bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Structures and function of locked conformations of SARS-CoV-2 spike

Kun Qu[1,#], Xiaoli Xiong[1,2,#,*], Katarzyna A. Ciazynska[1], Andrew P. Carter[1], John A. G. Briggs[1,*]

1. Structural Studies Division, Medical Research Council Laboratory of Molecular Biology, Cambridge, United Kingdom.

2. Bioland Laboratory (Guangzhou Regenerative Medicine and Health - Guangdong Laboratory), Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, China.

# equal contribution

* correspondence to XX (xiong_xiaoli@gibh.ac.cn) or JAGB (jbriggs@mrc-lmb.cam.ac.uk)

Abstract

The spike protein (S) of SARS-CoV-2 has been observed in three distinct pre-fusion conformations: locked, closed and open. Of these, the locked conformation was not previously observed for SARS-CoV-1 S and its function remains poorly understood. Here we engineered a SARS-CoV-2 S protein construct "S-R/x3" to arrest SARS-CoV-2 spikes in the locked conformation by a disulfide bond. Using this construct we determined high-resolution structures revealing two distinct locked states, with or without the D614G substitution that has become fixed in the globally circulating SARS-CoV-2 strains. The D614G mutation induces a structural change in domain D from locked-1 to locked-2 conformation to alter spike dynamics, promoting transition into the closed conformation from which opening of the receptor binding domain is permitted. The transition from locked to closed conformations is additionally promoted by a change from low to neutral pH. We propose that the locked conformations of S are present in the acidic cellular compartments where virus is assembled and egresses. In this model, release of the virion into the neutral pH extracellular space would favour transition to the closed form which itself can stochastically transition into the open form. The S-R/x3 construct provides a tool for the further structural and functional characterization of the locked conformations of S, as well as how sequence changes might alter S assembly and regulation of receptor binding domain dynamics.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

The spike (S) protein of coronaviruses is responsible for interaction with cellular receptors and fusion with the target cell membrane (Li, 2016). It is the main target of the immune system, therefore the focus for vaccine and therapeutics development. Most candidate Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) vaccines utilise S protein or its derivatives as the immunogen (Krammer, 2020) and a number of antibodies targeting the SARS-CoV-2 S protein are under development for COVID-19 treatment (Klasse and Moore, 2020).

The S protein of SARS-CoV-2 is a large, 1273 residue, type I fusion protein, that can be processed by host proteases into two parts, S1 and S2. In the prefusion, trimeric form, S1 forms multiple folded domains. From N-terminus to C-terminus, they are the N-terminal domain (NTD, also called domain A, residues 13-307), receptor binding domain (RBD, also called domain B, residues 330-528), Domain C (residues 320-330 and 529-592), and Domain D (residues 308-319 and 593-699) (Walls et al., 2020; Wrapp et al., 2020; Xiong et al., 2020). Neutralising antibodies have been identified that target NTD and RBD (Gavor et al., 2020). S2 contains the fusion peptide (817-832), S2' cleavage site (811-815), and a fusion peptide proximal region (FPPR) C-terminal to the fusion peptide (previously referred to by us as the 833-855 motif (Xiong et al., 2020)) that we and others have previously suggested may play a role in regulating RBD opening (Bangaru et al., 2020; Cai et al., 2020; Xiong et al., 2020) (**Fig. 1a**).

SARS-CoV-2 spike protein has been captured in four distinct conformations by cryo-electron microscopy (cryoEM): three prefusion conformations – locked (Bangaru et al., 2020; Cai et al., 2020; Toelzer et al., 2020; Xiong et al., 2020), closed (Walls et al., 2020; Wrapp et al., 2020), open (Walls et al., 2020; Wrapp et al., 2020); and one post-fusion conformation (Cai et al., 2020; Fan et al., 2020). The different conformations of S expose different epitopes (Barnes et al., 2020) and therefore induce different immune responses (Carnell et al., 2021). The different conformations must therefore play distinct roles in the infection cycle.

In the locked and closed conformations, the three copies of RBD in each spike trimer lie down on the top of the spike such that each RBD interacts in trans with the NTD of the neighbouring S protomer, hiding the receptor binding site. In the locked conformation, the three RBDs are in very close proximity at the 3-fold axis, and a disulfide-bond stabilised helix-turn-helix motif (FPPR motif) is formed below the RBD that sterically prevents RBD opening (Bangaru et al., 2020; Cai et al., 2020; Toelzer et al., 2020; Xiong et al., 2020). Additionally, a linoleic acid is bound within the RBD and this ligand has been proposed to stabilise the locked conformation (Toelzer et al., 2020). The locked conformation is adopted in solution by the Novavax NVX-CoV2373 vaccine candidate (Bangaru et al., 2020). In structures of the locked conformation, both NTD and RBD are well resolved, suggesting that the locked spike has a rigid overall structure. In the closed conformation, the RBDs are slightly further apart than in the locked conformation and the FPPR motif is not folded (Bangaru et al., 2020; Cai et al., 2020; Toelzer et al., 2020; Xiong et al., 2020). As a result, the NTD and RBD exhibit considerable dynamics and are less well resolved than the core part of the protein in cryo-EM structures. Due to such dynamics, the RBD in the closed conformation can readily lift up to transit into the open conformation (Walls et al., 2020;

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Wrapp et al., 2020). In the open conformation, one or more RBDs are raised up to expose the receptor binding loops. ACE2 or antibody binding to these loops appears to maintain the spike in an open conformation (Benton et al., 2020; Xu et al., 2021). It is believed that opening of the cleaved spike ultimately leads to structural transition into the postfusion conformation. The virus uses the dramatic structural refolding from the prefusion conformation to the lowest-energy postfusion conformation to drive membrane fusion (Bosch et al., 2003).

Among the observed conformational states of SARS-CoV-2 spike, the locked conformation is unique in that it has not been observed for the SARS-CoV-1 spike. Other coronavirus spikes do have locked-like conformations featuring motifs that preventing RBD opening, but for many of them, stochastic formation of an open conformation has not been observed (Yuan et al., 2017). The early cryoEM studies of purified SARS-CoV-2 spike ectodomain did not identify the locked conformation, and we previously found that only a small fraction of trimeric ectodomains adopt the locked conformation (Xiong et al., 2020), suggesting this conformation is transient. However, other cryoEM studies showed that the spike protein can predominantly adopt the locked conformation under certain conditions: purified full-length spike proteins including the transmembrane region, primarily adopt locked conformations (Bangaru et al., 2020; Cai et al., 2020); insect-cell-expressed spike ectodomains also primarily adopt the locked conformation, perhaps due to stabilizing lipid in the insect cell media (Toelzer et al., 2020). The locked conformation was not observed in the high-resolution structures of spike proteins on virions (Ke et al., 2020).

Considering the above observations, the role of the locked conformation in the SARS-CoV-2 life cycle is unclear. Here, based on our previous locked spike structure (Xiong et al., 2020), we engineered a disulfide pair which stabilises the SARS-CoV-2 spike in the locked conformation. This construct provides a tool to characterize the structure and function of the locked form of S and to assess how mutations and other environmental factors could affect its function. Newly obtained information regarding the “locked” spike allowed us to consider its functional role in the context of SARS-CoV-2 life cycle.

Results and Discussion

Design of a spike protein ectodomain stabilized in the locked conformation

We previously did not observe a locked conformation for S-R (S with a deletion at the Furin cleavage site leaving only a single arginine residue), while for S-R/PP (S-R additionally containing the widely-used, stabilizing double-proline mutation) the locked conformation was rare, <10% of S trimers when imaged by cryoEM (Xiong et al., 2020). The scarcity of the locked conformation for the expressed ectodomain makes it difficult to study. To overcome this challenge, based on the locked SARS-CoV-2 S-R/PP and closed S-R/PP/x1 and S-R/x2 spike structures determined previously (Xiong et al., 2020), we engineered cysteines residues replacing positions D427 and V987 to generate a new disulfide-stabilised S protein S-R/x3. We predicted that S-R/x3 should be predominantly in the locked conformation (**Fig. 1a**). Under non-reducing conditions, S-R/x3 ran exclusively as trimer in SDS-PAGE gel, and the trimer is converted to monomer under reducing conditions (**Fig. 1b**). This behaviour is similar to our previously engineered S-R/x2, suggesting efficient formation of disulfide bond

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

between the 2 engineered cysteines. Negative stain electron microscopy (EM) images of purified S-R/x3 show well-formed S trimers (**Fig. 1c**). We additionally introduced a D614G substitution into the S-R/x3 spike to obtain S-R/x3/D614G. Similar to S-R/x3, purified S-R/x3/D614G ran as disulfide linked trimers in SDS-PAGE gel under non-reducing conditions, and negative stain EM images show well-formed trimers (**Fig. 1b,c**).

Cryo-EM structure of S-R/x3

We performed cryo-EM of S-R/x3 trimers as described previously for S-R/x2 (Xiong et al., 2020). Consistent with the predicted effect of the engineered disulfide bond, 3D classification of S-R/x3 particles showed that majority of the particles (77%) are in the locked conformation, while 23% of the particles are in the closed conformation (**Fig. S1**). We obtained a reconstruction of the locked conformation at 2.6 Å with well resolved NTD and RBD densities (**Fig. S2**). S trimers in the locked conformation are characterised by structural rigidity, showing minimal local dynamics in NTD and RBD (**Fig. S2**). The engineered disulfide bond can be identified in the cryoEM density (**Fig. 2a**). Consistent with our previous observation of the locked conformation (Xiong et al., 2020), the RBDs are tightly clustered around the three-fold symmetry axis (**Fig. S3**), linoleic acid is bound into the described pocket in the RBD (**Fig. 2a**), and residues 833-855 are folded into a helix-loop-helix fusion peptide proximal region (FPPR) motif (**Fig. 2a**). We suggest that folding of the FPPR motif sterically prevents RBD opening.

