LA bound in the hydrophobic pocket.

Mass spectrometry analysis

For mass spectrometry analysis, a Bruker maXis II ETD quadrupole-time-of-flight instrument with electrospray ionization (ESI) coupled to a Shimadzu Nexera HPCL was used. In order to extract the fatty acids from the protein sample, a chloroform extraction protocol was performed. For this,

100 µl of the protein sample was mixed with 400 µl chloroform for 2 hours on a horizontal shaker in a teflon-sealed glass vial at 25°C. The organic phase was then transferred to a new glass vial and the chloroform was evaporated for 30 min in a desiccator. Subsequently, 50 µl of a 20% (v/v) acetonitrile solution was added to dissolve the fatty acids. From this solution, a 1:100 dilution in 20% (v/v) acetonitrile was prepared and 20 µl were injected for LC-MS analysis. The samples were passed over a Phenomenex C4 Aeris column (100 × 2.1 mm, 3.6 µm, 200 Å) heated to 50°C using a gradient of 20% A to 98% B in 5 min (solvent A: H_2O + 0.1% formic acid; solvent B/: ACN + 0.1% formic acid). The system was operated with Bruker's O-TOF Control (V4.1) and Hystar (V4.1) software and data analysis as well as data post-processing was performed with Bruker's DataAnalysis software (V4.4, SR1).

ELISA activity assay

The SARS-CoV-2 sVNT kit was obtained from GenScript Inc. (New Jersey, USA) for ELISA activity assays. Serial dilution series were prepared (from 0-4096 nM) of purified SARS-CoV-2 spike proteins and RBD in PBS pH 7.5. This dilutions series was mixed with the same amount of HRP-RBD and incubated at 37ºC for 30 min. The mixtures were then added to ELISA plate wells, coated with ACE2 according to the manufacturer's protocol. Triplicates of each sample were made. The ELISA plate was then incubated at 37ºC for 15 min and subsequently washed 4 times with wash buffer. The signal was then developed by adding TMB solution to each well, incubating 15 minutes in the dark, followed by adding stop solution. The absorbance at 450 nm was immediately recorded. The data was plotted using Microsoft excel. The standard deviation of triplicates was added as error bars.

SARS-CoV-2 spike / ACE2 interaction analysis, size exclusion chromatography (SEC)

Purified SARS-CoV-2 spike and ACE2 proteins were combined in PBS pH 7.5 at a molar ratio of 1:1.5 spike trimer to ACE2 monomer. The mixture was incubated on ice for 2 hours and subjected to SEC using a Superdex 200 increase 10/300 column (GE Healthcare) equilibrated in PBS pH 7.5. Purified SARS-CoV-2 spike protein and ACE2 proteins were individually run on the same column in the same buffer as controls. The contents of peak fractions were analyzed by reducing SDS-PAGE.

Surface plasmon resonance (SPR) experiments

Interaction experiments using surface plasmon resonance (SPR) were carried out with a Biacore T200 system (GE Healthcare) according to the manufacturer's protocols and recommendations. For RBD / ACE2 interaction measurements, purified biotinylated ACE2 was immobilized on a streptavidin coated (SA) chip (GE Healthcare) at ~300 response units (RUs). Purified RBD was serially diluted and injected at concentrations 80 nM, 120 nM, 140 nM and 160 nM. Measurements were carried out in triplicates. For measuring ACE2 interactions with S and lipidex-treated S, respectively, purified biotinylated ACE2 was immobilized at ~50 RUs and purified proteins serially diluted and injected at concentrations of 40 nM, 60 nM, 80 nM, 120 nM and 160 nM. For LA / RBD measurements, biotinylated RBD was immobilized at ~3800 RUs as required, according to our calculations, for a sufficiently prominent signal upon binding of the small molecule LA (MW 280 Da). LA was purchased as a sodium salt (Sigma Aldrich, cat no. L8134) and dissolved

in PBS pH 7.5 at a concentration of 10 µM, serially diluted and injected at concentrations 4 µM, 6 µM, 8 µM and 10 µM. Running buffer for all measurements was PBS buffer pH 7.5. Sensorgams were analyzed and K_D, k_{on} and k_{off} values in all SPR experiments determined with the Biacore Evaluation Software (GE Healthcare), fitting the raw data using a 1:1 binding model.

Live SARS-CoV-2 experiments

Cell culture. A Vero E6 cell line modified to constitutively express the serine protease TMPRSS2 (Vero E6/TMPRSS2; (*9*) (obtained from NIBSC, UK) and the human gut epithelial cell line Caco-2 expressing ACE2 (Caco-2-ACE2) (a kind gift of Dr Yohei Yamauchi, University of Bristol) were cultured at 37ºC in 5% CO_2 in Dulbecco's modified Eagle's medium plus GlutaMAX (DMEM, Gibco, ThermoFisher) supplemented with 10% fetal bovine serum (FBS, Gibco, ThermoFisher) and 0.1 mM non-essential amino acids (NEAA, Sigma Aldrich).

SARS-CoV-2 viral propagation. A stock of infectious virus was produced by inoculation of Vero E6/TMPRSS cells with the SARS-CoV-2 isolate SARS-CoV-2/human/Liverpool/REMQRQ0001/2020 (*55*) in infection medium (Eagle's minimum essential medium plus GlutaMAX (MEM, Gibco, ThermoFisher) supplemented with 2% FBS and NEAA). Cells were incubated at 37ºC in 5% CO_2 for 72 hours before being harvested and filtered through a 0.2 µm filter. The viral titer was determined by $TCID_{50}$ on Vero E6 cells as previously described (*56*).

SARS-CoV-2 infection assay. Caco-2-ACE2 cells were seeded for 18 hours in DMEM in µClear 96-well Microplates (Greiner Bio-one). The culture supernatant was removed and cells were infected with SARS-CoV-2 diluted in infection medium to a MOI of 0.1 for 1 hour. The virus was removed and replaced with infection medium containing eight-fold serial dilutions of LA, remdesivir (RMD; Cambridge Bioscience, UK) or LA+RMD. After 96 hours incubation at 37ºC in 5% CO_2, supernatants were collected for qRT-PCR and cells fixed in 4% PFA for 60 minutes for image analysis. All work with infectious SARS-CoV-2 virus was conducted in a Class III microbiological safety cabinet in a containment level 3 facility at the University of Bristol.

Cell viability assay. Cell viability and compound toxicity were determined by MTT assay. Samples were treated identically to viral infection assays, with the exclusion of the viral inoculation step. After 96 hours incubation at 37ºC, media was removed and replaced by Thiazolyl Blue Tetrazolium Bromide (MTT; Sigma Aldrich) solution (1 mg/ml) and incubated for 3 hours at 37ºC. The MTT reagent was then removed and cells lysed by addition of 100% ethanol. Samples were analyzed by reading the absorbance at 600 nm.

qRT-PCR. Culture supernatants were collected and RNA extracted using the QIAamp Viral RNA Isolation kit (Qiagen) according to the manufacturer's instructions with elution of RNA in 60 µl of AVE solution. 5 µl of each sample was used for qRT-PCR analysis using a TaqPath™ 1-Step RT-qPCR Master Mix, CG (Applied Biosystems, ThermoFisher) following the manufacturer's instructions and the previously described SARS-CoV-2 E gene primer set (*57*) using a Mx3005P Real-Time PCR System (Agilent). SARS-CoV-2 RNA genome equivalence was determined using the MxPro qPCR Software (Agilent) and RNA extracted from titered virus and a AccuPlexTM SARS-CoV-2 Reference Material Kit (SeraCare, Milford, US) as standards.

Immunofluorescence analysis. Fixed cells were stained for DAPI (Sigma Aldridge) and SARS-CoV-2 nucleocapsid (N) protein (200-401-A50, Rockland) following permeabilization with

Triton-X100 and blocking with BSA. Images were acquired on the ImageXpress Pico Automated Cell Imaging System (Molecular Devices) using the 10X objective. Stitched images of 9 fields covering the central 36% of the well were analyzed for infected cells using Cell Reporter Xpress software (Molecular Devices). Briefly, cell number was determined by automated counting of DAPI stained nuclei, infected cells are determined as those cells in which positive N staining was detected associated with a nuclei. The percentage of infected cells relative to control wells in which no drug was added (virus only) were calculated and variable slope non-linear fit curves were assigned using GraphPad Prism v8.4.3.

Computational Studies

Modelling

Loops for the unstructured regions of the locked (linoleic acid bound) and open (apo) cryo-EM structures were built using UCSF Chimera (*48*). Disulphide bonds within reasonable reach were reconstructed (42 for the closed structure and 43 for the open) and each chain sequence was used in an EBI-blast check to verify wild type spike sequence post build. Procheck (*58*) was then used to check the quality of the resulting structure prior to simulation. Acpype (*59*) was used to prepare the topologies for the linoleic acid.

