The SARS-CoV-2 Spike harbours a lipid binding pocket which modulates stability of the prefusion trimer

Loic Carrique[1*], Helen ME Duyvesteyn[1,2*], Tomas Malinauskas[1*], Yuguang Zhao[1*], Jingshan Ren[1], Daming Zhou[1], Thomas S Walter[1], Julika Radecke[3], Jiandong Huo[1,4,5], Reinis R Ruza[1], Pranav NM Shah[1], Elizabeth E Fry[1†], David I Stuart[1,3†]

[1] Division of Structural Biology, University of Oxford, The Wellcome Centre for Human Genetics, Headington, Oxford, OX3 7BN, UK.

[2] Centre for Translational Immunology, Chinese Academy of Medical Sciences Oxford Institute, University of Oxford, Oxford OX3 7FZ, UK.

[3] Diamond Light Source Ltd, Harwell Science & Innovation Campus, Didcot, OX11 0DE, UK.

[4] The Rosalind Franklin Institute, Harwell Campus, OX11 0FA, UK.

[5] Protein Production UK, Research Complex at Harwell, Harwell Science & Innovation Campus, Didcot, OX11 0FA, UK.

* These authors contributed equally to this work.

† Correspondence: David I. Stuart, phone: +44 1865 287567, e-mail: dave@strubi.ox.ac.uk, or Elizabeth E. Fry, e-mail: liz@strubi.ox.ac.uk

Highlights

- **SARS-CoV-2 Spike can adopt a 'locked' conformation with all receptor binding domains (RBDs) down, likely to represent the prefusion resting state**

- **This 'locked' conformation is compact and stable, braced by lipid bound within a potentially druggable pocket**

- **Key neutralization epitopes are shielded in the 'locked' form**

- **Loss of lipid may trigger a cascade of events that lead to cell entry analogous to the role of lipids in enterovirus cell entry**

Summary

Large trimeric Spikes decorate SARS-CoV-2 and bind host cells via receptor binding domains (RBDs). We report a conformation in which the trimer is 'locked' into a compact well-ordered form. This differs from previous structures where the RBD can flip up to recognise the receptor. In the 'locked' form regions associated with fusion transitions are stabilised and the RBD harbours curved lipids. The acyl chains bind a hydrophobic pocket in one RBD whilst the polar headgroups attach to an adjacent RBD of the trimer. By functional analogy with enteroviral pocket factors loss of the lipid would destabilise the 'locked' form facilitating receptor attachment, conversion to the postfusion state and virus infection. The nature of lipids available at the site of infection might affect the antigenicity/pathogenicity of released virus. These results reveal a potentially druggable pocket and suggest that the natural prefusion state occludes neutralising RBD epitopes, achieving conformational shielding from antibodies.

2

Introduction

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), the causative pathogen of COVID-19, emerged in late 2019 and has caused unprecedented global damage to human health and the economy (Nicola et al., 2020). Coronaviruses (CoVs) are pleomorphic enveloped RNA viruses possessing an unusually large positive sense, single stranded genome. They are classified into four genera, the α-, β-, γ- and δ-coronaviruses (Fehr and Perlman, 2015). SARS-CoV-2 belongs to group β, alongside two other viruses responsible for recent disease outbreaks, MERS-CoV (2012) and SARS-CoV-1 (2001-2003) (Gorbalenya et al., 2020). Although many coronaviruses are not pathogenic to humans, zoonotic coronaviruses have repeatedly jumped species barriers. Furthermore coronaviruses are prone to recombination, making the search for cross-protective antivirals and vaccines urgent (Su et al., 2016).

The SARS-CoV-2 viral envelope is decorated with glycosylated Spike, a large (>1200 residues per protomer) trimeric type I fusion protein. The majority of neutralising antibodies appear to target Spike and it is the primary focus of attention in the development of non-classical (e.g. RNA, DNA and VLP based) vaccines (Amanat and Krammer, 2020). The protein has at least two major functions. The first is to attach to receptor(s), with ACE2 identified as having high affinity for one domain of Spike, termed the receptor binding domain (RBD), and many neutralising antibodies function by abrogating this interaction (Wu et al., 2020). The second function is to act as the machine which fuses the viral and host cell membranes allowing the viral genome to enter the cell (Rey and Lok, 2018). This multistage process is not fully understood but it is thought that receptor attachment destabilizes the prefusion state and cleavage of Spike by host protease into two major fragments, S1 and S2, releases the metastable protein to undergo massive conformational changes, discarding the S1 portion which includes

3

the RBD. This allows S2 to convert into an extended helical bundle projecting a fusion peptide outwards to engage the host cell membrane before then hauling the membranes together to merge cytoplasmic and viral compartments. The use of enzymatic cleavage to release the potential energy of metastable fusion machines is common to many enveloped viruses, and coronaviruses belong to a very large set, bearing so-called 'type I fusion proteins', which share a common helical bundle mechanism, but diverge widely in cell attachment methods (Rey and Lok, 2018). Structural studies to date have largely used a highly engineered version of Spike (benefitting from experience with stabilising SARS-CoV-1 and MERS Spikes) to obtain a stable prefusion form, which although non-functional, maintains the authentic antigenicity (Pallesen et al., 2017). The most pertinent change was the introduction of two proline residues (2P) in the hairpin region linking the HR1 and Central helices (Figure S1) that become a single helix in the postfusion state. The prolines prevent formation of the extended helix and were characterized as stabilizing the prefusion Spike (Pallesen et al., 2017). This engineered 'prefusion' Spike has a rather loose structure, with significant portions of the trimer subunit disordered, and is found with the RBDs (residues 333-527) arrayed in different configurations of two orientations, either 'up' or 'down' (PDBIDs:6VSB, 6VYB, 6VXX, 6X2A,B,C,9) (Henderson et al., 2020; Walls et al., 2020; Wrapp et al., 2020). The RBD sits top-centre on the Spike, surrounded by the larger NTD, also part of S1. In the 'up' conformation the ACE2 binding site is accessible, however this conformation is thought to destabilize the Spike, facilitating fusion with the host cell membrane (Yan et al., 2018). It remains a puzzle how the Spike interrogates its surroundings for receptors whilst maintaining integrity of the prefusion state, however the accepted model has been that RBDs stochastically flip up to transiently reveal the ACE2 binding site (Yuan et al., 2017). This view of the Spike structure has been challenged by the observation that purified wild-type Spike ectodomain exists as a mixture of just two states: a 'locked' (all RBDs down) prefusion state not previously seen and a postfusion

conformation (Cai et al., 2020).