The improved resolution relative to our previous reconstruction of the S-R/PP locked structure, allowed us to build an atomic model including Domain D loop 633-641 (**Fig. 2**) where side chains were previously poorly resolved. Residues between 617-633 form a large disordered loop, which is not resolved in the cryoEM density. In structures of an unusual dimer of locked spikes, this loop has been seen to make contacts across the dimer interface (Bangaru et al., 2020).

Cryo-EM Structure of S-R/x3/D614G

We also determined the structure of the S-R/x3/D614G trimers. The proportion of locked conformations is greatly reduced compared to S-R/x3, constituting only ~20% of total particles (**Fig. S1**). Despite the presence of the x3 disulfide bond, 80% of the particles are in a closed conformation. The closed conformation accommodates the disulfide bond by motion of the RBD towards S2 (**Fig. S3**). In the S-R/x3/D614G locked structure, features observed in the S-R/x3 locked structure are present including the more tightly clustered RBD around the 3-fold axis, linoleic acid in the RBD and a folded FPPR motif. However, the S-R/x3/D614G locked structure differs from the S-R/x3 locked structure at domain D (**Fig. 2,3**). Within this domain, residues 617-641, which adopts a disordered loop structure in locked S-R/x3, refold into two small a-helices connected by a loop in locked S-R/x3/D614G. This change also affected structural features interacting with domain D, which are described below.

Comparison of the two different locked conformations

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

By comparing the locked conformations in S-R/x3 and S-R/x3/D614G, we observed that SARS-CoV-2 spike protein can adopt 2 different locked conformations (**Fig. 3**). We will refer to the conformation observed for 614D (S-R/x3) as "locked-1" while the one observed for 614G (S-R/x3/D614G) as "locked-2". The two conformations differ primarily at residues 617-641 within domain D, and due to this structural perturbation, there are differences in the FPPR motif and in residues 316-325 which are located in the hinge region connecting domain C and D. In the 614D locked-1 structure, the long sidechains of R634 and Y636 interdigitate with the sidechains of F318 and Y837 (**Fig. 3a**). In this interaction, R634 in domain D is sandwiched between F318 of the hinge region, and Y837 of the FPPR motif, bridging these three structural features. It appears likely that this zip-locking interaction maintains residues 617-641 in a loop structure. This arrangement positions loop residue R634 approximately 9 Å above D614.

In the 614D locked-1 structure, we and others observed a salt bridge between D614 and K854 within the FPPR motif (Cai et al., 2020; Toelzer et al., 2020; Xiong et al., 2020) (**Fig. 3a**). The D614G mutation abolishes this salt bridge in the locked-2 structure (**Fig. 3b**). More significantly, the 617-641 loop refolds into two small alpha-helices (**Fig. 3b**). We suggest that the mutation of negatively charged D614 to neutral G and loss of the salt bridge alters the local electrostatic interactions, preventing R634 binding between F318 and Y837, and triggering refolding of the 617-641 loop. A similar helical structure has been observed for this region in domain D of other beta-coronavirus spike proteins including Murine hepatitis virus (MHV) (Walls et al., 2016), Severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1) (Yuan et al., 2017), OC43-CoV (Tortorici et al., 2019), Middle-East Respiratory Syndrome coronavirus (MERS-CoV) (Walls et al., 2019)spikes (**Fig. S4**). Transition from the locked-1 to the locked-2 conformation breaks the interactions that link residues in domain D, the hinge region between domain C and domain D, and the FPPR motif. As a result, Y837 in the FPPR motif moves away from the RBD by approximately 4 Å while F318 rotates towards the inside of the spike, resulting a movement of approximately 7Å in the domain C-D hinge region (**Fig. 3c**). Further movement is required at the hinge region for the spike to transition from locked-2 to the closed conformation, this change results in a twisting motion of domain C and the RBD, loosening RBD packing, and thereby presumably readying the RBD for opening. Therefore, free movement of the F318 and the surrounding hinge region appears to be important for structural transition from locked-1 to locked-2 and from locked-2 to closed/open conformations. The absence of the interlocking interactions involving R634, Y636, F318 and Y837 in the S-R/x3/D614G structures provides an explanation for the observation that D614G promotes RBD opening.

Comparison to other Locked spike structures

Most SARS-CoV-2 spike structures deposited in the PDB are in the open or closed conformations and have a disordered FPPR motif and unresolved 617-641 region (https://www.ebi.ac.uk/pdbe/pdbe-kb/protein/P0DTC2). Only a few studies have captured S trimers in locked conformations. Two studies imaged purified full-length spike protein by cryoEM, with or without the PP mutation: these full-length structures are in the locked-1 conformation (Bangaru et al., 2020; Cai et al., 2020) (**Fig. S5a,h**). Soluble S-GSAS/PP trimers (where the furin cleavage site is replaced with a GSAS sequence) purified from insect cells are predominantly in the locked-1 conformation (Toelzer et al., 2020) (**Fig. S5c,d**), while

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

there are two studies describing S-GSAS/PP trimers purified from mammalian cells that are also predominantly in the locked-1 conformation (Wrobel et al., 2020; Xu et al., 2021) (**Fig. S5e-g**). We have previously described that a small proportion of S-R/PP and S-R/PP/x1 trimers are in the locked-1 conformation (Xiong et al., 2020). In all these studies, where the locked-1 conformation has been observed, the spike features a more constricted closed RBD and binding of lipid within the RBD as well as a folded FPPR motif. In all these structures, residues 617-641 form a loop with a disordered central region (618-632, dashed line in Fig. 3), and residues 634 and 636 interact with F318, as we observed here for the locked-1 conformation. Ward and colleagues have observed an unusual spike dimer in which the otherwise disordered 617-641 loop forms a long, extended loop structure that contacts the dimeric partner (Bangaru et al., 2020) (**Fig. S5a**). The reason why locked conformations are observed under some conditions but not others is not completely clear. The purification procedure; the presence of membrane anchor in the full-length spike (Bangaru et al., 2020; Cai et al., 2020); differences in expression conditions - added lipid in insect cell media, low pH condition of insect cell media (Toelzer et al., 2020); EM grid preparation methods – use of a high octyl-glucoside concentration (Wrobel et al., 2020), and differences in reagents as concluded by Xu et al (Xu et al., 2021), may have all played a part.

Zhou and colleagues investigated the effect of low pH on ACE2 and antibody binding to the SARS-CoV-2 spike protein (Zhou et al., 2020). Their structure at pH 4 is most similar to the locked-2 conformation: the RBD adopts the more constricted position, the FPPR motif is ordered, and the 617-641 loop adopts an a-helical structure (**Fig S5i-l**). Low pH can therefore convert the "closed" conformation to locked-2 conformation. In contrast to the locked-2 conformation that we have observed, the low pH forms do not contain lipid bound within the RBD.

The locked-1 conformation can transition to the closed conformation over time.

We observed that after storing the S-R/x3 spike at 4°C for 40 days the number of well-formed trimeric spikes in negative stain EM was reduced by 30-40%. CryoEM imaging revealed that the remaining trimers had transitioned into the "closed" conformation (**Fig. S1**), and lipid was no longer bound in the RBD. This suggests that the locked-1 conformation is metastable at pH 7.4, and spontaneously transitions into the closed conformation. It is consistent with the general observation that most published ectodomain constructs (not stabilised by the x3 crosslink) adopt the closed/open conformations rather than the locked conformation. We incubated the now "closed" spike at pH 5.0 overnight before determining its structure by cryoEM. We found that low pH treatment had converted most of the closed spike to the locked-2 conformation (**Fig. S1**), without lipid bound in the RBD, consistent with the observations of Zhou et al (Zhou et al., 2020). A comparison of the closed and locked-2 conformations suggests to us that H519 within RBD and H49 within domain A, which are located close to where the RBD and NTD, respectively, contact the trimer core, may play a role in mediating pH-dependent structural changes (**Fig. S6**).

These observations demonstrate that although low pH can maintain the RBD in the locked conformation, low pH alone is not sufficient to revert the closed conformation back to the locked-1 conformation. It further demonstrates that the 614D spike, like the 614G spike, is able to adopt the locked-2 conformation. We speculate that breaking the interdigitation of

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

residues in the 617-641 loop results in a locked-1 to closed transition and that this is a largely irreversible process. Under conditions, such as low pH, which favour transition from the closed back to locked conformation, the complex interdigitation cannot refold, resulting in a trimer in the locked-2 conformation.

The functional role of the locked-1 and locked-2 structures.

The disulfide-bond engineered into the S-R/x3 construct stabilizes the S trimer in the locked-1 conformation. In this conformation F318, R634, Y636 and Y837 interdigitate to link domain D, the hinge connecting domain C and D, and the FPPR motif. This interaction bridges the three structural features and we hypothesise that it locks the trimer, preventing RBD opening. At neutral pH, this structure is metastable and transitions over time into the closed conformation, from which stochastic RBD opening can take place. Lowering the pH of the closed spike, which promotes the close-apposition of RBDs in the locked conformation, leads to reversion into the locked-2 conformation, perhaps because the complex interdigitation of residues in domain D in the locked-1 conformation cannot easily be reformed.