Molecular Dynamics Simulations

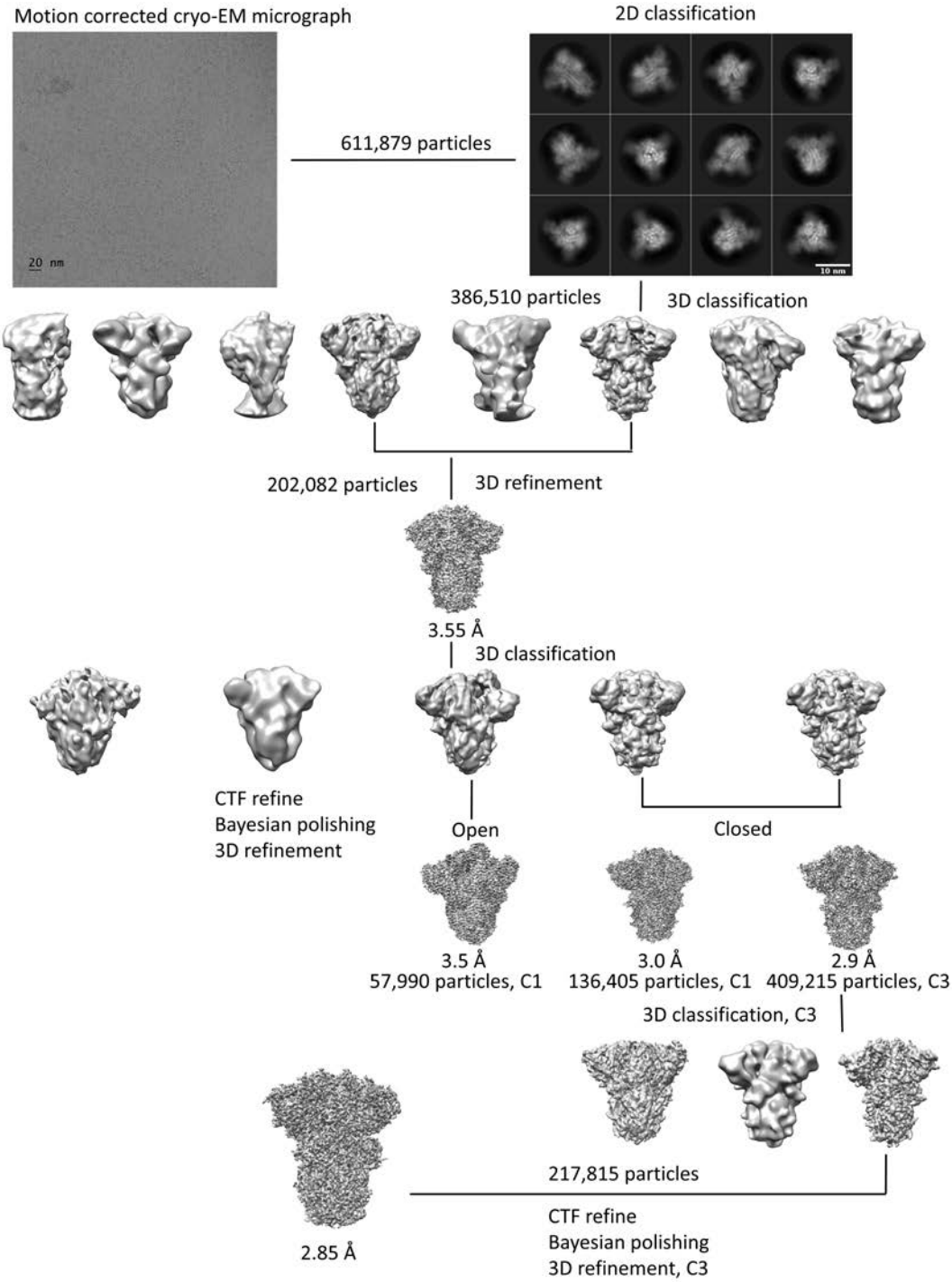
Simulation details: All simulations were performed under the AMBER99SB-ILDN (*60-62*) forcefield as NPT ensembles at 310 K using periodic boundary conditions. Hydrogen atoms, consistent to pH 7, were added to the complex. Short range electrostatic and van der Waals interactions were truncated at 1.4 nm while long range electrostatics were treated with the particle-mesh Ewald method and a long-range dispersion correction applied. A simulation box extending 2 nm from the protein was filled with TIP3P waters and 150 mM Na and Cl ions added to attain a neutral charge overall. Pressure was controlled by the Berendsen barostat and temperature by the V-rescale thermostat. The simulations were integrated with a leap-frog algorithm over a 2 fs time step, constraining bond vibrations with the P-LINCS method. Structures were saved every 0.1 ns for analysis in each of the simulations over 100 ns. Simulation data were accumulated on the Bristol supercomputer BlueCrystal Phase 4, the BrisSynBio BlueGEM, and the UK supercomputer, ARCHER.

Software. The GROMACS-2019.2 (*24*) suite of software was used to set up and perform the molecular dynamics simulations and analyses for BlueGEM and BlueCrystal Phase 4 runs and GROMACS-2019.1 (*24*) for the ARCHER runs. Molecular graphics manipulations and visualisations were performed using VMD-1.9.1 (*63*) and Chimera-1.10.2 (*48*) and movies made using VMD (*63*).

Fig. S1.

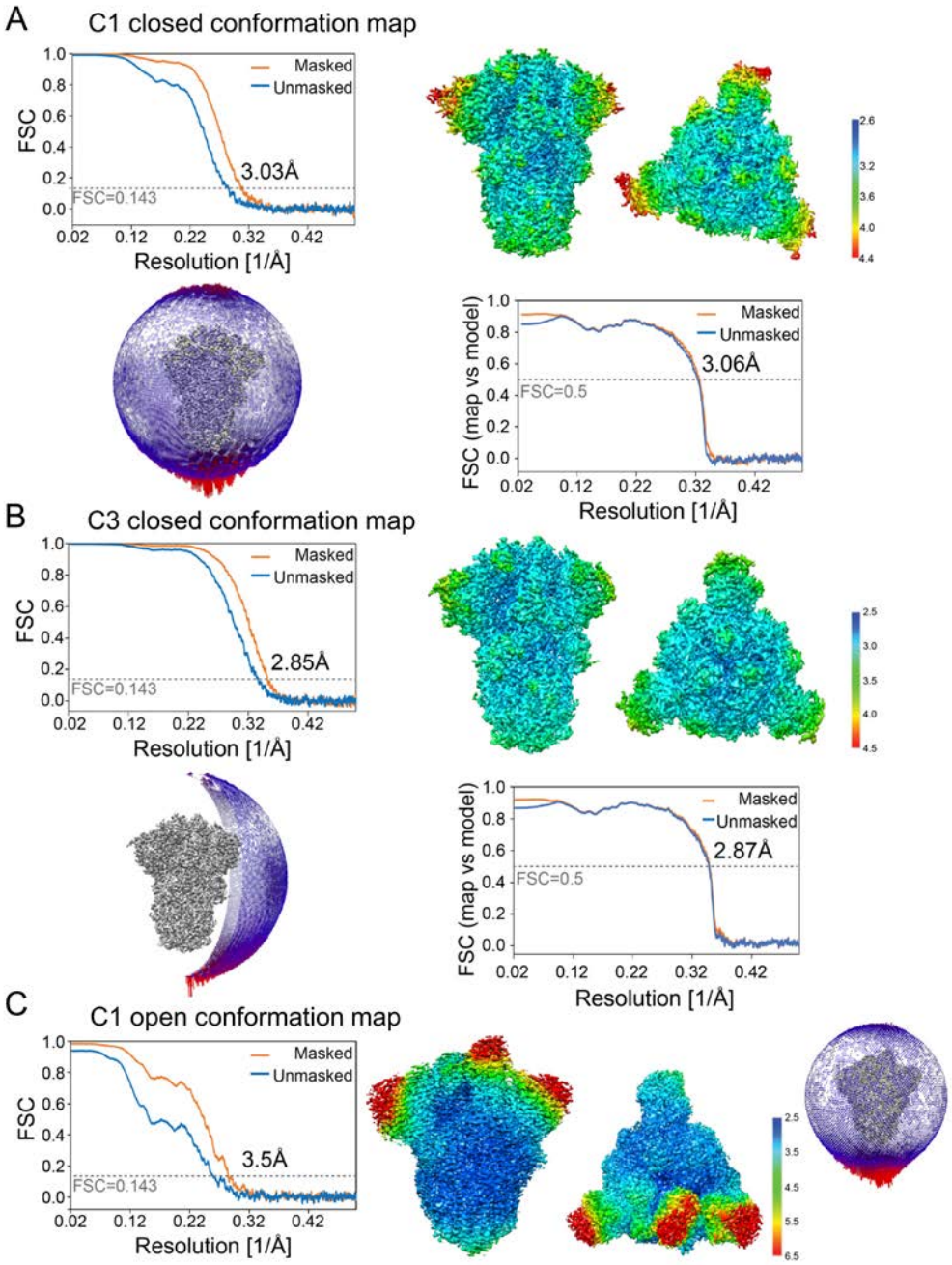
Purification and quality control of SARS-CoV-2 spike protein. (**A**) Size-exclusion chromatogram of affinity-purified SARS-CoV-2 spike protein using a Superdex 200 column. Absorption was detected at 280 nm (blue line) and 260 nm (red line). Peak fractions are indicated. The first peak at 8.6 mL elution volume corresponds to the void volume of the column comprising negligible amounts of spike protein as confirmed by the SDS-PAGE analysis in panel B (fractions 1 and 2). The spike trimer elutes at 10.8 mL. (**B**) SDS PAGE analysis of the SEC fractions from panel A. Lane 1: input fraction, lane 2: molecular weight marker, lane 3-15: fractions 1 to 13 from SEC. (**C**) Negative stain EM of SEC peak fraction 7 (scale bar: 100 nm). In panels A and B peak fraction 7 is highlighted with a red box; this fraction was used for negative stain EM and cryo-EM.

Fig. S2.



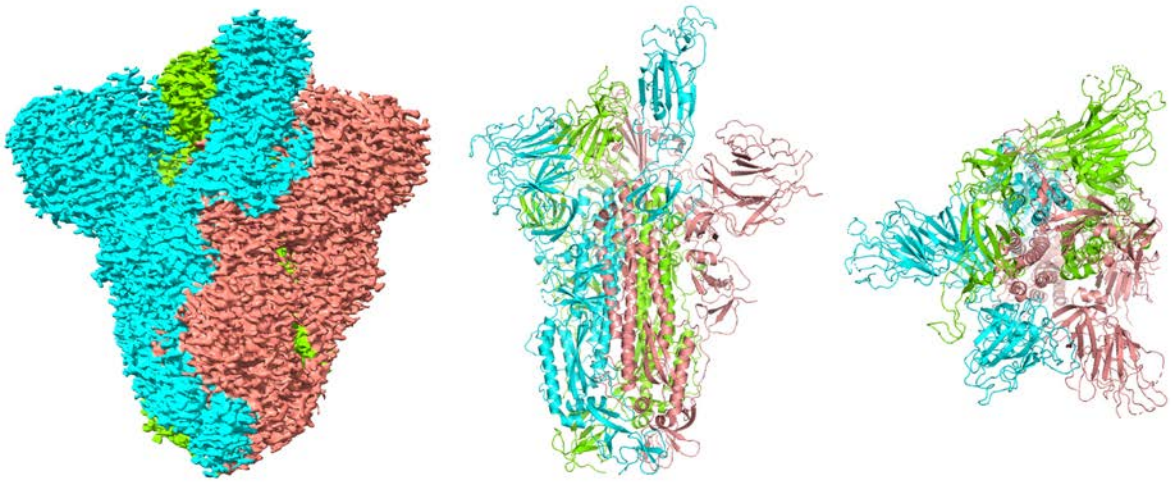
Cryo-EM image processing workflow. A motion-corrected cryo-EM micrograph is shown (scale bar 20 nm), reference-free 2D class averages (scale bar 10 nm), 3D classification and refinement resulting in cryo-EM maps corresponding to the open conformation and the closed conformation (not symmetrized and C3 symmetrized).

Fig. S3.