Efforts to meet the need for COVID-19 small molecule therapeutics have been primarily directed against essential viral proteases and polymerases (Khan et al., 2020; Zhang et al., 2020), with much interest, and some success, centred on opportunities for drug repurposing (Yin et al., 2020). An alternative approach is antibody therapy, either through the use of convalescent serum or monoclonal antibodies (Duan et al., 2020; Huo et al., 2020). By analogy with the search for AIDS therapeutics we would expect that, if there is sufficient will and resource, effective antivirals will be found that target a range of proteins, including non-enzyme targets, as exemplified by, for HIV, several clinically approved entry inhibitors (Esté and Telenti, 2007). Also repurposed drugs have been found that neutralize Ebolavirus by destabilising the functional equivalent of Spike (Zhao et al., 2016) and for coronaviruses EK1, a mimetic of part of the S2 fragment of Spike, has been shown to act as an entry inhibitor (Xia et al., 2019). Understanding Spike structure and conformational changes is therefore important to define the anatomy of the target not only for neutralizing antibodies (Jiang et al., 2020), but also new or repurposed chemical entities (Fleishman et al., 2011).

Here, we report a high-resolution cryo-EM structure for the 2P version of SARS-CoV-Spike ectodomain, revealing a tightly packed arrangement of the Spike trimer with all RBDs down. This 'locked' form orders substantial portions of structure including regions adjacent to the fusion peptide, suggesting it hinders conversion to the postfusion state. Remarkably, the 'locked' conformation, which is probably dis-favoured by the proline mutations, is stabilized by hydrophobic ligands, probably one or a mixture of several unsaturated fatty acids, buried in a hydrophobic pocket in the RBD. Stabilisation is achieved by RBD-RBD cross-linking by the head group of the lipid and also by contacts mediated by conserved glycans. Consistent with

5

this observation, we show that lipid-like molecules modulate the stability of the Spike. These data reveal that the RBD of SARS-CoV-2 Spike has a potentially druggable pocket and suggest alternative models for Spike dynamics and receptor attachment from those currently accepted.

Results

Locked prefusion SARS-CoV-2 Spike dis-favours conversion to postfusion state.

A detailed analysis was performed of cryo-EM data collected during a study of the binding of the neutralizing antibody CR3022 (Huo et al., 2020) to engineered Spike ectodomain, which has residues K986 and V987 replaced with prolines (2P), an abrogated furin cleavage motif and a C-terminal foldon domain to maintain trimeric association (Wrapp et al., 2020) (Methods, Figures S1-S3, Table S1). This revealed an further Spike conformation (referred to here as 'locked') in addition to the 'up' or 'down' configurations previously observed (predominantly two, one with all three RBDs down ('all down') and the other with one RBD rotated up and two down ('one up') (Walls et al., 2020; Wrapp et al., 2020) (Methods, Figure S2). The 'locked' configuration, comprising <10% of the total picked particles, constitutes the most tightly packed form observed for this SARS-CoV-2 Spike protein (key features are shown in Figure 1A-E). This is reflected in the increased number of residues (a total of 1,089 for the 1247 residue ectodomain construct) that can be fitted into the density map, especially in the NTD. Compared to the 'looser' all-down version of the trimer (PDB: 6VXX (Walls et al., 2020)) an extra 120 residues are visualized and the additional stability of the protein has permitted a high-resolution analysis (3.1 Å) from just 36,000 particles. In both the 'all-down' and the 'locked' conformations the RBDs point downward in a 3-fold symmetric arrangement. Whilst they are broadly similar in structure, there are significant differences. The comparison shows that for the 'locked' form, the NTDs are rotated by approximately 7° and slide inwards towards the RBDs by ~ 7-10 Å. The RBDs in turn are shifted towards the central axis by 2-3 Å (the distance

6

between the N-termini of adjacent RBDs decreases by almost 6 Å) (Figure 1A, Video S1). The tighter packing around the 3-fold axis locks the NTD loops into a more ordered structure, as reflected in the improved density maps. RBD packing around the 3-fold axis also leads to the ordering of a segment of S2 (T827- F855) (Figure S4) where Y837 interacts with the main chain of the adjacent protomer residues P589' and R634' (' denotes neighbouring protomer) and the side chains L841 and F855 form a hydrophobic patch with residues V551', T553' and F592'. This region of S2 is adjacent to the fusion peptide and we suggest that this structure in the 'locked' form helps secure the prefusion state of the molecule.

In the 'locked' conformation, the '2P' mutations K986P and V987P, introduced to prevent conversion to the postfusion conformation (Pallesen et al., 2017), are close to residue D427 (Figure S4). D427 is part of a conserved patch on the RBD that 'caps' the 986-987 hinge region, preventing premature conversion to the postfusion form (Zhou et al., 2020). In the wild-type Spike, K986 could form a salt bridge with D427' to stabilize the 'locked' form but these salt bridges would be lost in the more open forms of the Spike (Figure S4). This latch in the wild-type Spike suggests that the 'locked' conformation is a more realistic representation of the 'resting' prefusion Spike conformation present at the viral surface, with the previously reported 'all-down' and 'one-up' structures transient conformations occurring just prior to ACE2 engagement, perhaps following a cellular cue. Consistent with this, a recent structural analysis of wild-type Spike lacking the 2P mutation revealed a similar such 'locked' state, with no sign of the more open prefusion Spike structures reported for the 2P stabilized molecule (Cai et al., 2020), however the coordinates are not available precluding a detailed comparison.

Glycosylation is often considered to mask potentially immunogenic protein epitopes. Three conserved glycosylation sites, located at N165, N234 and N343, have been reported for

stabilized SARS-CoV-2 Spike expressed in 293F cells (Watanabe et al., 2020). All three lie close to RBDs in the 'locked' conformation, well positioned to shield the ACE2 binding site from immune recognition. In addition, the first or second glycosidic rings of the N165 and N234 glycans make polar interactions underneath the tip of the RBD, opposite to the ACE2 binding site: GlcNac$_{N165}$ with Y351' and GlcNac$_{N234}$ with R457' and S459' (Casalino et al., 2020) (Figure 1D). We would expect these to be recapitulated in the highly processed sugars of wild-type virus (Watanabe et al., 2020).