Removal of an electrostatic interaction between D614 and R634 in the D614G spike hinders formation of the interdigitating stabilising interaction in domain D. Therefore, the D614G spike adopts the locked-2 conformation, where domain D residues 617-641 fold into a helix-turn-helix motif similar to that observed in other beta-CoV spikes. This motif is located between the NTD and the hinge region leading to domain C and RBD. We speculate that the alternate conformations observed for domain D are part of a control mechanism regulating RBD opening and modulating receptor binding.

The absence of the interdigitating interaction network favours the transition to the closed conformation, as illustrated by the observation that the majority of D614G spike particles adopt the “closed” conformation. This is consistent with the observation that constructs bearing the D614G substitution are more susceptible to RBD opening, and provides a structural explanation for the change in conformational dynamics (Gobeil et al., 2021; Yurkovetskiy et al., 2020). A similar observation has also been made for full-length SARS-CoV-2 spike protein where the D614G substitution also induces disordering of the FPPR in domain D and increased number of open spikes (Zhang et al., 2020).

Beta-coronaviruses have recently been shown to assemble in the low pH, high lipid environment of the ERGIC, and egress through the acidic environment of endosome (Ghosh et al., 2020) (**Fig 4a**). Given that low pH and lipid binding both favour locked conformations, we speculate that during virus assembly, both the 614D and D614G spikes are in their respective locked-1 and locked-2 conformations, and that this provides a mechanism to prevent premature transition into open or post-fusion conformations during virus assembly. Once viruses are released into the neutral pH environment outside of the cell, over time the spike will transition to the closed conformation where the RBD can open and mediate binding to target cells (**Fig 4b**). The D614G spike, which lacks the interactions in domain D to stabilise the RBD in the locked-1 conformation, is likely to exhibit a different structural dynamics that may leads to more rapid transition to the closed and subsequently the open form. The resulting change in receptor binding, cell entry and immune presentation

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

characteristics has presumably provided a transmission advantage leading to global dominance of the mutant virus (Korber et al., 2020).

Acknowledgements
We thank the staff of the MRC-LMB EM Facility their support; Jake Grimmett and Toby Darling for scientific computing infrastructure; and Patricia Edwards for supporting cell culture. This study was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-CoG-648432 MEMBRANEFUSION to JAGB), the Medical Research Council as part of United Kingdom Research and Innovation (MC_UP_A025_1011 to APC; MC_UP_1201/16 to JAGB). XX acknowledges start-up grants from the Chinese Academy of Sciences and Bioland Laboratory (GRMH-GL).

Author Contributions
XX and JAGB conceived the study; KQ, XX and JAGB planned the study. XX, APC and JAGB designed constructs which were made by APC. Protein purification was performed by XX and KAC. XX performed negative stain and prepared EM grids. KQ performed cryo-EM and processed EM data. KQ and XX built molecular models which KQ, XX, and JAGB interpreted. JAGB managed the project. XX, KQ and JAGB made figures and wrote the paper with input from all authors.

References

Afonine, P. V., Poon, B.K., Read, R.J., Sobolev, O. V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. Sect. D Struct. Biol. *74*, 531–544.

Bangaru, S., Ozorowski, G., Turner, H.L., Antanasijevic, A., Huang, D., Wang, X., Torres, J.L., Diedrich, J.K., Tian, J.-H., Portnoff, A.D., et al. (2020). Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. Science *370*, 1089–1094.

Barnes, C.O., Jette, C.A., Abernathy, M.E., Dam, K.-M.A., Esswein, S.R., Gristick, H.B., Malyutin, A.G., Sharaf, N.G., Huey-Tubman, K.E., Lee, Y.E., et al. (2020). SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. Nature *588*, 682–687.

Benton, D.J., Wrobel, A.G., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and Gamblin, S.J. (2020). Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. Nature *588*, 327–330.

Bosch, B.J., van der Zee, R., de Haan, C.A.M., and Rottier, P.J.M. (2003). The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. J. Virol. *77*, 8801–8811.

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. Science *369*, 1586–1592.

Carnell, G.W., Ciazynska, K.A., Wells, D.A., Xiong, X., Aguinam, E.T., McLaughlin, S.H., Mallery, D., Ebrahimi, S., Ceron-Gutierrez, L., James, L.C., et al. (2021). SARS-CoV-2 spike protein arrested in the closed state induces potent neutralizing responses. BioRxiv 2021.01.14.426695.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

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

Gavor, E., Choong, Y.K., Er, S.Y., Sivaraman, H., and Sivaraman, J. (2020). Structural Basis of SARS-CoV-2 and SARS-CoV Antibody Interactions. Trends Immunol. *41*, 1006–1022.

Ghosh, S., Dellibovi-Ragheb, T.A., Kerviel, A., Pak, E., Qiu, Q., Fisher, M., Takvorian, P.M., Bleck, C., Hsu, V.W., Fehr, A.R., et al. (2020). β-Coronaviruses Use Lysosomes for Egress Instead of the Biosynthetic Secretory Pathway. Cell *183*, 1520-1535.e14.

Gobeil, S.M.-C., Janowska, K., McDowell, S., Mansouri, K., Parks, R., Manne, K., Stalls, V., Kopp, M.F., Henderson, R., Edwards, R.J., et al. (2021). D614G Mutation Alters SARS-CoV-2 Spike Conformation and Enhances Protease Cleavage at the S1/S2 Junction. Cell Rep. *34*, 108630.

Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., McKeane, L., Nakane, T., Zivanov, J., Neufeldt, C.J., Cerikan, B., et al. (2020). Structures and distributions of SARS-CoV-2 spike proteins on intact virions. Nature *588*, 498–502.

Kidmose, R.T., Juhl, J., Nissen, P., Boesen, T., Karlsen, J.L., and Pedersen, B.P. (2019). Namdinator - Automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. IUCrJ *6*, 526–531.

Klasse, P., and Moore, J.P. (2020). Antibodies to SARS-CoV-2 and their potential for therapeutic passive immunization. Elife *9*.

Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E.E., Bhattacharya, T., Foley, B., et al. (2020). Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell *182*, 812-827.e19.

Krammer, F. (2020). SARS-CoV-2 vaccines in development. Nature *586*, 516–527.

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

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

Toelzer, C., Gupta, K., Yadav, S.K.N., Borucu, U., Davidson, A.D., Kavanagh Williamson, M., Shoemark, D.K., Garzoni, F., Staufer, O., Milligan, R., et al. (2020). Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. Science *370*, 725–730.

Tortorici, M.A., Walls, A.C., Lang, Y., Wang, C., Li, Z., Koerhuis, D., Boons, G.-J., Bosch, B.-J., Rey, F.A., de Groot, R.J., et al. (2019). Structural basis for human coronavirus attachment to sialic acid receptors. Nat. Struct. Mol. Biol. *26*, 481–489.

Walls, A.C., Tortorici, M.A., Bosch, B.J., Frenz, B., Rottier, P.J.M., DiMaio, F., Rey, F.A., and Veesler, D. (2016). Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. Nature *531*, 114–117.

Walls, A.C., Xiong, X., Park, Y.J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., et al. (2019). Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. Cell *176*, 1026-1039.e15.

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.

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

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

conformation. Science *367*, 1260–1263.

Wrobel, A.G., Benton, D.J., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and Gamblin, S.J. (2020). SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. Nat. Struct. Mol. Biol. *27*, 763–767.

Xiong, X., Qu, K., Ciazynska, K.A., Hosmillo, M., Carter, A.P., Ebrahimi, S., Ke, Z., Scheres, S.H.W., Bergamaschi, L., Grice, G.L., et al. (2020). A thermostable, closed SARS-CoV-2 spike protein trimer. Nat. Struct. Mol. Biol. *27*, 934–941.

Xu, C., Wang, Y., Liu, C., Zhang, C., Han, W., Hong, X., Wang, Y., Hong, Q., Wang, S., Zhao, Q., et al. (2021). Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. Sci. Adv. *7*, eabe5575.

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*, 1–9.

Yurkovetskiy, L., Wang, X., Pascal, K.E., Tomkins-Tinch, C., Nyalile, T.P., Wang, Y., Baum, A., Diehl, W.E., Dauphin, A., Carbone, C., et al. (2020). Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. Cell *183*, 739-751.e8.

Zhang, J., Cai, Y., Xiao, T., Lu, J., Peng, H., Sterling, S.M., Walsh, R.M., Rits-Volloch, S., Sliz, P., and Chen, B. (2020). Structural impact on SARS-CoV-2 spike protein by D614G substitution. BioRxiv Prepr. Serv. Biol.

Zhou, T., Tsybovsky, Y., Gorman, J., Rapp, M., Cerutti, G., Chuang, G.-Y., Katsamba, P.S., Sampson, J.M., Schön, A., Bimela, J., et al. (2020). Cryo-EM Structures of SARS-CoV-2 Spike without and with ACE2 Reveal a pH-Dependent Switch to Mediate Endosomal Positioning of Receptor-Binding Domains. Cell Host Microbe *28*, 867-879.e5.

Zivanov, J., Nakane, T., and Scheres, S.H.W. (2019). A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. IUCrJ *6*, 5–17.