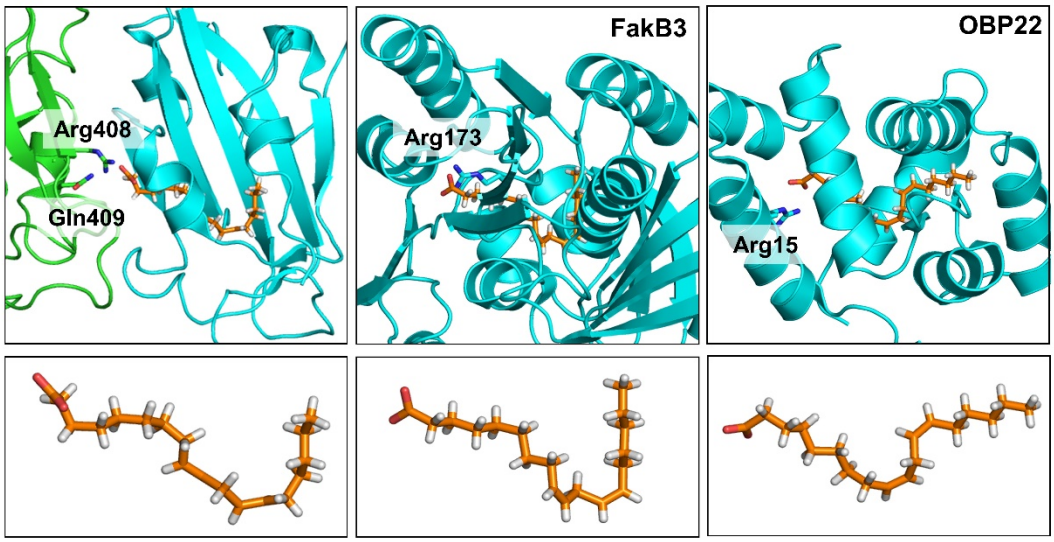
Cryo-EM structure validation. Top left: Fourier Shell Correlation (FSC) curve after gold standard refinement. Top right: Cryo-EM reconstruction colored according to the local resolution from a side and top view. Below left: Orientation distribution of views that contributed to this map. Longer red rods represent orientations that comprise more particles. Below right: Cross-validation FSC curves for the refined model versus the final masked and unmasked maps, shown for (**A**) the closed unsymmetrized C1 map (**B**) the closed C3-symmetrized map and (**C**) the open conformation map. For the open conformation, the views distribution is shown on the right, next to the local resolution map.

Fig. S4.



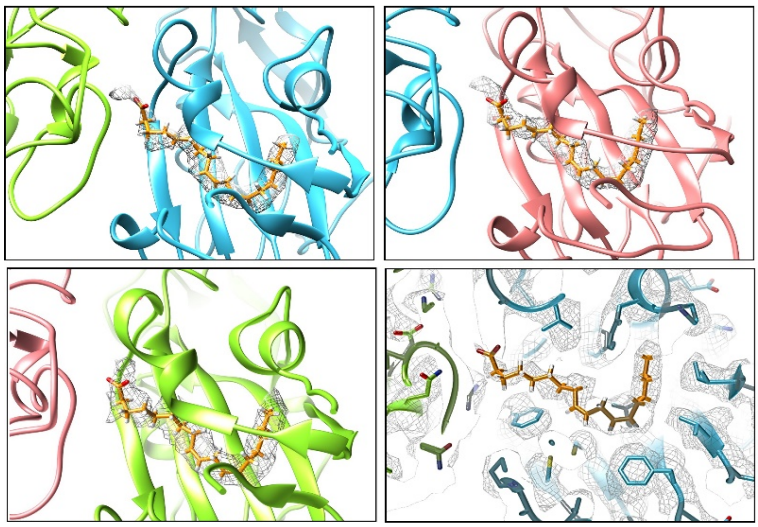
Cryo-EM structure of SARS-CoV-2 S in the open conformation. The 3.5 Å cryo-EM density of the SARS-CoV-2 spike trimer is shown (left). Monomers are colored in cyan, green and pink, respectively. A quasi-atomic model is depicted in a cartoon representation in a front (middle) and top view (right) using the same colors (c.f. Fig.1). RBDs were placed as rigid bodies, and LA was not included in the model.

Fig. S5.



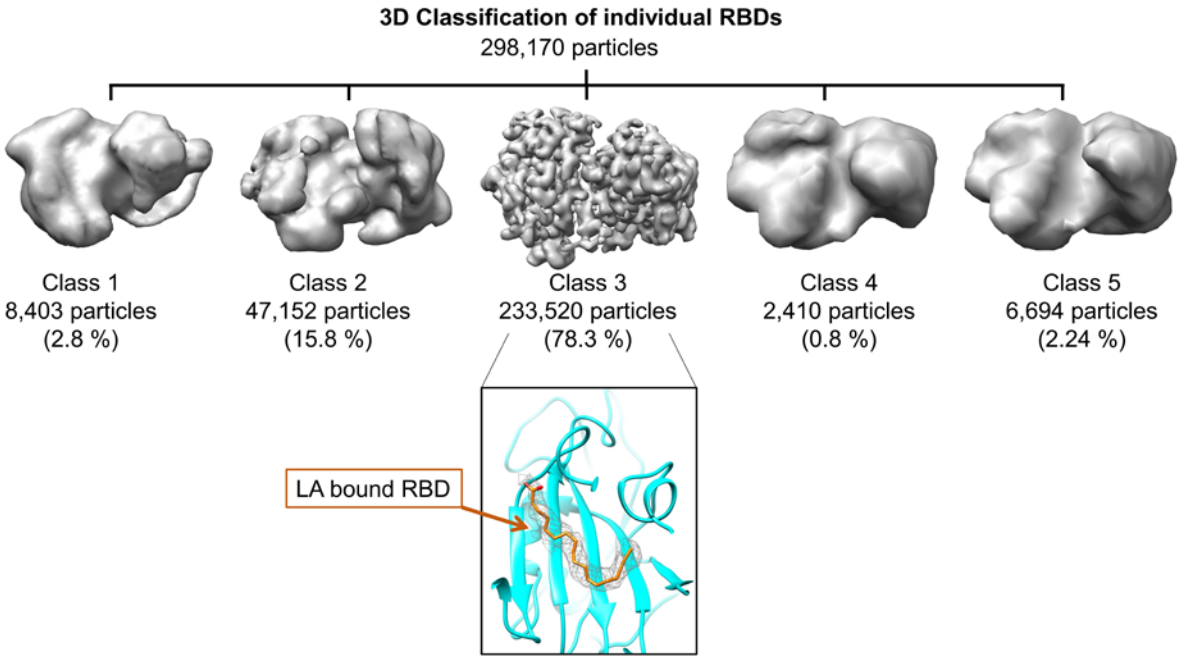
LA binding protein complexes. The structures of LA bound to SARS-CoV-2 S protein (this study, left), *Streptococcus pneumoniae* Fatty Acid Kinase (Fak) B3 (PDBID 6CNG (*20*), middle) and *Aedes aegypti* Odorant Binding Protein 22 (OBP22) (PDBID 6OGH, (*21*), right) are shown for comparison (top panel). Amino acids interacting with the polar headgroup of LA are labelled. Despite markedly different protein architectures, the conformation adopted by the LA in each complex is similar, featuring a pronounced kink in the structures (close-up views below).

Fig. S6.



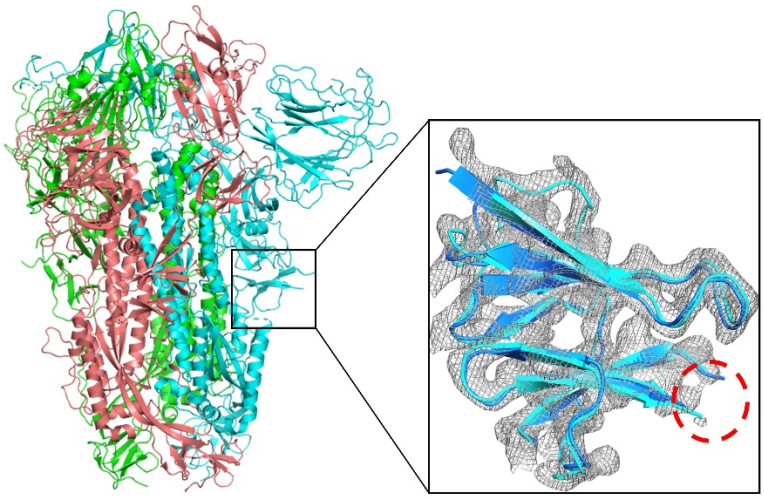
LA ligands bound to the SARS-CoV-2 spike RBDs. The three independent RBDs present in SARS-CoV-2 spike trimer are shown. Processing and refinement in C1 (no symmetry applied) results in tube-shaped density within all three RBDs, indicating that a LA ligand is present in each fatty acid-binding pocket. EM density is depicted as a mesh colored in grey. RBDs are colored cyan, green and pink (c.f. Figure 1). LA ligand is shown as sticks colored in orange. The density of amino acid residues in the binding pocket surrounding LA in the C3 structure is shown for comparison in a clipped view (bottom right).

Fig. S7.



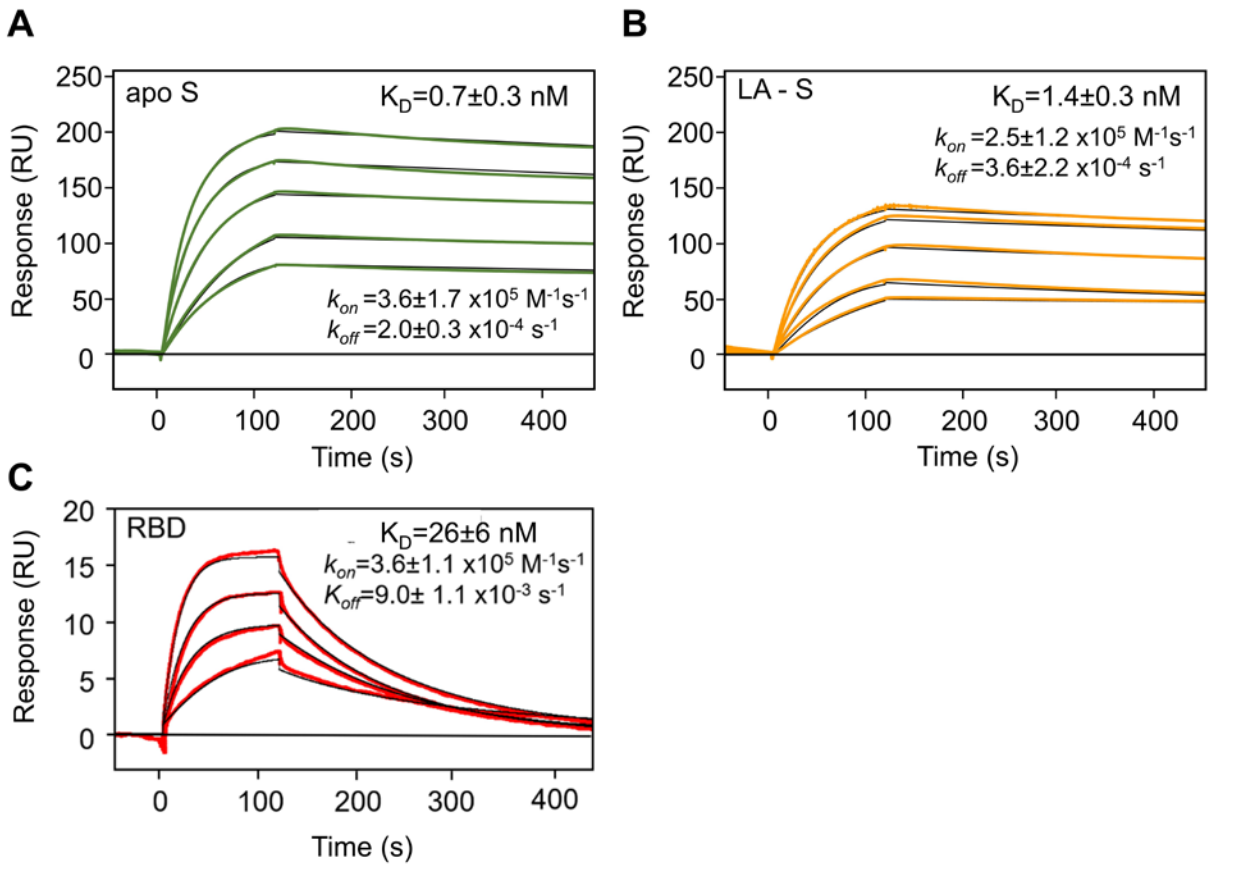
Masked 3D classification focusing on the Receptor Binding Domain of the S glycoprotein into 5 classes. Class 3 comprised 78.3% of all RBDs. This class presents LA bound RBD. The other 3D classes did not reach sufficient resolution to determine the presence or absence of LA in the RBD reconstructions.