The 'locked' conformation is stabilized by a pocket factor which cross-links the RBDs

Careful inspection of the 'locked' structure revealed the presence of elongated continuous density extending deep inside the hydrophobic core of the RBD (Figures 1 and 2). The density occupies a curved pocket lodged between the ß-sheet and outer α-helices of some 589 Å^3 volume (Figure 2) not found in the previously reported 'all-down' and 'one-up' structures. The length of the density is consistent with an 18-carbon fatty acid such as stearic acid, and its curvature suggests an omega-3 or omega-6 fatty acid with at least two unsaturated carbons (Figure 1C). In the absence of a formal chemical analysis of the molecule we have modelled it as stearic acid, however note, by analogy with enterovirus pocket factors, that there might be a mixture of lipids (Smyth et al., 2003). The fatty acid fits almost entirely within the pocket (85% of the surface area buried) and is stabilized by a series of hydrophobic interactions (Figure 2E-F), whilst the polar/charged headgroup is exposed at the end of the pocket where it faces a neighbouring RBD, interacting with R408' and indirectly with Q409' (Figure 2F). Interestingly, while the tip of the RBD, which includes the ACE2 binding site, is highly variable across β-coronaviruses, the residues composing the hydrophobic pocket are largely conserved (Figure 3). Of the 27 residues shielding the lipid, 23 are conserved and for none of the variable residues does the side-chain interact directly with the lipid. This strong

8

conservation suggests that the pocket could be a common feature across β-coronaviruses. Residues D364-F374 move outwards, from the positions seen in previous structures, by ~6 Å to form the pocket entrance (Figure 2A-B, Video S2) adopting a conformation that has never previously been observed for the engineered 2P prefusion Spike (Walls et al., 2020; Wrapp et al., 2020) or the isolated RBD domain (for instance PDBIDs: 6VXX, 6VSB, 6VYB) (Figure 2C), but may be consistent with the form reported recently for the wild-type prefusion Spike (Cai et al., 2020) (coordinates not available). This reorganization allows favourable packing against the neighbouring RBD, especially *via* Y365 and Y369.

The locked conformation was observed alongside Spike trimers with "all-down" and "one-up" conformations in our single particle cryoEM data set. Inspection of the RBD density for the 'down' and 'up' conformations in our analyses and those reported previously (Huo et al., 2020; Walls et al., 2020; Wrapp et al., 2020) showed no evidence of pocket binder. However, since the density for RBDs in the 'up' conformation was of poor quality, we performed focused refinement of our data, isolating 'up' RBDs and their associated adjacent NTDs. This yielded an improved map with a resolution of 3.7 Å for these domains, sufficient to unambiguously assign side chains (Figure S3). The characteristic structure of residues D364-F374 confirmed no fatty acid was bound. Thus, the binding pocket is only found in the 'locked' conformation.

Hydrophobic molecules perturb stability of the prefusion Spike

We hypothesized that hydrophobic molecules might replace the bound pocket factor (or occupy empty pockets). If the pocket factor modulates stability of the 'locked' trimer these events might alter the stability of the prefusion Spike. We therefore assembled a range of potential ligands which were tested in a thermofluor assay (Walter et al., 2012) (Figure S5, Methods). A number of molecules were assayed, based on the observed density in the pocket, being mainly

fatty acids, pocket factors designed to replace fatty acids in enterovirus capsids (Tuthill et al., 2010) and compounds observed to bind Notum, which harbours a lipid binding pocket reminiscent of the one reported here (Kakugawa et al., 2015). The Spike exhibited two distinct melts, the first at ~53 °C, presumably a structural rearrangement, and a second melt, presumably the total denaturation of the protein, at ~68 °C. A number of compounds significantly reduced the first melting temperature (by 1.0 °C or more relative to a Spike with DMSO control). For both saturated and unsaturated fatty acids there was a clear trend with chain length (Figure S5). However no compounds showed a significant stabilisation of the Spike and we have no evidence that the compounds bind in the RBD pocket.

Discussion

Biological relevance of the 'locked' conformation.

We have described a well-ordered closed conformation of prefusion Spike that appears to result from the incorporation of an endogenous 'pocket-factor'. This 'locked' Spike conformation is stabilized by a series of specific, conserved interactions which tightly pack the pocket-filled RBD and NTD around the molecular 3-fold axis. It appears that RBD interactions generated in this structure stabilize regions of S2 involved in the structural transition to the postfusion state, notably the turn at residues 986-987 and the region adjacent to the fusion peptide (827- 855), thereby preventing premature conversion of the prefusion state. These data, together with the recent observation of a similar structure in wild-type Spike ectodomain (Cai et al., 2020), strongly suggest that before recognition by a cellular receptor, the 'locked' conformation will be the predominant prefusion form on the viral surface. We believe that this has escaped notice primarily because attention has focused on a so-called stabilized Spike ectodomain construct, that has been engineered to prevent the switch to helical conformation in residues 986 and 987. Unwittingly these changes actually destabilize the 'locked' Spike, by abrogating the attachment

10

of the region on the RBD which in the prefusion 'all-down conformation' appears to interact with the turn between the HR1 and CH helices forming a 'cap' or 'lid' to prevent the jack-knife helix extension and conversion to the postfusion form. Whilst the chemical identity of the lipid molecule we see bound in the pocket remains speculative, C18 fatty acids fit well with unsaturated Omega-3 or Omega-6 likely. It should also be borne in mind that the fatty acids available in the expression system might dictate what is recruited into the pocket, as has been observed for the analogous case of picornaviruses when VLPs of mammalian viruses were produced in insect or plant cells (Marsian et al., 2017; Ren et al., 2015), but note that our expression system uses mammalian cells. Fatty acid interactions between adjacent RBDs seem to play a key role in the stability of the 'locked' conformation. Interestingly fatty acids appear to bind to conserved regions of the Spike of an α-coronavirus (Kirchdoerfer et al., 2020), although binding does not occur in a pocket, and an array of hydrophobic compounds have been identified as stabilizers for MERS-CoV (ß subgroup C) (Su et al., 2016). The conservation of the residues forming the pocket we have identified (Figure 3) suggests that this locked conformation may occur across the β-coronaviruses as a 'resting conformation'. This seclusion of the RBD might facilitate release of the virus from the cell and there would also be a selective advantage in keeping the immunogenic sites, including the ACE2 binding site, as inaccessible as possible. It is possible that the lipid then slowly diffuses out or a trigger, for instance engagement of an abundant attachment receptor (such as sialic acid or heparan sulphate), might facilitate the release of the fatty acid, initiating the switch to the 'open' state and presentation of the ACE2 binding site. Our observations also prompt the suspicion that production and purification protocols currently used may remove most of the lipid, contributing to the observed instability of the Spike.