Zivanov, J., Nakane, T., and Scheres, S.H.W. (2020). Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. IUCrJ *7*, 253–267.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Methods

Expression Constructs

Expression constructs were generated essentially as described in Xiong et al. 2020 (Xiong et al., 2020). Starting from the construct S-R (Xiong et al., 2020), introduction of cysteines to form x3 disulfide and introduction of D614G substitution were carried out using Q5 polymerase PCR with primers containing desired substitutions, followed by In-Fusion HD (Takara Bio) assembly.

Protein Production and Purification

Proteins were expressed in Expi293 cells and purified by metal exchange chromatography exactly as described in Xiong et al. 2020 (Xiong et al., 2020). Except where otherwise indicated, proteins were flash frozen in liquid nitrogen and stored at -80°C.

Negative staining EM

3 µl of proteins (~0.05 mg/ml) diluted with PBS buffer were deposited onto carbon-coated grids (EMS CF200-Cu) glow-discharged for 15 seconds at 25 mA in air. After 60s incubation, excess proteins were wicked by filter paper. Grids were washed once in buffer, and stained twice in Nano-W stain (Nanoprobes) with blotting in between. The grids were air dried on filter paper and imaged using a Tecnai T12 Spirit operated at 120 kV. Micrographs were taken in Digital Micrograph (Gatan).

Cryo-EM

Grid preparation and image collection were performed for S-R/x3 and S-R/x3/D614G spike proteins essentially as described in Xiong et al. 2020 (Xiong et al., 2020). C-Flat 2/2 3C grids (Protochips) were glow-discharged for 45 seconds at 25 mA. 3 µl of freshly purified protein at 0.6 mg/ml supplemented with 0.01% octyl-glucoside (OG) was applied to the grids, which were plunge-frozen in liquid ethane using a Vitrobot (Thermo Fisher Scientific).

An aliquot of S-R/x3 freshly purified spike at 1.0 mg/ml was stored at 4˚C for 40 days. 10 µl of the stored protein was subjected to plunge-freezing at 1.0 mg/ml following the same procedure as for the freshly purified S trimers. Another 10 µl of the stored protein was incubated with 1 µl of citrate acid (pH 4.8) overnight and then plunge-frozen.

Grids were stored in liquid nitrogen and loaded into a Titan Krios electron microscope (Thermo Fisher Scientific) operated at 300kV. Movies with 48 frames were collected with a Gatan K3 BioQuantum direct electron detector with the slit retracted. Three shots per hole were achieved with beam-image shift controlled in SerialEM-3.7.0 (Mastronarde, 2005). An accumulated dose of 50 electrons/Å^2 were acquired in counting mode at the magnification of 81,000 X, corresponding to a calibrated pixel size of 1.061 Å/pixel. Detailed data acquisition parameters are summarized in Extended Data Table 1.

CryoEM image processing

Real-time data processing was performed in RELION-3.1's Scheduler as described in (Xiong, 2020). Motion correction, contrast transfer function (CTF) estimation, template particle picking and initial 3D classification were carried out while micrographs were being collected. An EM structure of the SARS-CoV-2 S protein in open form was filtered to 20 Å resolution as

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

a 3D reference for template picking. Initial 3D classification with an open S model was accomplished at bin4 in batches of 500,000 particles to identify S protein trimers. Subsets of S trimers in the 3D classes which displayed clear secondary structures were pooled and subjected to one round of 2D classification cleaning. Subsequently, a second round of 3D classification was used to assess the ratio of closed and locked states (**Fig. S1**). Auto refinement, Bayesian polishing and CTF refinement were performed iteratively on classified closed and locked subsets, respectively (Zivanov et al., 2019, 2020). Following the final round of 3D auto-refinement, map resolutions were estimated at the 0.143 criterion of the phase-randomization-corrected Fourier shell correlation (FSC) curve calculated between two independently refined half-maps multiplied by a soft-edged solvent mask. Final reconstructions were sharpened and locally filtered in RELION *post-processing* (**Fig. S2**). The estimated B-factors of each map are listed in Extended Data Table 1.

The radiation damage caused fading of the disulphide bond between the two engineered cysteines in S-R/x3 and S-R/x3/D614G. In order to recover the bond density, EM maps of closed and locked S trimers were reconstructed from the first 4 frames in the movies of S-R/x3 and the first 5 frames of S-R/x3/D614G, respectively. Distinct densities of the disulphide bond were observed in the EM structures reconstructed from early exposed frames and are shown in **Fig. 2**.

Model building and refinement

For the closed conformation structures, the SARS-CoV-2 S protein ectodomain structure (PDBID: 6ZOX (Xiong et al., 2020)) was fitted into the EM density as a starting model. For the locked conformations, S structures from our previous study (PDBID: 6ZP2, 6ZOZ (Xiong et al., 2020)) were used as starting models. Model building and adjustment were performed manually in Coot-0.9 (Emsley and Cowtan, 2004). Steric clash and sidechain rotamer conformations were improved using the Namdinator web server (Kidmose et al., 2019). After further manual adjustment, the locked and closed structures were refined and validated in PHENIX-1.18.261 (Afonine et al., 2018) to good geometry. Refinement statistics are given in Extended Data Table 1.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

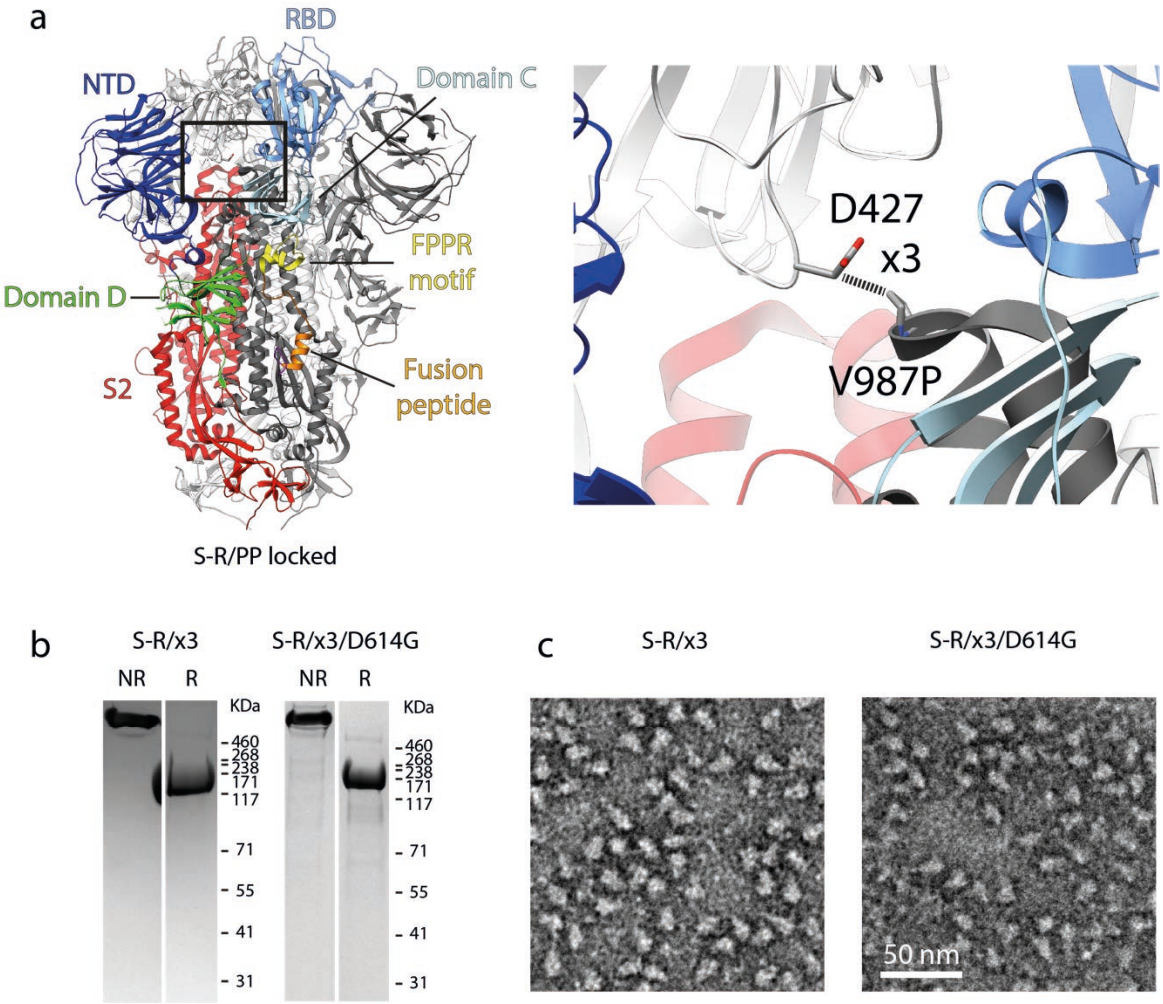

Figure 1. Design of the x3 disulfide bond to stabilise the SARS-CoV-2 spike in the "locked" state. a, Structure of the S-R/PP spike in the locked state. Structural domains are coloured and labelled. The box indicates the location of the zoomed-in view in the right panel. Residues 427 and 987 that were mutated to cysteine to form the x3 disulfide bond are indicated. **b,** Coomassie-stained SDS-PAGE gels to assess expression of the S-R/x3 and S-R/x3/D614G S proteins and confirm formation of a disulphide bond. NR or R indicates protein samples were prepared in non-reducing or reducing conditions. **c,** Negative stain EM images of the purified S-R/x3 and S-R/x3/D614G spikes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