Fig. S8.



Polybasic cleavage site deletion in SARS-CoV-2 spike. The construct used in this study has a RRAR to A deletion in the polybasic cleavage site (boxed in black). This site is part of a flexible loop in the S protein which is not observed by cryo-EM (circled in red in close-up view). Comparison of the LA-bound SARS-CoV-2 spike protein (cyan) and the 'apo' form (dark blue, PDBID 6VXX, RRAR->SGAG (*7*)) shown in a close up view superimposed on cryo-EM density (grey mesh) evidences that the structures are virtually identical in this region.

Fig. S9.

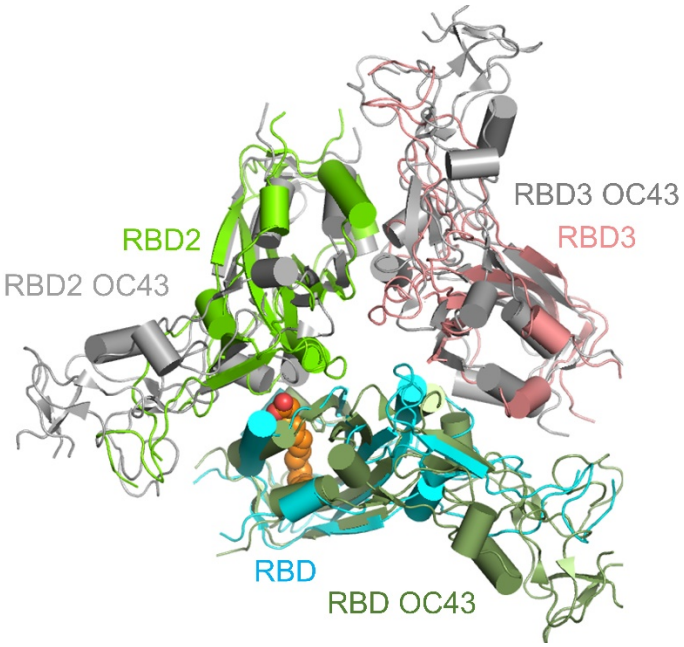


Binding of spike glycoprotein and RBD to immobilized ACE2 ectodomain by surface plasmon resonance (SPR). Apo spike glycoprotein **(A)** and LA-bound spike **(B)** were diluted to concentrations between 40 nM and 160 nM and flowed over 50 RU of biotinylated ACE2 immobilized on a streptavidin-coated sensor chip. Black lines correspond to a global fit of the data using a 1:1 binding model. Each experiment was repeated independently three times. All protein concentrations were used to calculate the K_D, k_{on} and k_{off} values. For the RBD **(C)**, 300 RU ACE2 were immobilized, and RBD concentrations from 80 nM to 160 nM were flowed over the ACE2 surface.

Fig. S10.

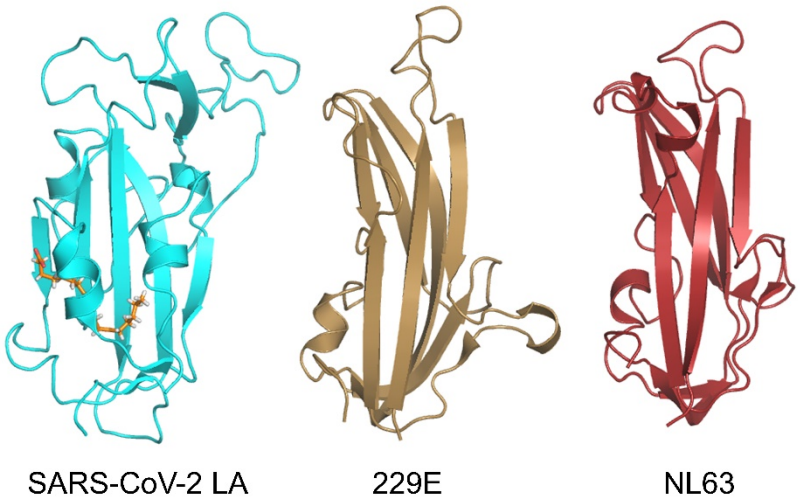
Binding of LA to immobilized RBD by surface plasmon resonance (SPR). LA was diluted to concentrations between 4 μM and 10 μM and flowed over 3,800 RU of biotinylated and lipidex-treated RBD immobilized on a streptavidin-coated sensor chip. Black lines correspond to a global fit of the data using a 1:1 binding model. Each experiment was repeated independently three times. All LA concentrations were used to calculate the $K_{D,}$ k_{on} and k_{off} values.

Fig. S11.



Superimposition of RBDs in OC43 S protein trimer and LA-bound SARS-CoV-2 S. The OC43 RBD colored in dark green and the LA-bound SARS-CoV-2 RBD colored in cyan were used for structure alignment. Two further RBDs in the trimer in OC43 S are colored in grey. Corresponding RBDs in the LA-bound SARS-CoV-2 S are colored in green and pink (c.f. Fig. 1). The three RBDs of OC43 are locked in a tightly packed conformation with close contacts at the RBD-RBD interfaces (PDBID 6NZK (*18*)).

Fig. S12.



Spike protein RBD structures. LA-bound SARS-CoV-2 S RBD (cyan, this study, left) and RBDs of commonly circulating human coronaviruses 229E (PDBID 6U7H (*27*), middle) and NL63 (PDBID 5SZS (*28*), right) are shown. The RBD architectures of 229E and BL63 are substantially different.

Table S1.

Cryo-EM data collection and refinement statistics.

	Closed conformation, C3 symmetrized	Closed conformation, C1	Open conformation, C1
Voltage (kV)	200	200	200
Magnification (nominal)	130,000	130,000	130,000
Pixel size (Å/pix)	1.05 (0.525)	1.05 (0.525)	1.05 (0.525)
Flux (e^- /pix/sec)	6.1	6.1	6.1
Frames per exposure	55	55	55
Exposure (e^- /Å^2)	60.5	60.5	60.5
Defocus range (μm)	-0.8 to -2	-0.8 to -2	-0.8 to -2
Micrographs collected	3289	3289	3289
Particles final	217,815	136,405	57,990
Map sharpening B-factor (Å^2)	-86.8	-89	-87
Masked resolution at 0.143 FSC (Å)	2.85	3.03	3.5

Refinement

	Closed conformation, C3 symmetrized	Closed conformation, C1
Composition		
Amino acids	3096	3133
Glycans	26	27
Ligands	3	3
RMSD bonds (Å)	0.013	0.008
RMSD angles (º)	0.843	0.807
Mean B-factors (Å^2)		
Amino acids	25.25	32.78
Ligands	36.91	38.86
Ramachandran		
Favoured (%)	92.20	90.17
Allowed (%)	7.80	9.66
Outliers (%)	0.00	0.16
Rotamer outliers (%)	3.27	5.84
Clash score	4.16	4.26
C-beta outliers (%)	0.00	0.00
CaBLAM outliers (%)	3.46	4.31
CC (mask)	0.84	0.84
MolProbity score	2.07	2.33
EMRinger score	3.59	3.79
Model resolution (Å) 0.5 FSC threshold	2.9	3.1

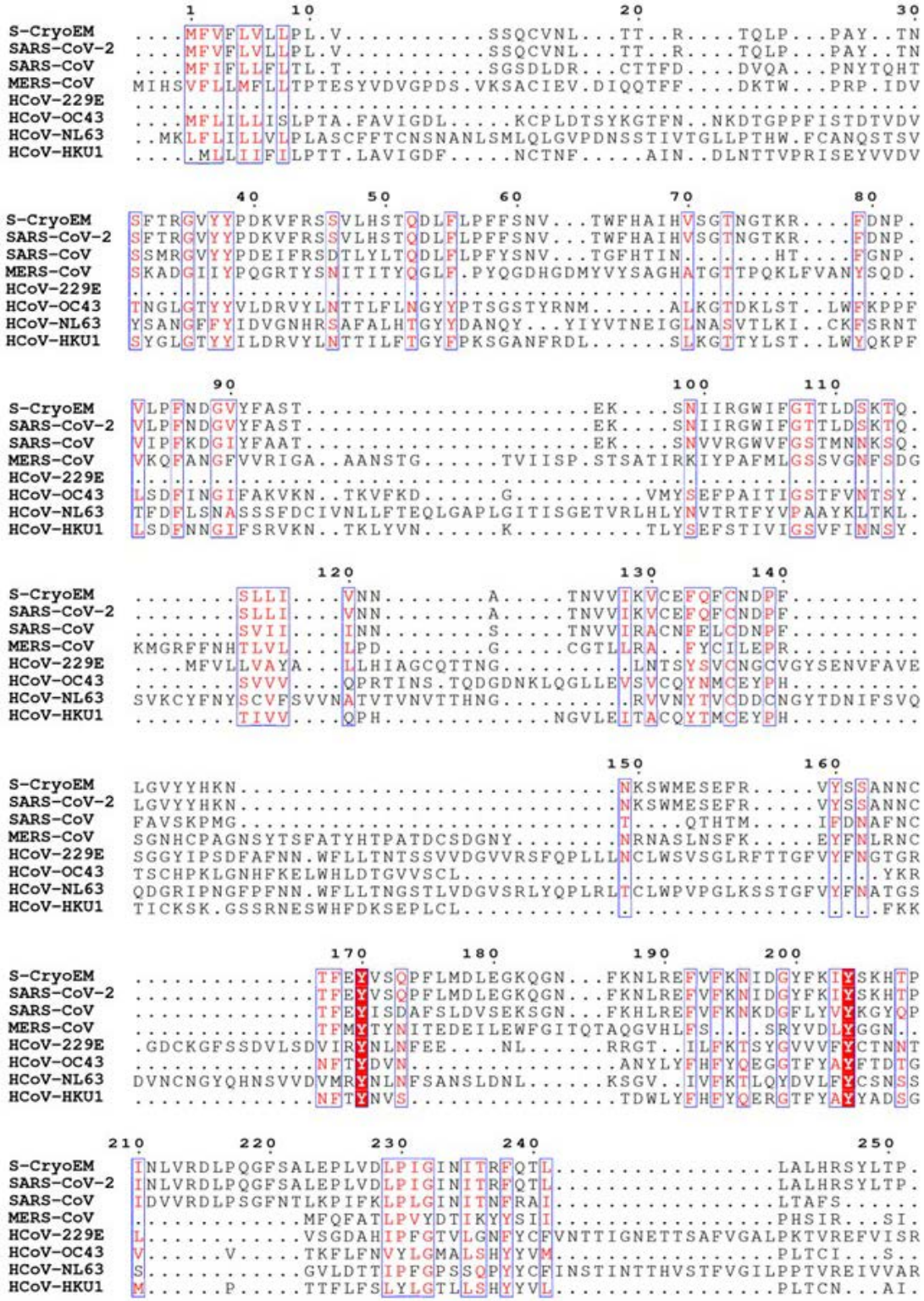
Table S2.
N-linked glycosylation sites in the SARS-CoV-2 spike protein