11

The "locked" Spike conformation has considerable potential relevance for therapeutic design. A mechanism, whereby pocket-bound endogenous fatty acids stabilize a metastable state ahead of cell entry is analogous to the strategy used by some other RNA viruses, namely enteroviruses, where release of pocket-bound fatty-acids initiates a conformational change which primes the virion for genome release (Wang et al., 2012), and pocket factor mimics that fail to release have potential as drugs (De Colibus et al., 2014). In the case of Spike, we suggest that pocket binders which either stabilize or destabilize the Spike would offer a plausible route to a therapeutic, and we demonstrate that several compounds significantly destabilize Spike. The pocket is therefore a putative target for antiviral design. Knowledge of the mechanism of stabilization of the prefusion form of Spike is also important for understanding the context of epitopes presented *in vivo* and hence optimisation of presentation in non-classical vaccines. It is remarkable that the locked state of Spike serves to hide three key neutralizing epitopes, namely those of (i) ACE2 competing antibodies, for instance (Wu et al., 2020), (ii) antibodies that bind the S2 'capping' epitope (exemplified by CR3022 (Huo et al., 2020)), and (iii) the extremely potent S309 antibody epitope (Pinto et al., 2020) (the major change in the conformation of the N343 glycan would be expected to disrupt antibody binding as the sugar contributes part of the epitope, Figure 2D, 3C). This suggests that a vaccine which presents these epitopes effectively, such as one using polyvalent displayed RBD domains might be more effective than a genuinely stable ('locked') prefusion Spike trimer. In summary, our findings are a further step towards understanding the structural dynamics of this complex and rather unstable molecule with implications for vaccines and antiviral therapy.

Figure Legends

Figure 1. Overall structure of locked trimeric structure of SARS-CoV-2 Spike. (A) Comparison of the lipid bound Spike with the previously observed RBD 'all down' (PDB ID, 6VXX) and 'one up' Spike structures (PDB ID, 6Z97). The protomers are coloured in pale blue, grey and purple respectively. Relative rotation and translation of the NTD and RBD are indicated. (B) Top view of the RBD and NTD domains of the 'locked' Spike showing the relative positions of the bound fatty acids (green surfaces), the ACE2 binding sites are shown in red, α2 helices that have largest movement in brown, glycans at N165 and N234 of the NTD and N343 of RBD as yellow surfaces. (C) Density of the bound fatty acid. (D) Closeup of the interactions between glycans (yellow) on N165 and N234 of protomer B NTD (grey) and protomer A RBD (cyan).

Figure 2. Details of the fatty acid binding site. (A) The fatty acid binding site in the locked Spike RBD. The polypeptide chain is drawn as blue ribbons, ACE2 epitope in red, and the segment with large conformational difference compared with the known Spike structures in orange. The fatty acid is shown as green sticks with density as surface, and protein side chains in contacts with the lipid as blue and orange sticks. (B) The collapsed fatty acid binding pocket of the RBD in the 'all-down' form of Spike. (C) Comparison of the fatty acid (green sticks) bound RBD of the 'locked' Spike (blue) with six crystal structures of RBD complexed with either ACE2 or Fab or nanobody (PDB IDs, 6YLA, 6ZCZ, 6M0J, 6LZG, 6YZ5 and 7BZ5). (D) Overlay of the fatty acid bound RBD (blue) with that from S309 Fab bound Spike (grey; PDB ID, 6WPS) showing the clashes of α2 helix with the glycans (grey sticks) that are part of the S309 binding epitope. (E) Ligplot showing the interactions of the lipid with residues lining the pocket. Hydrophobic interactions and hydrogen bond are shown as dotted lines. Residue

R408 from a neighbouring RBD marked with a *. (F) The head group of the bound fatty acid is stabilized by polar or charged residues from neighbouring RBD.

Figure 3. Conservation of the lipid binding site in the beta linage of coronaviruses and the hidden RBD epitopes of the locked Spike. (A) Sequence alignment of RBDs of three beta-coronaviruses. Conserved amino acids are shown in red background, residues shielding the fatty acid marked by blue or red (from neighbouring RBD) disks. (B), Sequence conservation the RBD determined by CONSERV showing strong conservation of residues forming the lipid binding pocket. (C) Surface representation of the RBD domains of the locked Spike with ACE2, CR3022 and S309 binding epitopes coloured in magenta, blue and cyan, respectively.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

Genetic constructs of spike ectodomain. The gene encoding amino acids 1-1208 of the SARS-CoV-2 spike glycoprotein ectodomain, with mutations of RRAR > GSAS at residues 682-685 (the furin cleavage site) and KV > PP at residues 986-987, as well as inclusion of a T4 fibritin trimerisation domain, a HRV 3C cleavage site, a His-8 tag and a Twin-Strep-tag at the C-terminus. As reported by Wrapp *et al*. (Wrapp et al., 2020). All vectors were sequenced to confirm clones were correct.

Protein production. Recombinant Spike ectodomain was expressed by transient transfection in HEK293S GnTI- cells (ATCC CRL-3022) for 9 days at 30 °C. Conditioned media was dialysed against 2x phosphate buffered saline pH 7.4 buffer. The Spike ectodomain was purified by immobilized metal affinity chromatography using Talon resin (Takara Bio) charged with cobalt followed by size exclusion chromatography using HiLoad 16/60 Superdex 200 column in 150 mM NaCl, 10 mM HEPES pH 8.0, 0.02% NaN_3 at 4 °C.