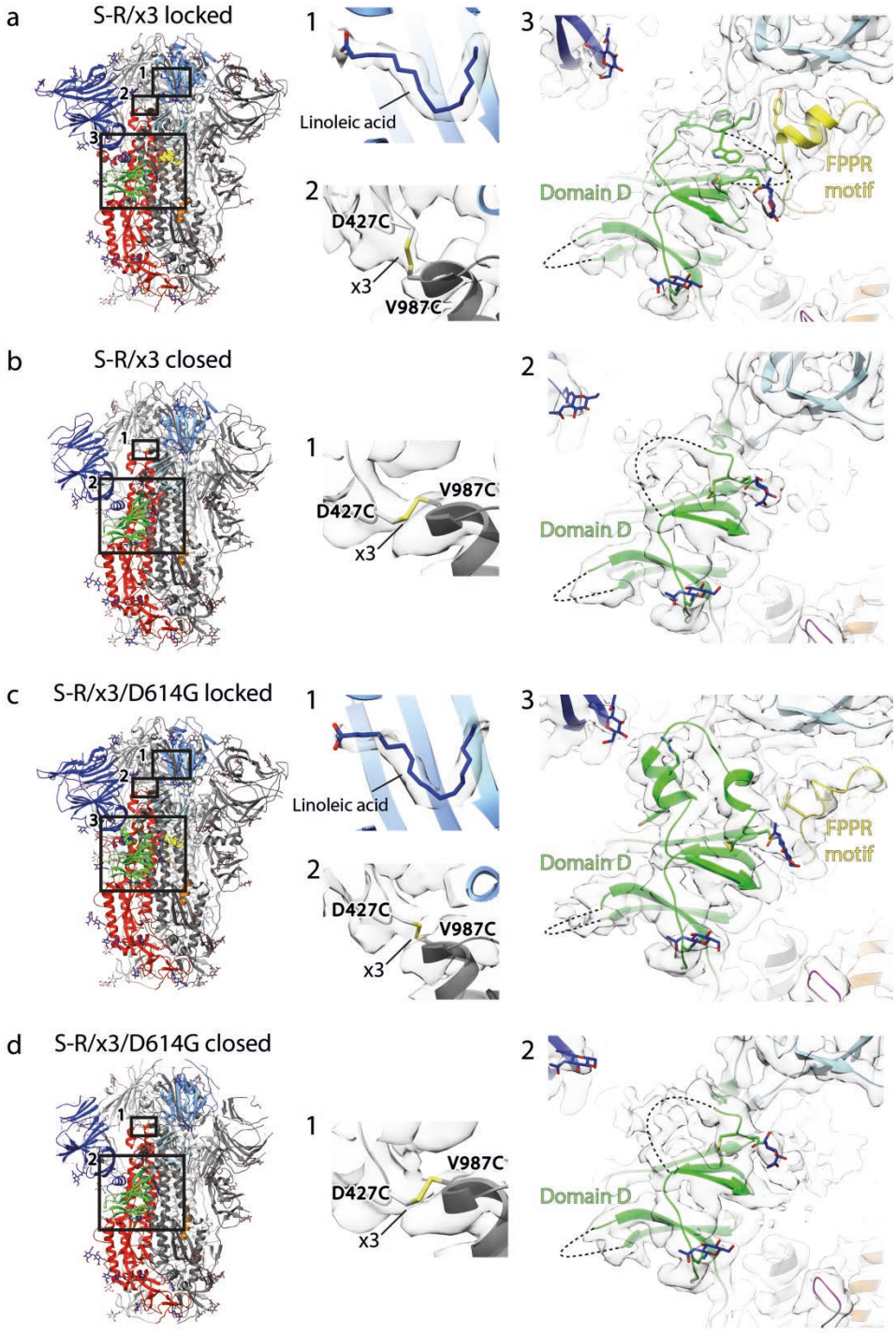

Figure 2. Structural features of spike proteins in different conformational states as determined by cryoEM. a-d, left panels show side views of indicated spike proteins and conformational states, the structural domains are coloured as in Figure 1. The numbered boxes indicate the zoomed-in views as shown in the right hand panels. The right hand panels show cryoEM densities of linoleic acid, the x3 disulfide bond, and structural regions around domain D. Disordered loops are indicated by dashed lines, glycans are shown as stick representations in blue.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

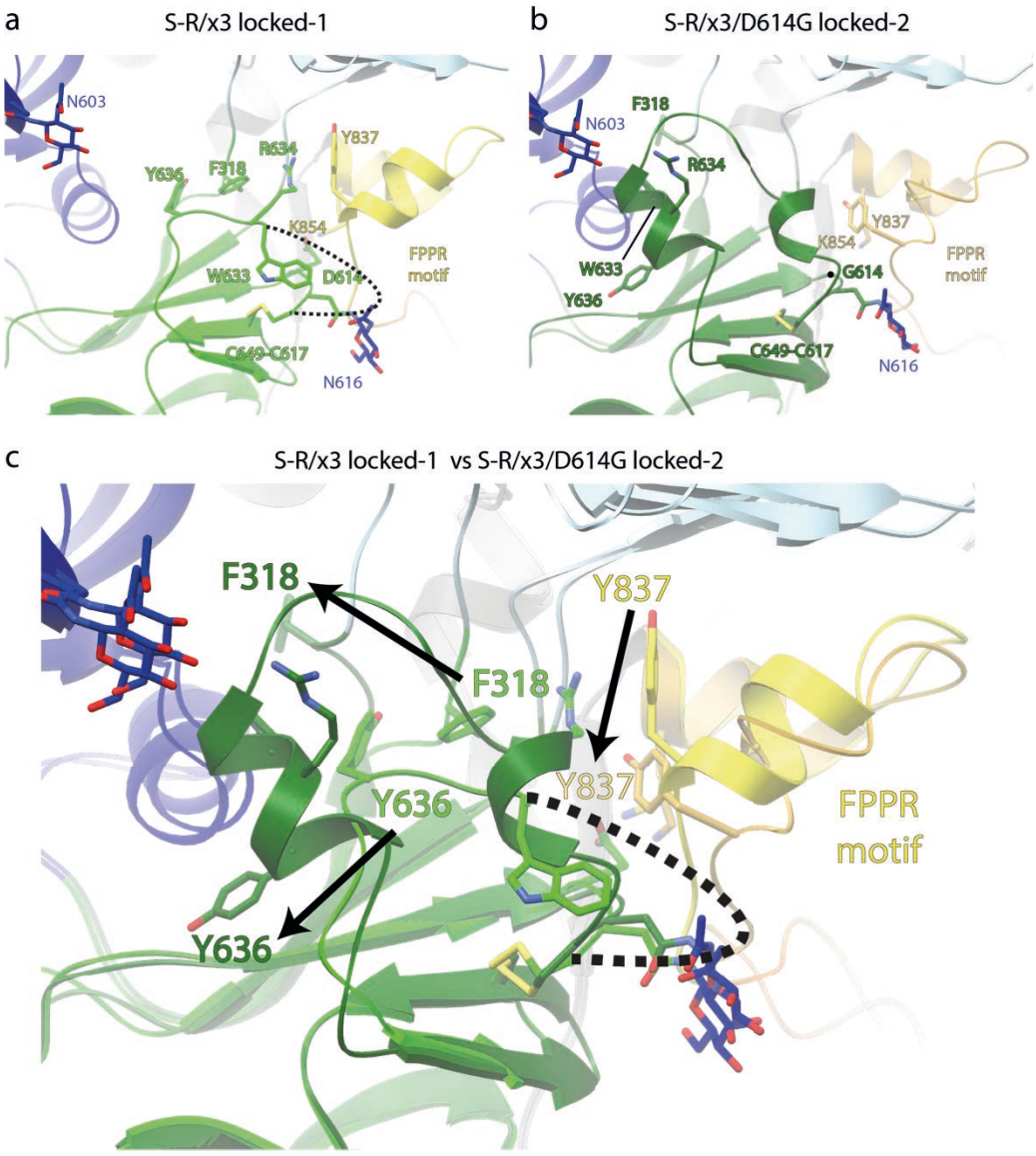

Figure 3. Structural changes in domain D between spikes in locked-1 and locked-2 conformational states. a, structure of domain D in the S-R/x3 locked-1 spike structure. Selected amino acid sidechains and disulfide bonds are marked. The disordered loop between W633 and C617 is represented by a dashed line. Y636, F318, R634, and Y837 interdigitate to stabilise the disordered loop. **b,** structure of domain D in the S-R/x3/D614G locked-2 spike structure. The loop region containing the selected residues has refolded to adopt a different structure. **c,** overlay of domain D structures in spikes adopting the locked-1 and locked-2 conformational states. The movements of residues F318, Y636, and Y837 between the two conformational states are indicated by arrows. Structural domains are colored as in as in Figure 1.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