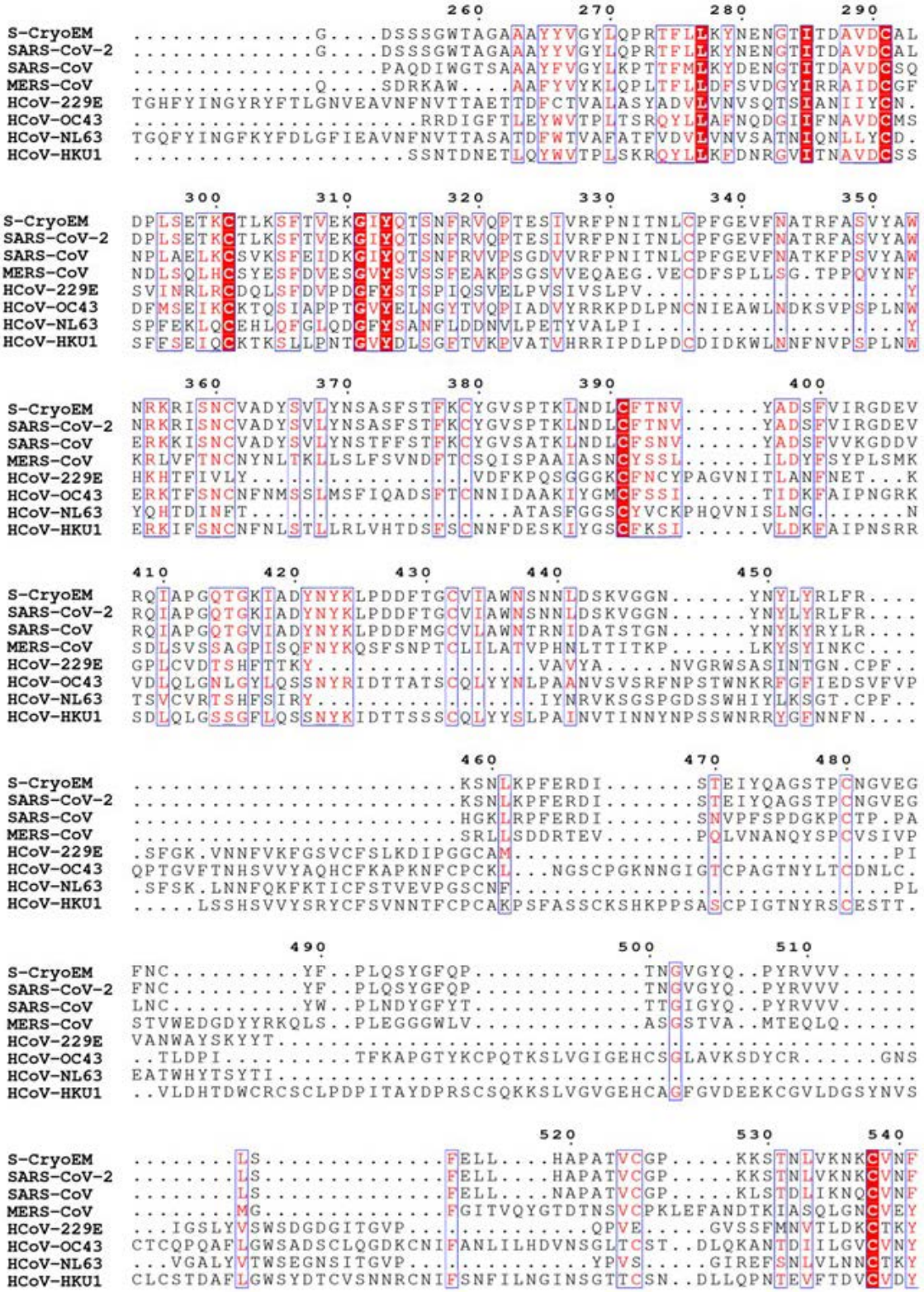
WT SARS-CoV-2 spike *	Recombinant SARS-CoV-2 spike expressed in	
Predicted (UNIPROT)	**Freestyle 293F (7) ****	**Hi5 (this study) ****
N_{17}LT		N_{17}LT
N_{61}VT	N_{61}VT	N_{61}VT
N_{74}GT		
N_{121}NA #		
N_{122}AT	N_{122}AT	N_{122}AT
N_{149}KS		
N_{165}CT	N_{165}CT	N_{165}CT
N_{234}IT	N_{234}IT	N_{234}IT
N_{282}GT	N_{282}GT	N_{282}GT
N_{331}IT	N_{331}IT	N_{331}IT
N_{343}AT	N_{343}AT	N_{343}AT
N_{370}SA #		
N_{603}TS	N_{603}TS	
N_{616}CT	N_{616}CT	N_{616}CT
N_{657}NS	N_{657}NS	
N_{709}NS	N_{709}NS	N_{706}NS
N_{717}FT	N_{717}FT	N_{714}FT
N_{801}FS	N_{801}FS	N_{798}FS
N_{1074}FT	N_{1074}FT	N_{1071}FT
N_{1098}GT	N_{1098}GT	N_{1095}GT
N_{1134}NT	N_{1134}NT	N_{1131}NT
N_{1158}HT		
N_{1173}AS		
N_{1194}ES		

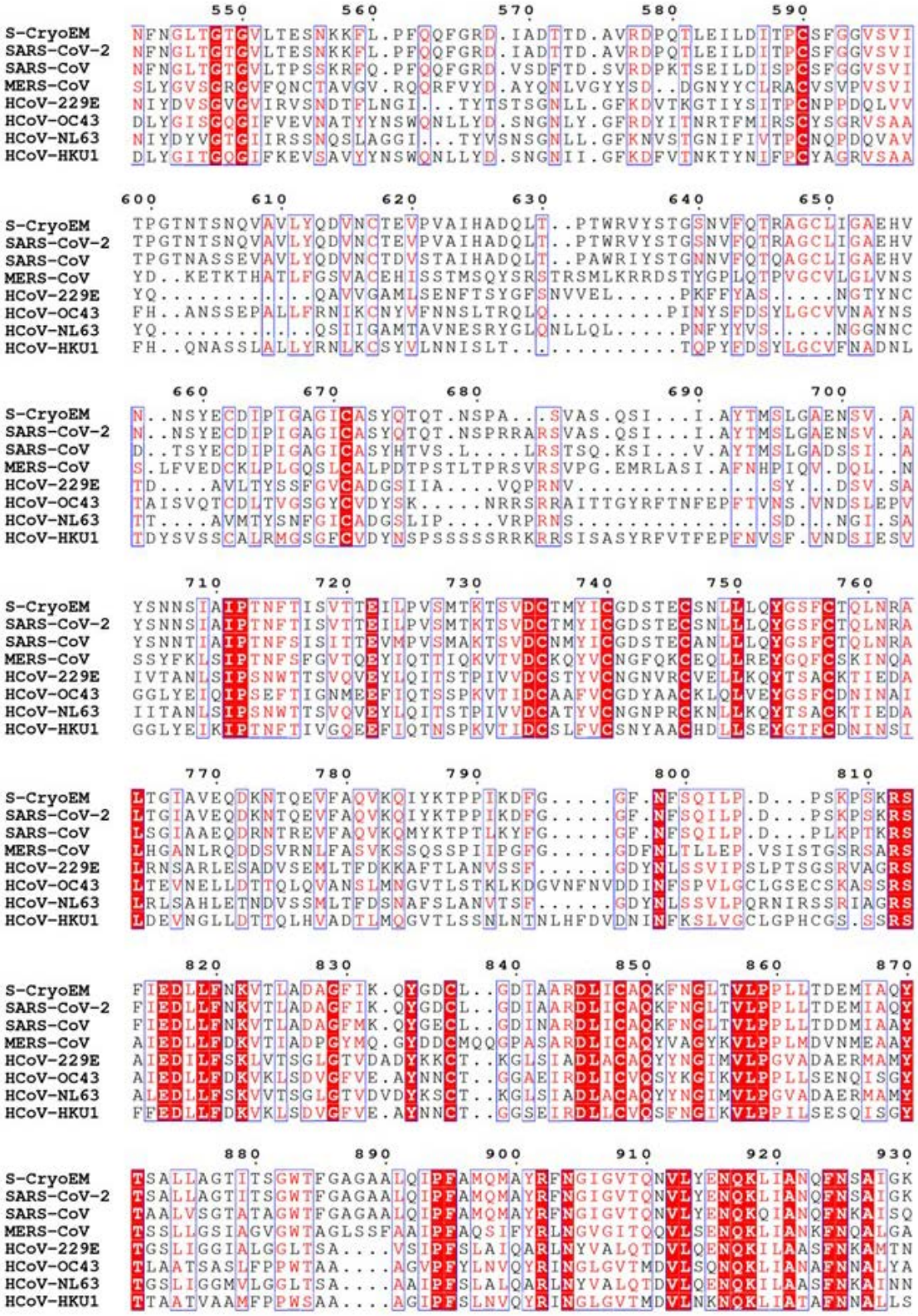
* *YP_009724390.1*
***Sites lacking glycosylation in cryo-EM maps are omitted (boxes colored in grey).*
Predicted based on SARS-CoV spike protein (7)