Complex Preparation and Cryo-EM Data Collection, Purified spike protein was buffer exchanged into 2 mM Tris pH 8.0, 200 mM NaCl, 0.02 % NaN_3 buffer using a desalting column (Zeba, Thermo Fisher). A final concentration of 0.2 mg/mL was incubated with CR3022 Fab (in the same buffer) in a 6:1 molar ratio (Fab to trimeric spike) at room temperature. Aliquots were taken at 50 minutes. Immediately an aliquot was taken 3 μL of it was applied to a holey carbon-coated 200 mesh copper grid (C-Flat, CF-2/1, Protochips) that had been freshly glow

15

discharged on high intensity for 20 s (Plasma Cleaner PDC-002-CE, Harrick Plasma) and excess liquid removed by blotting for 6 s with a blotting force of -1 using vitrobot filter paper (grade 595, Ted Pella Inc.) at 4.5 ºC, 100 % relative humidity. Blotted grids were then immediately plunge frozen using a Vitrobot Mark IV (Thermo Fisher). The analysis of the CR3022 bound fractions were analysed and published in an independent paper (BioRxiv: https://doi.org/10.1101/2020.05.05.079202).

Cryo-EM Data collection and processing. Data were collected on a Titan Krios G2 (Thermo Fischer) at 300 kV). Automated data collection was setup in EPU 2.5 using a K3 (Gatan) direct electron detector operating in super-resolution mode at 0.415 Å per pixel and a GIF Quantum energy filter (Gatan) with a 20 eV slit. A defocus range of -0.8 μm to -2.6 μm and total dose of ~ 42.0 e$^-$/Å^2 across 40 frames was applied to the sample. Sample-specific data collection parameters are summarized in Table S1.

Motion correction of super-resolution raw movies were processed in Relion 3.1 (Zivanov et al., 2018), using a five-by-five patch-based alignment, keeping all the frames and a two time data binning, resulting in a final pixel size of 0.83 Å per pixel. The contrast transfer function (CTF) of full-dose, non-weighted micrographs was estimated using Gctf-v.1.06 (Zhang, 2016). Poor-quality images were discarded after manual inspection. Particles were blob picked in Cryosparc v.2.14.1 (Punjani et al., 2017) and after inspection of 2D classes, two sets of classes of interest were selected to generate templates for complete particle picking. A total number of 3,423,951 particles were picked and a final number of 468,112 particles were selected after two rounds of 2D classification. Three ab-initio models were generated and refined using heterogenous refinement. One good class containing around 328,000 particles was refined using non-uniform refinement to 3.3 Å and particles were exported to RELION 3.1 using pyEM

16

csparc2star.py scripts (Asarnow et al., 2019). Bayesian polishing and per particle CTF refinement with beam tilt correction improved map resolution to 3.0 Å. For the 3D variability analysis, polished particles and model were imported into cryoSPARC v.2.14.0 and refined using the non-uniform refinement. The 3D variability analysis was then performed on the full map, asking for solving the three main modes on a structure low pass filtered to 6 Å. The results of this principal component analysis (PCA) were clustered in ten sub-populations and models were reconstructed for each of the individual clusters. One cluster was showing what we named the 'locked' conformation which was refined using the non-uniform refinement and C3 symmetry to 3.0 Å from 36,000 particles. In parallel, all the particles having a strong density for one of the RBD in the 'up' conformation were selected and refined using the non-uniform refinement and C1 symmetry to 3.1 Å from 191,000 particles. Local refinement after particle subtraction was performed on the RBD 'up' and some part of the S2 to allow good image alignment and another one focusing on the other part of the Spike, which were refined to 3.6 Å and 3.0 Å, respectively. Structures were modelled by first rigid body fitting the PDB ID, 6VYB (Walls et al., 2020) for the "up" conformation and by rigid body fitting individual domains for the "locked" conformation into the locally filtered map using UCSF Chimera (Pettersen et al., 2004) One cycle of rigid body real space refinement followed by manual adjustment in Coot (Emsley and Cowtan, 2004) was performed to correctly position the Cα chain into the density. Finally, cycles of PHENIX (Liebschner et al., 2019) real space refinement and manual building in Coot (Emsley and Cowtan, 2004) were used to improve model geometry. Map-to-model comparison in PHENIX (Liebschner et al., 2019) mtriage validated that no over-fitting was present in the structures. Model geometry was validated for all the refined models using MolProbity (Davis et al., 2004). All map and model statistics are detailed in Table S1.

Thermofluor assays were performed in triplicate in white 96-well PCR plates (4titude) using a RT-PCR machine (MX3005P, Agilent) with excitation and emission filters of 492 and 585nm. Experiments of 50 µL well volumes contained 0.2 µM protein (1-1.5 µg), a final dye concentration of 3X (SYPRO orange, Invitrogen) and 0.1M HEPES pH7.5, 0.1M NaCl. All screening compounds were dissolved in DMSO and used in large (50 00up to ~1000-fold) molar excess with a final concentration of 4% DMSO. A smaller excess of compound (~50-fold) was used if the background signal was seen to be so high that the melts could not be clearly observed: this was the case for the unsaturated fatty acids for example. To limit the self-quenching effect of the dye at elevated temperatures an "expanding sawtooth" profile was used in which the sample was held for 30 seconds at an incrementally increasing temperature from 25 to 98 °C before reading the fluorescence at 25 °C. Raw fluorescence data were exported and analysis was performed by fitting a 5-parameter sigmoid curve using the online software JTSA (Sacchi et al., 2015). The Tm was taken as being the midpoint of the curve. Further analysis was performed in Excel spreadsheets and Prism.

Data availability. All data are available from the corresponding authors and/or included in the manuscript or Supplementary Information. Cryo-EM density maps with the corresponding atomic coordinates have been deposited in the Electron Microscopy Data Bank and the Protein Data Bank, respectively, with accession codes EMD-XXX, PDB ID XXX 'locked' Spike (C3), EMD-XXX, PDB ID XXX ('locked' Spike RBD/NTD focused), EMD-XXX, PDB ID XXX ('all down' Spike (C3)), EMD-XXX, PDB ID XXX ('one up' Spike (C1)), EMD-XXX, PDB ID XXX ('one up' Spike RBD/NTD focused), EMD-XXX, PDB ID XXX ('one up' Spike core/RBD down focused).

MAIN TEXT STATEMENTS

Acknowledgements

Author Information

These authors contributed equally: L. Carrique, H.M.E. Duyvesteyn, T. Malinauskas, Yuguang Zhao.

Contributions

T.M., Y.Z., D.Z. and J.H. prepared the samples. H.M.E.D., R.R.R and P.N.M.S. performed

cryo-EM sample preparation. J.Radecke performed cryo-EM data collection. H.M.E.D. and

L.C. processed cryo-EM data. T.S.W. performed thermofluor expriments and analysed the

data. L.C., H.M.E.D., J.R., T.S.W, E.E.F. and D.I.S. analysed the results and wrote the

manuscript. All authors read and approved the manuscript.