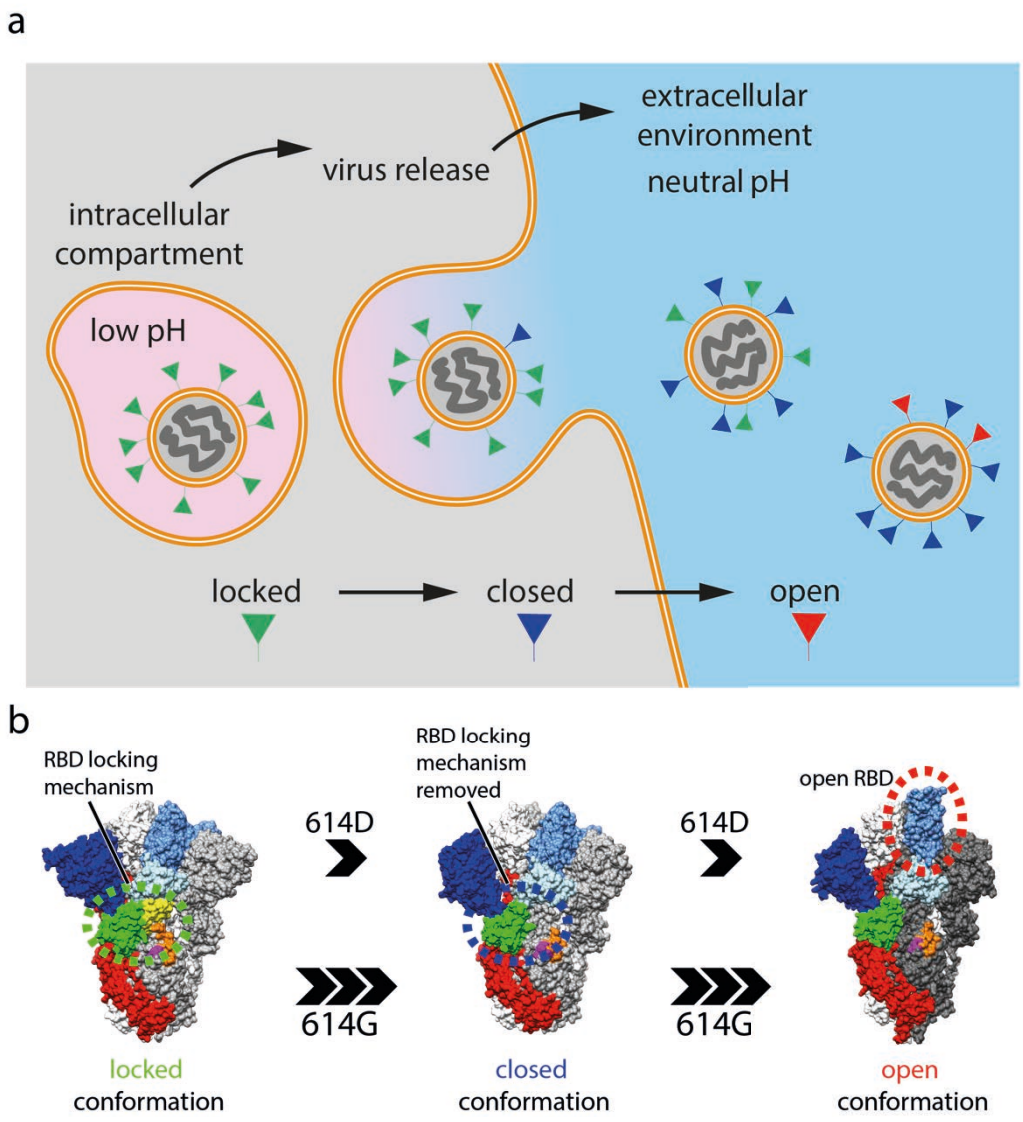

Figure 4. SARS-CoV-2 virus particle release and the accompanied structural transitions of surface spike protein. **a,** schematic diagram illustrating the release of nascent SARS-CoV-2 virus particles from the cell. Labelled boxes indicate the predicted conformational states of spike proteins as the virus travel through the acidic intracellular compartment before it was released into the neutral extracellular environment. **b,** the three prefusion conformational states observed for the SARS-CoV-2 spike protein and factors that influence the structural transitions between the conformational states.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

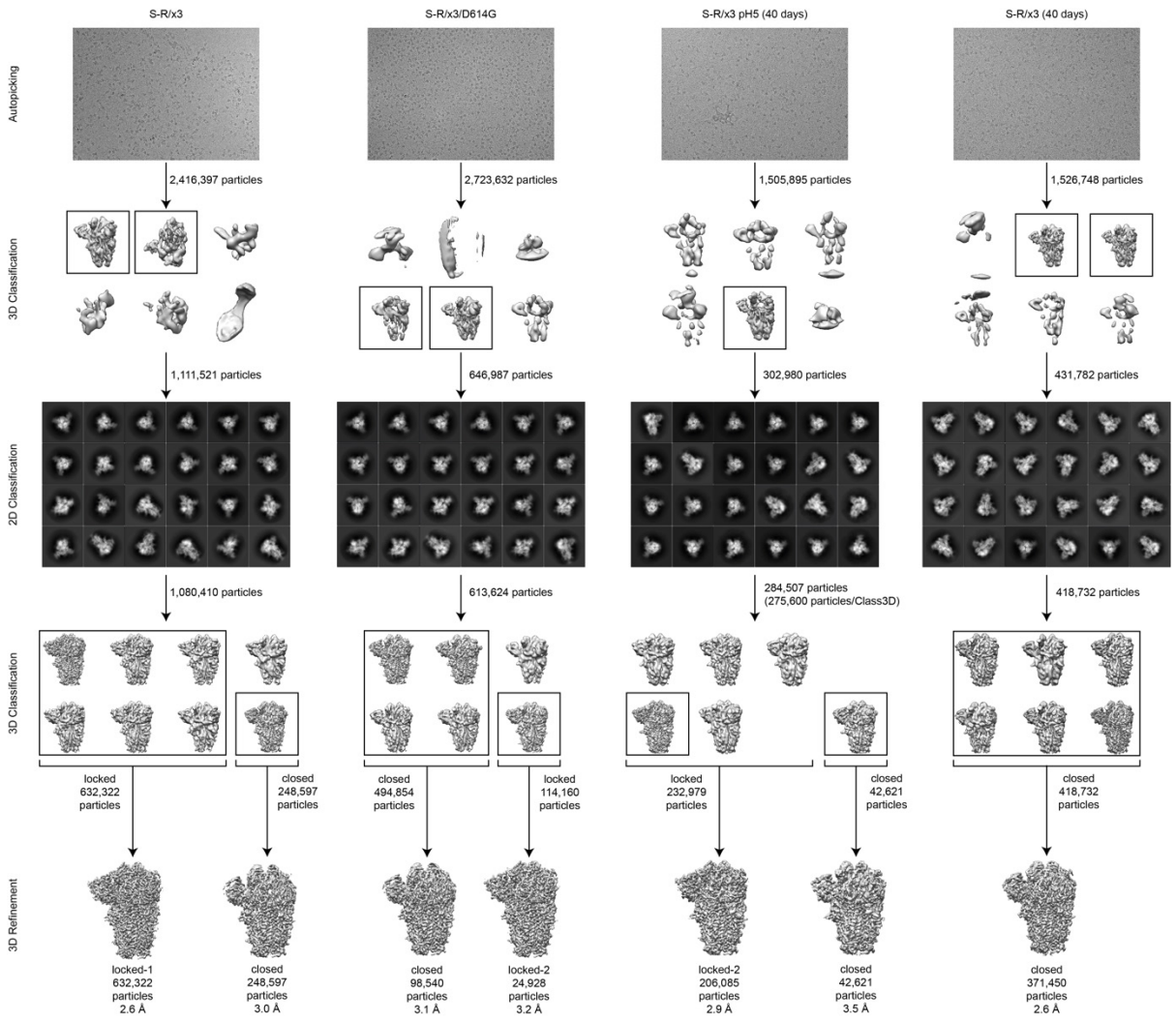

Figure S1. Pipeline used for picking and classification of cryo-EM data. Pipeline is illustrated for S-R/x3, S-R/x3/D614G, S-R/x3 after 40 days, and the latter after transfer to pH5.0 buffer. After automated picking, 3D and 2D classification steps were used to remove contaminating objects. 3D classification was then used to sort the data into locked and closed conformations.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