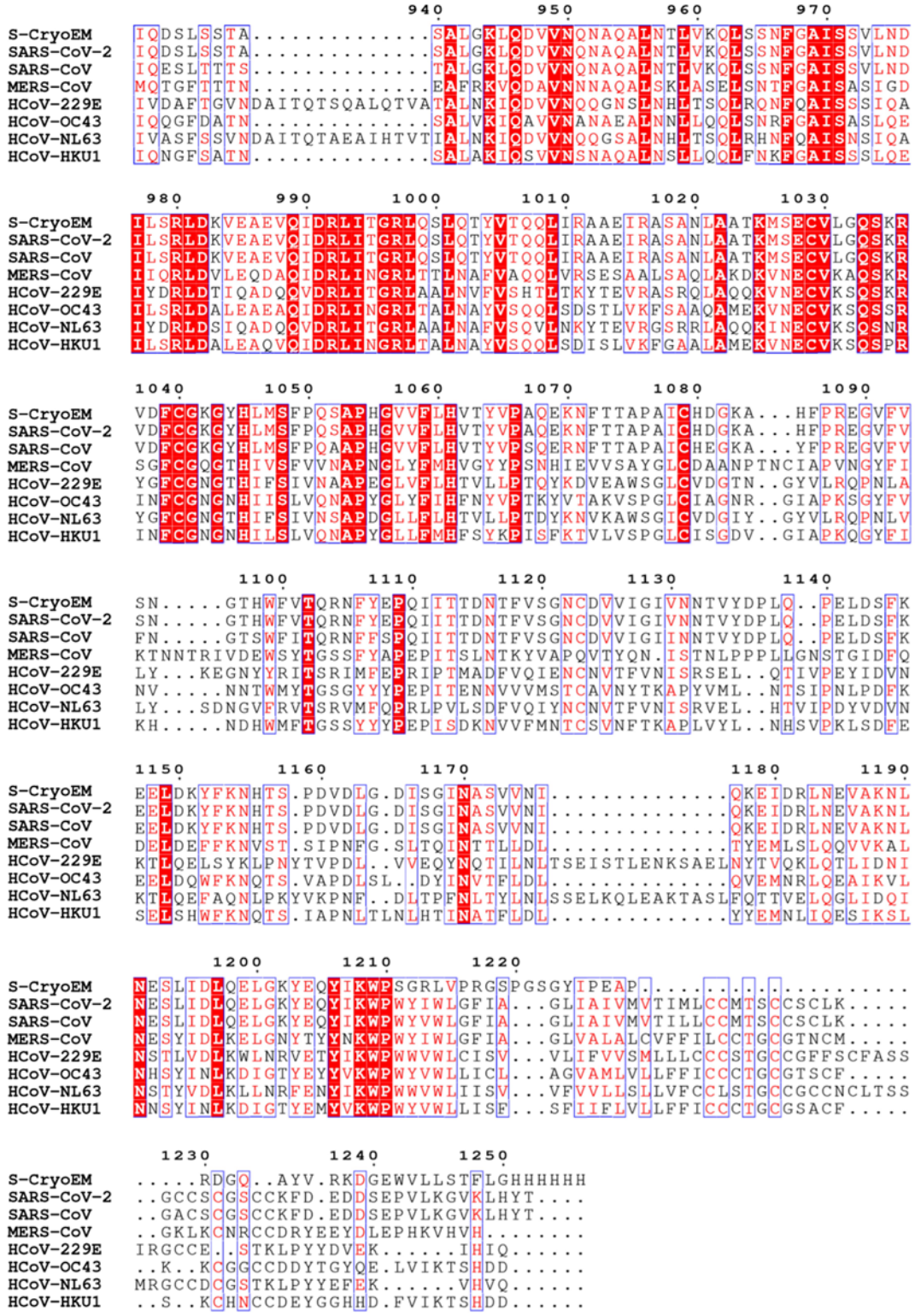
Table S3.

Alignment of coronavirus S protein sequences (S-CryoEM denotes construct in this study)




260 270 280 290
S-CryoEM .....G....DSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCAL
SARS-CoV-2 .....G....DSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCAL
SARS-CoV .....PAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQ
MERS-CoV .....Q....SDRKAW....AAFYVYKLQPLTFLLDFSVDGYIRRAIDCGF
HCoV-229E TGHFYINGYRYFTLGNVEAVNFNVTTAETTDFCTVALASYADVLVNVSQTSIANIIYCN.
HCoV-OC43 .....RRDIGFTLEYWVTPLTSRQYLLAFNQDGIIFNAVDCMS
HCoV-NL63 TGQFYINGFKYFDLGFIEAVNFNVTTASATDFWTVAFATFVDVLVNVSATNIQNLLYCD.
HCoV-HKU1 .....SSNTDNETLQYWVTPLSKRQYLLKFDNRGVITNAVDCSS
300 310 320 330 340 350
S-CryoEM DPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAW
SARS-CoV-2 DPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAW
SARS-CoV NPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAW
MERS-CoV NDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEG.VECDFSPLLSG.TPPQVYNF
HCoV-229E SVINRLRCDQLSFDVPDGFYSTSPIQSVELPVSIVSLPV.....Y
HCoV-OC43 DFMSEIKCKTQSIAPPTGVYELNGYTVQPIADVYRRKPDLPNCNIEAWLNDKSVPSPLNW
HCoV-NL63 SPFEKLQCEHLQFGLQDGFYSANFLDDNVLPETYVALPI.....Y
HCoV-HKU1 SFFSEIQCKTKSLLPNTGVYDLSGFTVKPVATVHRRIPDLPDCDIDKWLNNFNVPSPLNW
360 370 380 390 400
S-CryoEM NRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNV.....YADSFVIRGDEV
SARS-CoV-2 NRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNV.....YADSFVIRGDEV
SARS-CoV ERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNV.....YADSFVVKGDDV
MERS-CoV KRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSL.....ILDYFSYPLSMK
HCoV-229E HKHTFIVLY.....VDFKPQSGGGKCFNCYPAGVNITLANFNET...K
HCoV-OC43 ERKTFSNCNFNMSSLMSFIQADSFTCNNIDAAKIYGMCFSSI.....TIDKFAIPNGRK
HCoV-NL63 YQHTDINFT.....ATASFGGSCYVCKPHQVNISLNG.....N
HCoV-HKU1 ERKIFSNCNFNLSTLLRLVHTDSFSCNNFDESKIYGSCFKSI.....VLDKFAIPNSRR
410 420 430 440 450
S-CryoEM RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGN.....YNYLYRLFR....
SARS-CoV-2 RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGN.....YNYLYRLFR....
SARS-CoV RQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGN.....YNYKYRYLR....
MERS-CoV SDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKP.....LKYSYINKC....
HCoV-229E GPLCVDTSHFTTKY.....VAVYA.....NVGRWSASINTGN.CPF..
HCoV-OC43 VDLQLGNLGYLQSSNYRIDTTATSCQLYYNLPAANVSVSRFNPSTWNKRFGFIEDSVFVP
HCoV-NL63 TSVCVRTSHFSIRY.....IYNRVKSGSPGDSSWHIYLKSGT.CPF..
HCoV-HKU1 SDLQLGSSGFLQSSNYKIDTTSSSCQLYYSLPAINVTINNYNPSSWNRRYGFNNFN....
460 470 480
S-CryoEM .....KSNLKPFERDI.....STEIYQAGSTPCNGVEG
SARS-CoV-2 .....KSNLKPFERDI.....STEIYQAGSTPCNGVEG
SARS-CoV .....HGKLRPFERDI.....SNVPFSPDGKPCTP.PA
MERS-CoV .....SRLLSDDRTEV.....PQLVNANQYSPCVSIVP
HCoV-229E .SFGK.VNNFVKFGSVCFSLKDIPGGCAM.....PI
HCoV-OC43 QPTGVFTNHSVVYAQHCFKAPKNFCPCKL...NGSCPGKNNGIGTCPAGTNYLTCDNLC.
HCoV-NL63 .SFSK.LNNFQKFKTICFSTVEVPGSCNF.....PL
HCoV-HKU1 .....LSSHSVVYSRYCFSVNNTFCPCAKPSFASSCKSHKPPSASCPIGTNYRSCESTT.
490 500 510
S-CryoEM FNC.....YF..PLQSYGFQP.....TNGVGYQ..PYRVVV.....
SARS-CoV-2 FNC.....YF..PLQSYGFQP.....TNGVGYQ..PYRVVV.....
SARS-CoV LNC.....YW..PLNDYGFYT.....TTGIGYQ..PYRVVV.....
MERS-CoV STVWEDGDYYRKQLS..PLEGGGWLV.....ASGSTVA..MTEQLQ.....
HCoV-229E VANWAYSKYYT.....
HCoV-OC43 ..TLDPI.....TFKAPGTYKCPQTKSLVGIGEHCSGLAVKSDYCR.....GNS
HCoV-NL63 EATWHYTSYTI.....
HCoV-HKU1 ..VLDHTDWCRCSCLPDPITAYDPRSCSQKKSLVGVGEHCAGFGVDEEKCGVLDGSYNVS
520 530 540
S-CryoEM .....LS.....FELL....HAPATVCGP.....KKSTNLVKNKCVNF
SARS-CoV-2 .....LS.....FELL....HAPATVCGP.....KKSTNLVKNKCVNF
SARS-CoV .....LS.....FELL....NAPATVCGP.....KLSTDLIKNQCVNF
MERS-CoV .....MG.....FGITVQYGTDTNSVCPKLEFANDTKIASQLGNCVEY
HCoV-229E ...IGSLYVSWSDGDGITGVP.....QPVE.....GVSSFMNVTLDKCTKY
HCoV-OC43 CTCQPQAFLGWSADSCLQGDKCNIFANLILHDVNSGLTCST..DLQKANTDIILGVCVNY
HCoV-NL63 ...VGALYVTWSEGNSITGVP.....YPVS.....GIREFSNLVLNNCTKY
HCoV-HKU1 CLCSTDAFLGWSYDTCVSNNRCNIFSNFILNGINSGTTCSN..DLLQPNTEVFTDVCVDY