Competing interests. The authors declare no competing interests.

Correspondence and requests for materials should be addressed to DIS/EEF

References

Amanat, F., and Krammer, F. (2020). SARS-CoV-2 Vaccines: Status Report. Immunity *52*, 583–589.

Asarnow, D., Palovcak, E., and Cheng, Y. (2019). asarnow/pyem: UCSF pyem v0.5.

Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S.M., Walsh, R.M., Rawson, S., Rits-Volloch, S., and Chen, B. (2020). Distinct conformational states of SARS-CoV-2 spike protein. BioRxiv 2020.05.16.099317.

Casalino, L., Gaieb, Z., Dommer, A.C., Harbison, A.M., Fogarty, C.A., Barros, E.P., Taylor, B.C., Fadda, E., and Amaro, R.E. (2020). Shielding and Beyond: The Roles of Glycans in SARS-CoV-2 Spike Protein. BioRxiv 2020.06.11.146522.

De Colibus, L., Wang, X., Spyrou, J.A.B., Kelly, J., Ren, J., Grimes, J., Puerstinger, G., Stonehouse, N., Walter, T.S., Hu, Z., et al. (2014). More-powerful virus inhibitors from structure-based analysis of HEV71 capsid-binding molecules. Nat. Struct. Mol. Biol. *21*, 282–288.

Davis, I.W., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic Acids Res. *32*, W615-9.

Duan, K., Liu, B., Li, C., Zhang, H., Yu, T., Qu, J., Zhou, M., Chen, L., Meng, S., Hu, Y., et al. (2020). Effectiveness of convalescent plasma therapy in severe COVID-19 patients. Proc. Natl. Acad. Sci. U. S. A. *117*, 9490–9496.

Emsley, P., and Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. *60*, 2126–2132.

Esté, J.A., and Telenti, A. (2007). HIV entry inhibitors. Lancet *370*, 81–88.

Fehr, A.R., and Perlman, S. (2015). Coronaviruses: An overview of their replication and pathogenesis. In Coronaviruses: Methods and Protocols, (Springer New York), pp. 1–23.

Fleishman, S.J., Whitehead, T.A., Ekiert, D.C., Dreyfus, C., Corn, J.E., Strauch, E.M., Wilson, I.A., and Baker, D. (2011). Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science (80-.). *332*, 816–821.

Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., et al. (2020). The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. *5*, 536–544.

Henderson, R., Edwards, R.J., Mansouri, K., Janowska, K., Stalls, V., Gobeil, S., Kopp, M., Hsu, A.L., Borgnia, M.J., Parks, R., et al. (2020). Controlling the SARS-CoV-2 Spike Glycoprotein Conformation. BioRxiv 2020.05.18.102087.

Huo, J., Zhao, Y., Ren, J., Zhou, D., Duyvesteyn, H.M., Ginn, H.M., Carrique, L., Malinauskas, T., Ruza, R.R., Shah, P.N., et al. (2020). Neutralization of SARS-CoV-2 by destruction of the prefusion Spike. BioRxiv 2020.05.05.079202.

Jiang, S., Hillyer, C., and Du, L. (2020). Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. Trends Immunol. *41*, 355–359.

Kakugawa, S., Langton, P.F., Zebisch, M., Howell, S.A., Chang, T.H., Liu, Y., Feizi, T., Bineva, G., O'Reilly, N., Snijders, A.P., et al. (2015). Notum deacylates Wnt proteins to suppress signalling activity. Nature *519*, 187–192.

Khan, S.A., Zia, K., Ashraf, S., Uddin, R., and Ul-Haq, Z. (2020). Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 *via* integrated computational approach. J. Biomol. Struct. Dyn. 1–10.

Kirchdoerfer, R.N., Bhandari, M., Martini, O., Sewall, L.M., Bangaru, S., Yoon, K.-J., and Ward, A.B. (2020). Structure and immune recognition of the porcine epidemic diarrhea virus spike protein. BioRxiv 2020.02.18.955195.

Liebschner, D., Afonine, P. V., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze, B.,

Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. Acta Crystallogr. Sect. D Struct. Biol. *75*, 861–877.

Marsian, J., Fox, H., Bahar, M.W., Kotecha, A., Fry, E.E., Stuart, D.I., Macadam, A.J., Rowlands, D.J., and Lomonossoff, G.P. (2017). Plant-made polio type 3 stabilized VLPs-A candidate synthetic polio vaccine. Nat. Commun. *8*.

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., and Agha, R. (2020). The socio-economic implications of the coronavirus pandemic (COVID-19): A review. Int. J. Surg. *78*, 185–193.

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

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

Pinto, D., Park, Y.J., Beltramello, M., Walls, A.C., Tortorici, M.A., Bianchi, S., Jaconi, S., Culap, K., Zatta, F., De Marco, A., et al. (2020). Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. Nature 1–10.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Ren, J., Wang, X., Zhu, L., Hu, Z., Gao, Q., Yang, P., Li, X., Wang, J., Shen, X., Fry, E.E., et al. (2015). Structures of Coxsackievirus A16 Capsids with Native Antigenicity: Implications for Particle Expansion, Receptor Binding, and Immunogenicity. J. Virol. *89*, 10500–10511.

Rey, F.A., and Lok, S.M. (2018). Common Features of Enveloped Viruses and Implications

for Immunogen Design for Next-Generation Vaccines. Cell *172*, 1319–1334.

Sacchi, L., Capozzi, D., Bellazzi, R., and Larizza, C. (2015). JTSA: An Open Source Framework for Time Series Abstractions. Comput Methods Programs Biomed. *121*, 175–188.

Smyth, M., Pettitt, T., Symonds, A., and Martin, J. (2003). Identification of the pocket factors in a picornavirus. Arch. Virol. *148*, 1225–1233.

Su, S., Wong, G., Shi, W., Liu, J., Lai, A.C.K., Zhou, J., Liu, W., Bi, Y., and Gao, G.F. (2016). Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. Trends Microbiol. *24*, 490–502.

Tuthill, T.J., Groppelli, E., Hogle, J.M., and Rowlands, D.J. (2010). Picornaviruses. In Current Topics in Microbiology and Immunology, (NIH Public Access), pp. 43–89.

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

Walter, T.S., Ren, J., Tuthill, T.J., Rowlands, D.J., Stuart, D.I., and Fry, E.E. (2012). A plate-based high-throughput assay for virus stability and vaccine formulation. J. Virol. Methods *185*.