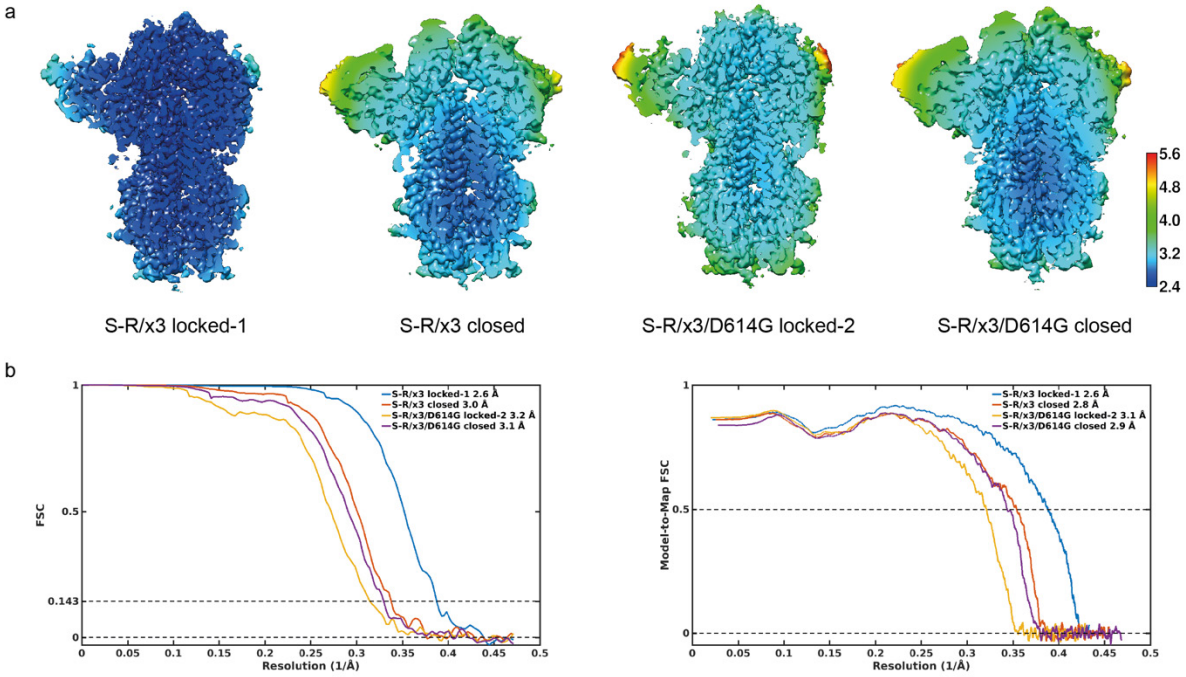

Figure S2. Resolution assessment of cryo-EM structures. a, Local resolution maps for all structures calculated using RELION. **b,** Global resolution assessment by Fourier shell correlation at the 0.143 criterion (left panel), and correlations of model vs map by Fourier shell correlation at the 0.5 criterion (right panel).

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

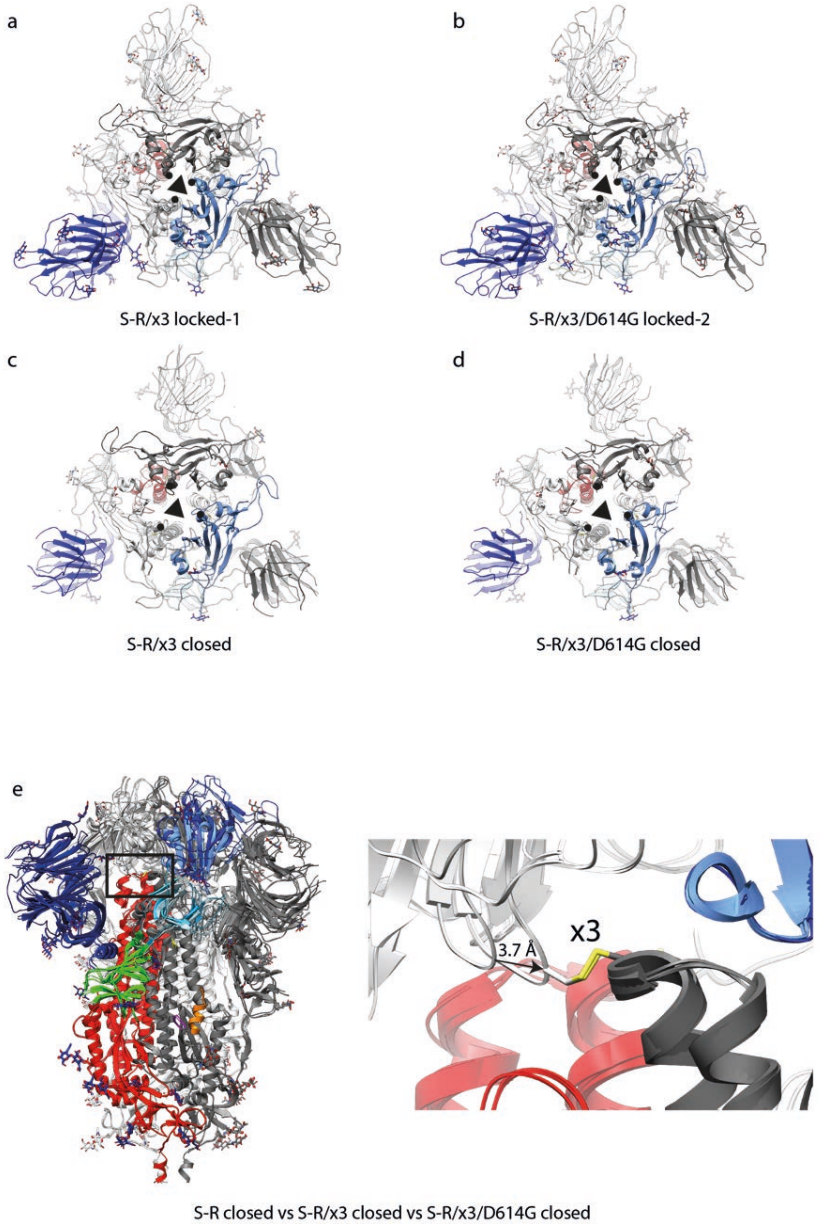

Figure S3. Structural features of S-R/x3 and S-R/x3/D614G S proteins. a-d, top views of S-R/x3 and S-R/x3/D614G spikes in locked-1, locked-2 and closed conformational states. The three-fold axes are indicated by triangles. The structural elements that are closest to the three-fold axes in the locked structures are indicated by dots (panel a and b). The same structural elements are indicated by dots in the closed conformations to highlight the movement of the receptor binding domains between the two conformational states (panel c and d). **e,** comparison of S-R, S-R/x3 and S-R/x3/D614G spikes in closed conformational states. Left panel, overlay of the three structures, the box indicates the zoomed-in view in the right hand panel. Right panel, a movement of 3.7 Å was observed between the closed S-R (without x3 disulphide) and the closed S-R/x3 and S-R/x3/D614G spikes (with x3 disulphide).

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

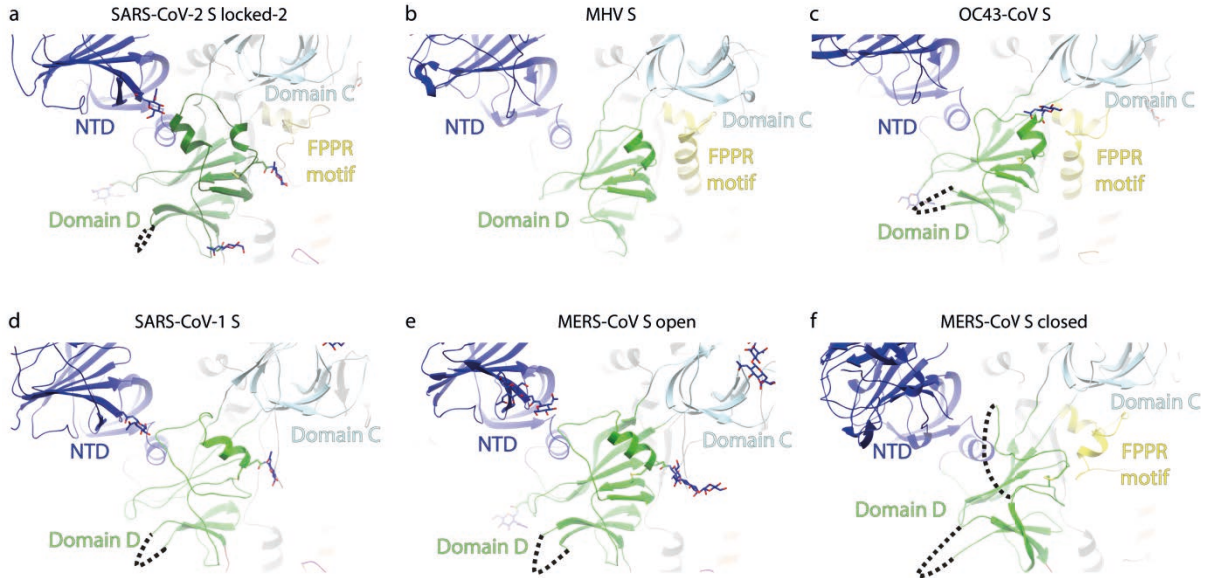

Figure S4. Comparison of domain D structures in S proteins of beta-coronaviruses. Structural elements are colored as in Figure 1. Disordered loops are indicated by dashed lines. **a,** SARS-CoV-2 S in locked-2 conformation; **b,** MHV S; **c,** OC43-CoV S; **d,** SARS-CoV-1 S; and **e,** MERS-CoV S in open conformation all feature helical structures in Domain D. There is a structural change in domain D between the open conformation of MERS-CoV S and **f,** closed MERS-CoV S.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