550 560 570 580 590
S-CryoEM NFNGLTGTGVLTESNKKFL.PFQQFGRD.IADTTD.AVRDPQTLEILDITPCSFGGVSVI
SARS-CoV-2 NFNGLTGTGVLTESNKKFL.PFQQFGRD.IADTTD.AVRDPQTLEILDITPCSFGGVSVI
SARS-CoV NFNGLTGTGVLTPSSKRFQ.PFQQFGRD.VSDFTD.SVRDPKTSEILDISPCSFGGVSVI
MERS-CoV SLYGVSGRGVFQNCTAVGV.RQQRFVYD.AYQNLVGYYSD..DGNYYCLRACVSVPVSVI
HCoV-229E NIYDVSGVGVIRVSNDTFLNGI...TYTSTSGNLL.GFKDVTKGTIYSITPCNPPDQLVV
HCoV-OC43 DLYGISGQGIFVEVNATYYNSWQNLLYD.SNGNLY.GFRDYITNRTFMIRSCYSGRVSAA
HCoV-NL63 NIYDYVGTGIIRSSNQSLAGGI...TYVSNSGNLL.GFKNVSTGNIFIVTPCNQPDQVAV
HCoV-HKU1 DLYGITGQGIFKEVSAVYYNSWQNLLYD.SNGNII.GFKDFVTNKTYNIFPCYAGRVSAA
600 610 620 630 640 650
S-CryoEM TPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT..PTWRVYSTGSNVFQTRAGCLIGAEHV
SARS-CoV-2 TPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT..PTWRVYSTGSNVFQTRAGCLIGAEHV
SARS-CoV TPGTNASSEVAVLYQDVNCTDVSTAIHADQLT..PAWRIYSTGNNVFQTQAGCLIGAEHV
MERS-CoV YD..KETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQTPVGCVLGLVNS
HCoV-229E YQ.....QAVVGAMLSENFTSYGFSNVVEL.....PKFFYASN.....NGTYNC
HCoV-OC43 FH..ANSSEPALLFRNIKCNYVFNNSLTRQLQ.....PINYSFDSYLGCVVNAYNS
HCoV-NL63 YQ.....QSIIGAMTAVNESRYGLQNLLQL.....PNFYYVSN.....NGGNNC
HCoV-HKU1 FH..QNASSLALLYRNLKCSYVLNNISLT..T.....TQPYFDSYLGCVFNADNL
660 670 680 690 700
S-CryoEM N..NSYECDIPIGAGICASYQTQT.NSPA..SVAS.QSI...I.AYTMSLGAENSV..A
SARS-CoV-2 N..NSYECDIPIGAGICASYQTQT.NSPRRARSVAS.QSI...I.AYTMSLGAENSV..A
SARS-CoV D..TSYECDIPIGAGICASYHTVS.L....LRSTSQ.KSI...V.AYTMSLGADSSI..A
MERS-CoV S.LFVEDCKLPLGQSLCALPDTPSTLTPRSVRSVPG.EMRLASI.AFNHPIQV.DQL..N
HCoV-229E TD....AVLTYSSFGVCADGSIIA....VQPRNV.....SY..DSV.SA
HCoV-OC43 TAISVQTCDLTVGSGYCVDYSK.....NRRSRRAITTGYRFTNFEPFTVNS.VNDSLEPV
HCoV-NL63 TT....AVMTYSNFGICADGSLIP....VRPRNS.....SD..NGI.SA
HCoV-HKU1 TDYSVSSCALRMGSGFCVDYNSPSSSSSRRKRRSISASYRFVTFEPFNVSF.VNDSIESV
710 720 730 740 750 760
S-CryoEM YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRA
SARS-CoV-2 YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRA
SARS-CoV YSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRA
MERS-CoV SSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQA
HCoV-229E IVTANLSIPSNWTTSVQVEYLQITSTPIVVDCSTYVCNGNVRCVELLKQYTSACKTIEDA
HCoV-OC43 GGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACKLQLVEYGSFCDNINAI
HCoV-NL63 IITANLSIPSNWTTSVQVEYLQITSTPIVVDCATYVCNGNPRCKNLLKQYTSACKTIEDA
HCoV-HKU1 GGLYEIKIPTNFTIVGQEEFIQTNSPKVTIDCSLFVCSNYAACHDLLSEYGTFCDNINSI
770 780 790 800 810
S-CryoEM LTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG.....GF.NFSQILP.D...PSKPSKRS
SARS-CoV-2 LTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG.....GF.NFSQILP.D...PSKPSKRS
SARS-CoV LSGIAAEQDRNTREVFAQVKQMYKTPTLKYFG.....GF.NFSQILP.D...PLKPTKRS
MERS-CoV LHGANLRQDDSVRNLFASVKSSQSSPIIPGFG.....GDFNLTLLEP.VSISTGSRSARS
HCoV-229E LRNSARLESADVSEMLTFDKKAFTLANVSSF.....GDYNLSSVIPSLPTSGSRVAGRS
HCoV-OC43 LTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFNVDDINFSPVLGCLGSECSKASSRS
HCoV-NL63 LRLSAHLETNDVSSMLTFDSNAFSLANVTSF.....GDYNLSSVLPQRNIRSSRIAGRS
HCoV-HKU1 LDEVNGLLDTTQLHVADTLMQGVTLSSNLNTNLHFDVDNINFKSLVGCLGPHCGS.SSRS
820 830 840 850 860 870
S-CryoEM FIEDLLFNKVTLADAGFIK.QYGDCL..GDIAARDLICAQKFNGLTVLPPLLTDEMIAQY
SARS-CoV-2 FIEDLLFNKVTLADAGFIK.QYGDCL..GDIAARDLICAQKFNGLTVLPPLLTDEMIAQY
SARS-CoV FIEDLLFNKVTLADAGFMK.QYGECL..GDINARDLICAQKFNGLTVLPPLLTDDMIAAY
MERS-CoV AIEDLLFDKVTIADPGYMQ.GYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAY
HCoV-229E AIEDILFSKLVTSGLGTVDADYKKCT..KGLSIADLACAQYYNGIMVLPGVADAERMAMY
HCoV-OC43 AIEDLLFDKVKLSDVGFVE.AYNNCT..GGAEIRDLICVQSYKGIKVLPPLLSENQISGY
HCoV-NL63 ALEDLLFSKVVTSGLGTVDVDYKSCT..KGLSIADLACAQYYNGIMVLPGVADAERMAMY
HCoV-HKU1 FFEDLLFDKVKLSDVGFVE.AYNNCT..GGSEIRDLLCVQSFNGIKVLPPILSESQISGY
880 890 900 910 920 930
S-CryoEM TSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGK
SARS-CoV-2 TSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGK
SARS-CoV TAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQ
MERS-CoV TSSLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGA
HCoV-229E TGSLIGGIALGGLTSA....VSIPFSLAIQARLNYVALQTDVLQENQKILAASFNKAMTN
HCoV-OC43 TLAATSASLFPPWTAA....AGVPFYLNVQYRINGLGVTMDVLSQNQKLIANAFNNALYA
HCoV-NL63 TGSLIGGMVLGGLTSA....AAIPFSLALQARLNYVALQTDVLQENQKILAASFNKAINN
HCoV-HKU1 TTAATVAAMFPPWSAA....AGIPFSLNVQYRINGLGVTMDVLNKNQKLIATAFNNALLS


940 950 960 970
S-CryoEM IQDSLSSTA.....SALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLND
SARS-CoV-2 IQDSLSSTA.....SALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLND
SARS-CoV IQESLTTTS.....TALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLND
MERS-CoV MQTGFTTTN.....EAFRKVQDAVNNNAQALSKLASELSNTFGAISASIGD
HCoV-229E IVDAFTGVNDAITQTSQALQTVATALNKIQDVVNQQGNSLNHLTSQLRQNFQAISSSIQA
HCoV-OC43 IQQGFDATN.....SALVKIQAVVNANAEALNNLLQQLSNRFGAISASLQE
HCoV-NL63 IVASFSSVNDAITQTAEAIHTVTIALNKIQDVVNQQGSALNHLTSQLRHNFQAISNSIQA
HCoV-HKU1 IQNGFSATN.....SALAKIQSVVNSNAQALNSLLQQLFNKFGAISSSLQE
980 990 1000 1010 1020 1030
S-CryoEM ILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR
SARS-CoV-2 ILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR
SARS-CoV ILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR
MERS-CoV IIQRLDVLEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQLAKDKVNECVKAQSKR
HCoV-229E IYDRLDTIQADQQVDRLITGRLAALNVFVSHTLTKYTEVRASRQLAQQKVNECVKSQSKR
HCoV-OC43 ILSRLDALEAEAQIDRLINGRLTALNAYVSQQLSDSTLVKFSAAQAMEKVNECVKSQSSR
HCoV-NL63 IYDRLDSIQADQQVDRLITGRLAALNAFVSQVLNKYTEVRGSRRLAQQKINECVKSQSNR
HCoV-HKU1 ILSRLDALEAQVQIDRLINGRLTALNAYVSQQLSDISLVKFGAALAMEKVNECVKSQSPR
1040 1050 1060 1070 1080 1090
S-CryoEM VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKA...HFPREGVFV
SARS-CoV-2 VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKA...HFPREGVFV
SARS-CoV VDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKA...YFPREGVFV
MERS-CoV SGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTNCIAPVNGYFI
HCoV-229E YGFCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLCVDGTN..GYVLRQPNLA
HCoV-OC43 INFCGNGNHIISLVQNAPYGLYFIHFNYVPTKYVTAKVSPGLCIAGNR..GIAPKSGYFV
HCoV-NL63 YGFCGNGTHIFSIVNSAPDGLLFLHTVLLPTDYKNVKAWSGICVDGIY..GYVLRQPNLV
HCoV-HKU1 INFCGNGNHILSLVQNAPYGLLFMHFSYKPISFKTVLVSPGLCISGDV..GIAPKQGYFI
1100 1110 1120 1130 1140
S-CryoEM SN.....GTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQ..PELDSFK
SARS-CoV-2 SN.....GTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQ..PELDSFK
SARS-CoV FN.....GTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQ..PELDSFK
MERS-CoV KTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQN.ISTNLPPPLLGNSTGIDFQ
HCoV-229E LY...KEGNYYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSEL..QTIVPEYIDVN
HCoV-OC43 NV.....NNTWMYTGSGYYYPEPITENNVVVMSTCAVNYTKAPYVML..NTSIPNLPDFK
HCoV-NL63 LY...SDNGVFRVTSRVMFQPRLPVLSDFVQIYNCNVTFVNISRVEL..HTVIPDYVDVN
HCoV-HKU1 KH.....NDHWMFTGSSYYYPEPISDKNVVFMNTCSVNFTKAPLVYL..NHSVPKLSDFE
1150 1160 1170 1180 1190
S-CryoEM EELDKYFKNHTS.PDVDLG.DISGINASVVNI.....QKEIDRLNEVAKNL
SARS-CoV-2 EELDKYFKNHTS.PDVDLG.DISGINASVVNI.....QKEIDRLNEVAKNL
SARS-CoV EELDKYFKNHTS.PDVDLG.DISGINASVVNI.....QKEIDRLNEVAKNL
MERS-CoV DELDEFFKNVST.SIPNFG.SLTQINTTLLDL.....TYEMLSLQQVVKAL
HCoV-229E KTLQELSYKLPNYTVPDL..VVEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNI
HCoV-OC43 EELDQWFKNQTS.VAPDLSL..DYINVTFLDL.....QVEMNRLQEAIKVL
HCoV-NL63 KTLQEFAQNLPKYVKPNF..DLTPFNLTYLNLSSELKQLEAKTASLFQTTVELQGLIDQI
HCoV-HKU1 SELSHWFKNQTS.IAPNLTLNLHTINATFLDL.....YYEMNLIQESIKSL
1200 1210 1220
S-CryoEM NESLIDLQELGKYEQYIKWPSGRLVPRGSPGSGYIPEAP.....
SARS-CoV-2 NESLIDLQELGKYEQYIKWPWYIWLGFIA...GLIAIVMVTIMLCCMTSCCSCLK.....
SARS-CoV NESLIDLQELGKYEQYIKWPWYVWLGFIA...GLIAIVMVTILLCCMTSCCSCLK.....
MERS-CoV NESYIDLKELGNYTYYNKWPWYIWLGFIA...GLVALALCVFFILCCTGCGTNCM.....
HCoV-229E NSTLVDLKWLNRVETYIKWPWWVWLCISV...VLIFVVSMLLLCCSTGCCGFFSCFASS
HCoV-OC43 NHSYINLKDIGTYEYYVKWPWYVWLLICL...AGVAMLVLLFFICCCTGCGTSCF.....
HCoV-NL63 NSTYVDLKLLNRFENYIKWPWWVWLIISV...VFVVLLSLLVFCCLSTGCCGCCNCLTSS
HCoV-HKU1 NNSYINLKDIGTYEMYVKWPWYVWLLISF...SFIIFLVLLFFICCCTGCGSACF.....
1230 1240 1250
S-CryoEM .....RDGQ..AYV.RKDGEWVLLSTFLGHHHHHH
SARS-CoV-2 ..GCCSCGSCCKFD.EDDSEPVLKGVKLHYT....
SARS-CoV ..GACSCGSCCKFD.EDDSEPVLKGVKLHYT....
MERS-CoV ..GKLKCNRCCDRYEEYDLEPHKVHVH.....
HCoV-229E IRGCCE..STKLPYYDVEK.....IHIQ.....
HCoV-OC43 ..K..KCGGCCDDYTGYQE.LVIKTSHDD.....
HCoV-NL63 MRGCCDCGSTKLPYYEFEK.....VHVQ.....
HCoV-HKU1 ..S..KCHNCCDEYGGHHD.FVIKTSHDD.....