Wang, X., Peng, W., Ren, J., Hu, Z., Xu, J., Lou, Z., Li, X., Yin, W., Shen, X., Porta, C., et al. (2012). A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71. Nat. Struct. Mol. Biol. *19*, 424–429.

Watanabe, Y., Allen, J.D., Wrapp, D., McLellan, J.S., and Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. Science (80-.). eabb9983.

Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.-L., Abiona, O., Graham, B.S., and McLellan, J.S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science (80-.). *367*, 1260 LP – 1263.

Wu, Y., Wang, F., Shen, C., Peng, W., Li, D., Zhao, C., Li, Z., Li, S., Bi, Y., Yang, Y., et al. (2020). A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding

to its receptor ACE2. Science (80-.). *368*, eabc2241.

Xia, S., Yan, L., Xu, W., Agrawal, A.S., Algaissi, A., Tseng, C.T.K., Wang, Q., Du, L., Tan, W., Wilson, I.A., et al. (2019). A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike. Sci. Adv. *5*, eaav4580.

Yan, L., Meng, B., Xiang, J., Wilson, I.A., and Yang, B. (2018). Crystal structure of the post-fusion core of the Human coronavirus 229E spike protein at 1.86 Å resolution. Acta Crystallogr. Sect. D Struct. Biol. *74*, 841–851.

Yin, W., Mao, C., Luan, X., Shen, D.-D., Shen, Q., Su, H., Wang, X., Zhou, F., Zhao, W., Gao, M., et al. (2020). Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science (80-.). eabc1560.

Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Lu, G., Wu, Y., Yan, J., Shi, Y., et al. (2017). Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. Nat. Commun. *8*.

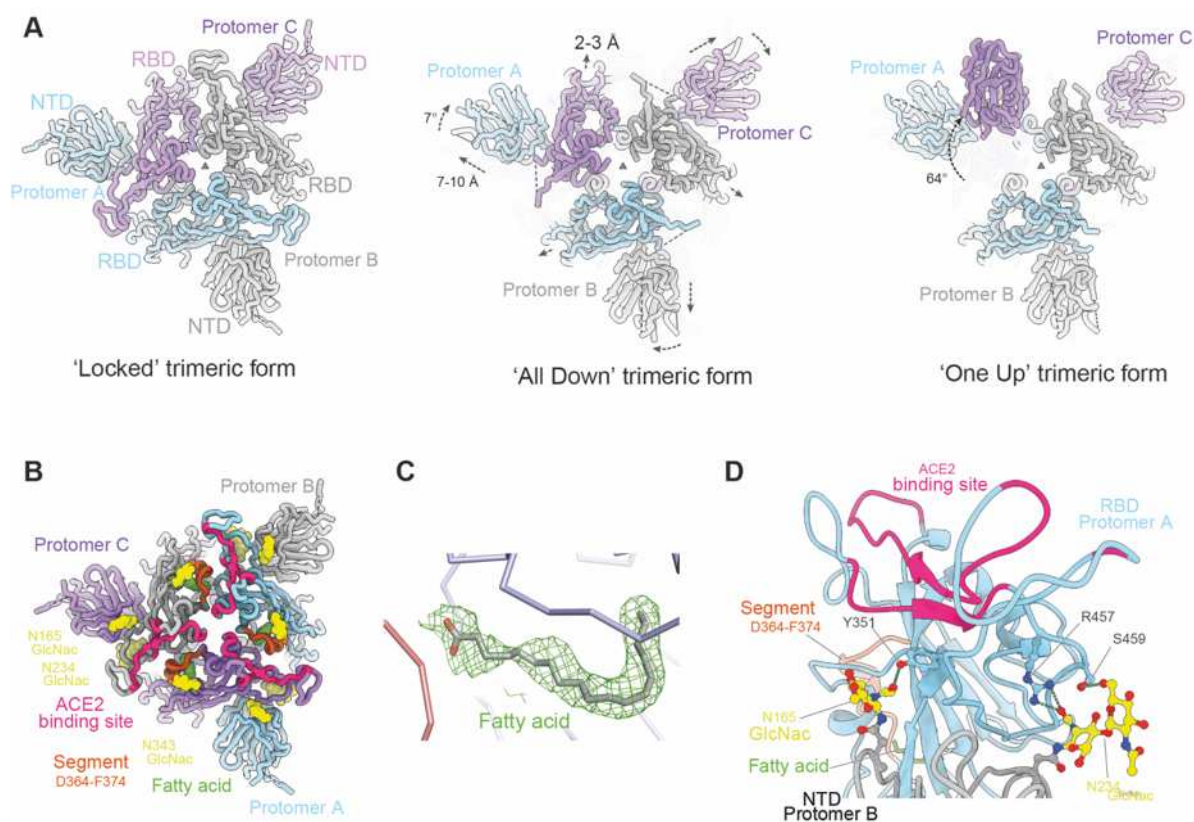Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., and Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved a-ketoamide inhibitors. Science (80-.). *368*, 409–412.

Zhao, Y., Ren, J., Harlos, K., Jones, D.M., Zeltina, A., Bowden, T.A., Padilla-Parra, S., Fry, E.E., and Stuart, D.I. (2016). Toremifene interacts with and destabilizes the Ebola virus glycoprotein. Nature *535*, 169–172.
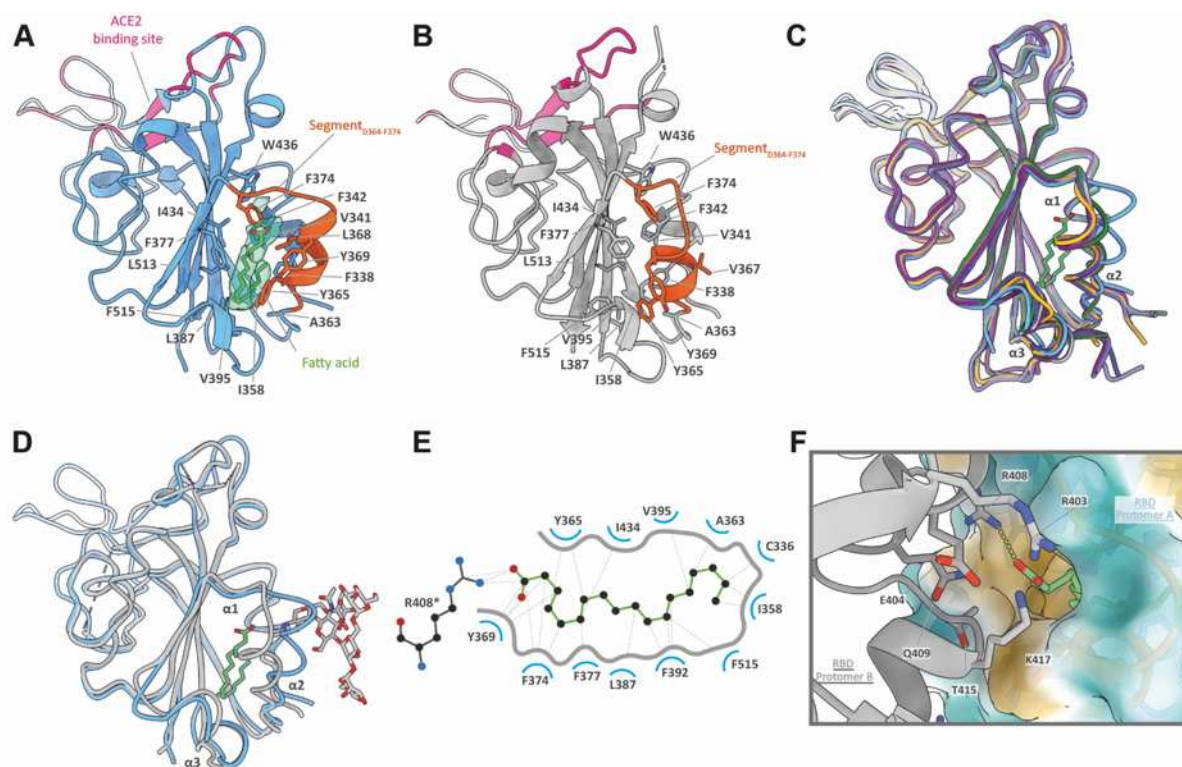
Zhou, D., Duyvesteyn, H.M., Chen, C.-P., Huang, C.-G., Chen, T.-H., Shih, S.-R., Lin, Y.-C., Cheng, C.-Y., Cheng, S.-H., Huang, Y.-C., et al. (2020). Structural basis for the neutralization of SARS-CoV-2 by an antibody from a convalescent patient. BioRxiv 2020.06.12.148387.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J.H., Lindahl, E., and

Scheres, S.H.W. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*.
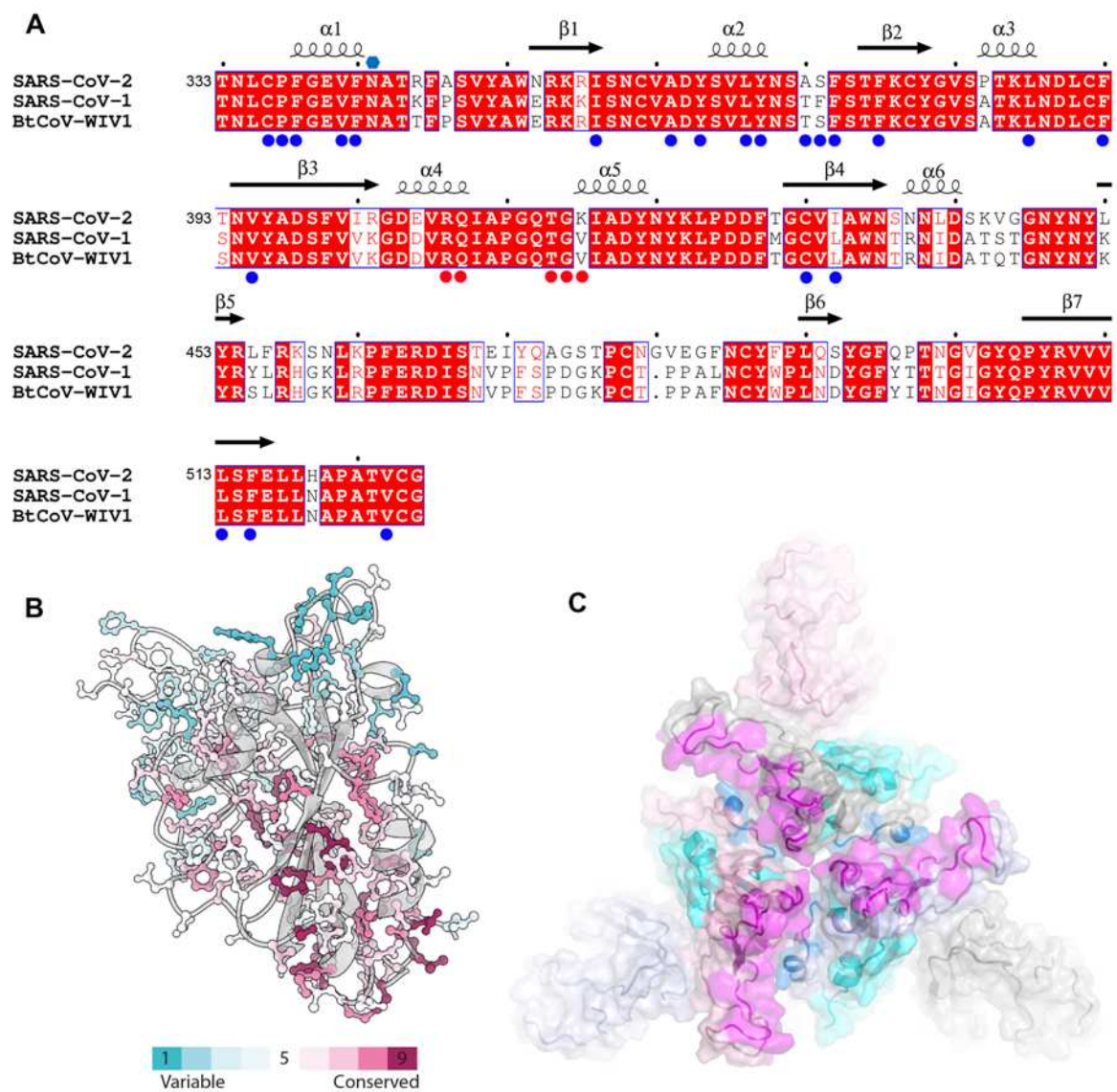
Figure 1

Figure 2

Figure 3