Movies Captions

Movie S1: Simulation of LA-bound SARS-CoV-2 S trimer colored in cyan, green and pink in the locked conformation, viewed from the side. LA is shown as spheres.

Movie S2: Simulation of SARS-CoV-2 S trimer in the locked form, close up view of LA-bound pocket. LA is shown as spheres. Amino acids interacting with the ligand are shown as sticks.

Movie S3: Simulation of SARS-CoV-2 S trimer colored in cyan, green and pink in the open form, viewed from the side. The RBD of one S monomer (colored in pink) adopts the 'up' position. LA is shown as spheres.

Movie S4: Animated structure of LA-bound SARS-CoV-2 S trimer colored in cyan, green and pink. Morphing of apo S (PDBID 6VXX (*7*)) to LA-bound S is illustrated.

References and Notes:

1. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
2. A. Zumla, J. F. Chan, E. I. Azhar, D. S. Hui, K. Y. Yuen, Coronaviruses - drug discovery and therapeutic options. *Nat Rev Drug Discov* **15**, 327-347 (2016).
3. Z. Varga *et al.*, Endothelial cell infection and endotheliitis in COVID-19. *Lancet* **395**, 1417-1418 (2020).
4. V. G. Puelles *et al.*, Multiorgan and Renal Tropism of SARS-CoV-2. *N Engl J Med* **383**, 590-592 (2020).
5. M. Hoffmann *et al.*, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278 (2020).
6. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* **5**, 562-569 (2020).
7. A. C. Walls *et al.*, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292 e286 (2020).
8. M. Hoffmann, H. Kleine-Weber, S. Pohlmann, A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol Cell* **78**, 779-784 e775 (2020).
9. S. Matsuyama *et al.*, Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc Natl Acad Sci U S A* **117**, 7001-7003 (2020).
10. B. Yan *et al.*, Characterization of the Lipidomic Profile of Human Coronavirus-Infected Cells: Implications for Lipid Metabolism Remodeling upon Coronavirus Replication. *Viruses* **11**, 73 (2019).
11. C. Qin *et al.*, Dysregulation of immune response in patients with COVID-19 in Wuhan, China. *Clin Infect Dis* **71**, 762-768 (2020).
12. M. Z. Tay, C. M. Poh, L. Renia, P. A. MacAry, L. F. P. Ng, The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol* **20,** 363-374 (2020).
13. C. G. K. Ziegler *et al.*, SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. *Cell* **181**, 1016-1035 e1019 (2020).
14. F. Amanat *et al.*, A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat Med* **26**, 1033-1036 (2020).
15. D. J. Fitzgerald *et al.*, Protein complex expression by using multigene baculoviral vectors. *Nat Methods* **3**, 1021-1032 (2006).
16. To produce spike glycoprotein, we used here baculovirus-infected insect cells (Hi5) cultured in ESF921 media which contains cod liver oil as a nutrient supplement (Corey Jacklin, Expression Systems, personal communication). Cod liver oil comprises hundreds of FFAs including LA (*38*) which could be a possible source of the fatty acid. In parallel to our work, cryo-EM structures were determined of spike expressed in a mammalian system (HEK293) in serum and protein-free media (*39*, *40*). The tube-shaped density we identified in our study is also present in those structures (but was not assigned or interpreted). We conclude that LA binding in the SARS-CoV-2 S pocket is thus not dependent on the expression system or media used.

17. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).
18. M. A. Tortorici *et al.*, Structural basis for human coronavirus attachment to sialic acid receptors. *Nat Struct Mol Biol* **26**, 481-489 (2019).
19. Y. Yuan *et al.*, Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun* **8**, 15092 (2017).
20. J. M. Gullett, M. G. Cuypers, M. W. Frank, S. W. White, C. O. Rock, A fatty acid-binding protein of *Streptococcus pneumoniae* facilitates the acquisition of host polyunsaturated fatty acids. *J Biol Chem* **294**, 16416-16428 (2019).
21. J. Wang, E. J. Murphy, J. C. Nix, D. N. M. Jones, *Aedes aegypti* Odorant Binding Protein 22 selectively binds fatty acids through a conformational change in its C-terminal tail. *Sci Rep* **10**, 3300 (2020).
22. J. Shang *et al.*, Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224 (2020).
23. J. F. Glatz, J. H. Veerkamp, Removal of fatty acids from serum albumin by Lipidex 1000 chromatography. *J Biochem Biophys Methods* **8**, 57-61 (1983).
24. H. J. C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43-56 (1995).
25. J. H. Beigel *et al.*, Remdesivir for the Treatment of Covid-19 - Preliminary Report. *N Engl J Med*, NEJMoa2007764 (2020).
26. X. Ou *et al.*, Crystal structure of the receptor binding domain of the spike glycoprotein of human betacoronavirus HKU1. *Nat Commun* **8**, 15216 (2017).
27. Z. Li *et al.*, The human coronavirus HCoV-229E S-protein structure and receptor binding. *Elife* **8**, e51230 (2019).
28. A. C. Walls *et al.*, Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol* **23**, 899-905 (2016).
29. J. Badger *et al.*, Structural analysis of a series of antiviral agents complexed with human rhinovirus 14. *Proc Natl Acad Sci U S A* **85**, 3304-3308 (1988).
30. M. A. Oliveira *et al.*, The structure of human rhinovirus 16. *Structure* **1**, 51-68 (1993).
31. V. Casanova, F. H. Sousa, C. Stevens, P. G. Barlow, Antiviral therapeutic approaches for human rhinovirus infections. *Future Virol* **13**, 505-518 (2018).
32. K. K. W. To, C. C. Y. Yip, K. Y. Yuen, Rhinovirus - From bench to bedside. *J Formos Med Assoc* **116**, 496-504 (2017).
33. B. Shen *et al.*, Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **182**, 59-72 (2020).
34. C. M. Goodwin, S. Xu, J. Munger, Stealing the Keys to the Kitchen: Viral Manipulation of the Host Cell Metabolic Network. *Trends Microbiol* **23**, 789-798 (2015).
35. J. W. Schoggins, G. Randall, Lipids in innate antiviral defense. *Cell Host Microbe* **14**, 379-385 (2013).

36. M. M. Zaman *et al.*, Linoleic acid supplementation results in increased arachidonic acid and eicosanoid production in CF airway cells and in cftr-/- transgenic mice. *Am J Physiol Lung Cell Mol Physiol* **299**, L599-606 (2010).
37. I. Kimura, A. Ichimura, R. Ohue-Kitano, M. Igarashi, Free Fatty Acid Receptors in Health and Disease. *Physiol Rev* **100**, 171-210 (2020).
38. S. Hauff, W. Vetter, Quantitation of cis- and trans-monounsaturated fatty acids in dairy products and cod liver oil by mass spectrometry in the selected ion monitoring mode. *J Agric Food Chem* **57**, 3423-3430 (2009).
39. Y. Cai *et al.*, Distinct conformational states of SARS-CoV-2 spike protein. *Science* (2020).
40. A. G. Wrobel *et al.*, SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat Struct Mol Biol* **27**, 763-767 (2020).
41. I. Berger, D. J. Fitzgerald, T. J. Richmond, Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol* **22**, 1583-1587 (2004).
42. M. M. Vork, J. F. Glatz, D. A. Surtel, G. J. van der Vusse, Assay of the binding of fatty acids by proteins: evaluation of the Lipidex 1000 procedure. *Mol Cell Biochem* **98**, 111-117 (1990).
43. Q. Wang *et al.*, Protocols and pitfalls in obtaining fatty acid-binding proteins for biophysical studies of ligand-protein and protein-protein interactions. *Biochem Biophys Rep* **10**, 318-324 (2017).
44. M. Fairhead, M. Howarth, Site-specific biotinylation of purified proteins using BirA. *Methods Mol Biol* **1266**, 171-184 (2015).
45. S. Q. Zheng *et al.*, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
46. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221 (2015).
47. S. H. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530 (2012).
48. T. D. Goddard, C. C. Huang, T. E. Ferrin, Visualizing density maps with UCSF Chimera. *J Struct Biol* **157**, 281-287 (2007).
49. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
50. T. C. Terwilliger, O. V. Sobolev, P. V. Afonine, P. D. Adams, Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr D Struct Biol* **74**, 545-559 (2018).
51. N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* **65**, 1074-1080 (2009).
52. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).
53. V. B. Chen *et al.*, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).

54. B. A. Barad *et al.*, EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat Methods* **12**, 943-946 (2015).
55. J. L. Daly *et al.*, Neuropilin-1 is a host factor for SARS-CoV-2 infection. *bioRxiv, doi: https://doi.org/10.1101/2020.06.05.134114*, (2020).
56. A. D. Davidson *et al.*, 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).
57. V. M. Corman *et al.*, Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* **25**, 2000045 (2020).
58. R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. P. Thornton, PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.* **26**, 283-291 (1993).
59. A. W. Sousa da Silva, W. F. Vranken, ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res Notes* **5**, 367 (2012).
60. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-1174 (2004).
61. D. A. Case *et al.*, The Amber biomolecular simulation programs. *J Comput Chem* **26**, 1668-1688 (2005).
62. K. Lindorff-Larsen *et al.*, Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950-1958 (2010).
63. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).