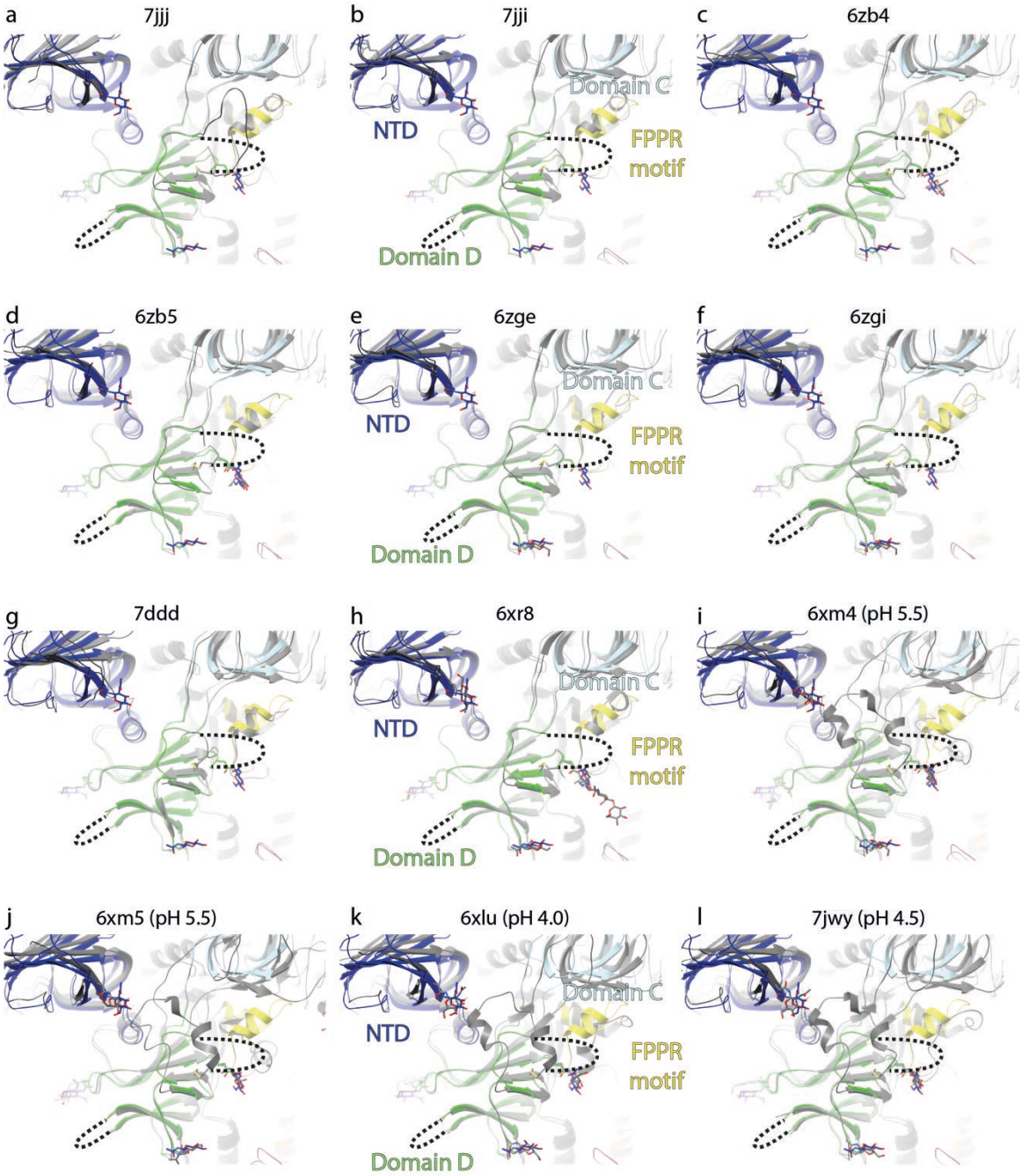

Figure S5. Comparison of domain D and surrounding structural features between S-R/x3 "locked-1" structure (coloured) and different deposited structures in which the 833-855 motif is ordered (gray). a-h, In most of these structures, the 833-855 motif adopts the same structure as in the locked-1 state. **i-l,** In the low-pH-induced structures (6xm4, 6xm5, 6xlu, 7jwy), domain D has refolded, the 833-855 motifs are formed but the structures vary.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

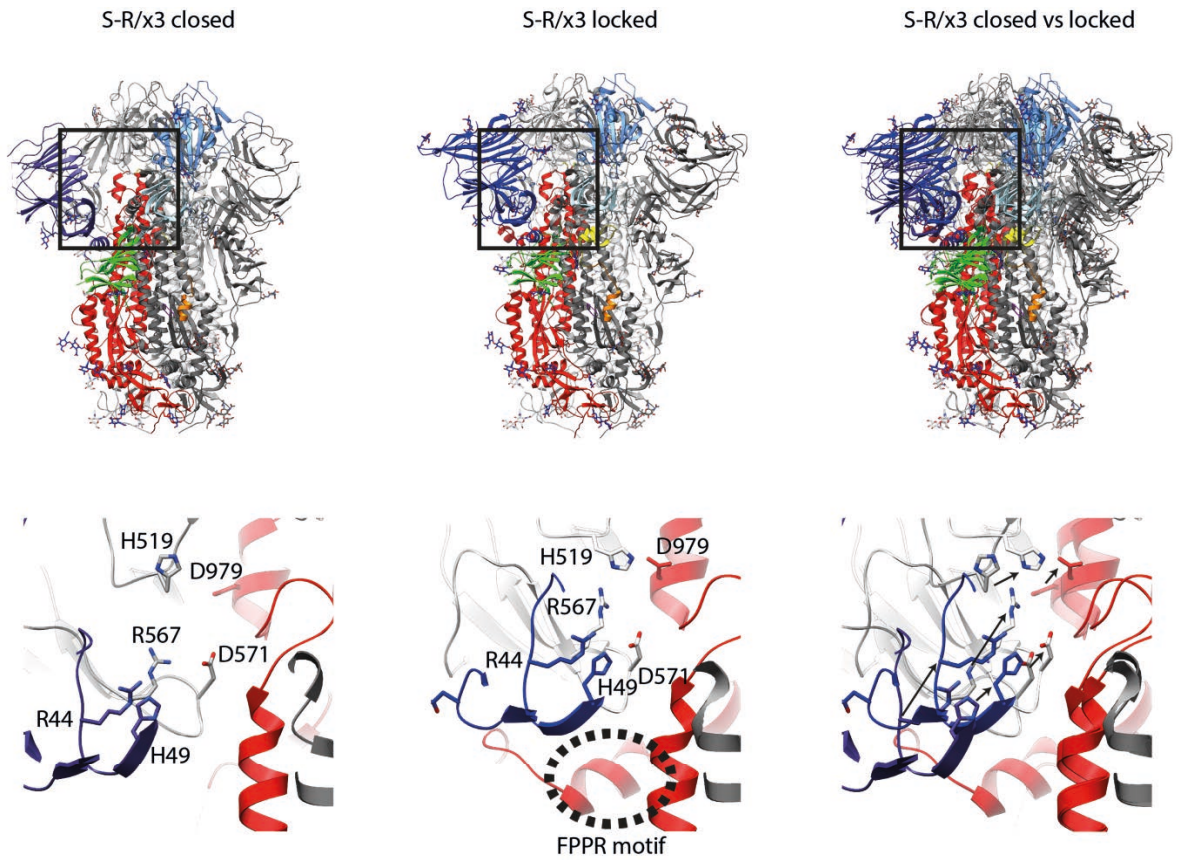

Figure S6. Speculative pH sensing mechanism within the SARS-CoV-2 spike. H49 and H519 undergo large movement between the “closed” and “locked” conformational states to engage in different interactions. These residues could serve as the pH sensor to induce conformational change in SARS-CoV-2 spike.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.10.434733; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

	S-R/x3 locked-1 (EMD-xxxx, PDB xxxx)	S-R/x3 closed (EMD-xxxx, PDB xxxx)	S-R/x3/D614G locked-2 (EMD-xxxx, PDB xxxx)	S-R/x3/D614G closed (EMD-xxxx, PDB xxxx)	S-R/x3 (pH 5) locked-2 (EMD-xxxx)	S-R/x3 (pH 5) closed (EMD-xxxx)
Data collection and processing						
Magnification	81,000		81,000		81,000	
Voltage (kV)	300		300		300	
Electron exposure (e^-/Å^2)	50		50		50	
Defocus range (μm)	1.0-2.4		1.0-2.4		1.0-2.4	
Pixel size (Å)	1.061		1.061		1.061	
Movies (no.)	6,642		3,036		1,698	
Symmetry imposed	*C3*		*C3*		*C3*	
Initial particle images (no.)	2,416,397		2,723,632		1,505,895	
Final particle images (no.)	632,322	248,597	24,928	98,540	232,979	42,621
Map resolution (Å)	2.6	3.0	3.2	3.1	2.9	3.5
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.44-4.19	2.73-5.92	2.97-8.82	2.78-6.98	2.66-5.86	3.12-8.63
Refinement						
Initial model used	PDB 6ZP2	PDB 6ZOX	PDB 6ZOZ	PDB 6ZOX		
Model resolution (Å)	2.6	2.8	3.1	2.9		
FSC threshold	0.5	0.5	0.5	0.5		
Map sharpening *B* factor (Å^2)	-38.8	-25.3	-30.0	-30.0	-21.8	-50
Model composition						
Non-hydrogen atoms	26724	24150	26940	22611		
Protein residues	3288	3009	3330	2811		
Ligands	72	45	66	42		
B factors (Å^2)						
Protein	70.95	103.16	83.32	85.36		
Ligand	94.40	112.38	102.98	100.14		
R.m.s. deviations						
Bond lengths (Å)	0.008	0.005	0.006	0.010		
Bond angles (°)	0.816	0.898	0.832	0.762		
Validation						
MolProbity score	1.30	1.52	1.66	1.60		
Clashscore	2.96	5.19	5.06	6.15		
Poor rotamers (%)	0.63	0.69	0.31	0.53		
Ramachandran plot						
Favored (%)	96.65	96.95	94.16	95.51		
Allowed (%)	3.35	3.05	5.75	4.49		
Disallowed (%)	0.00	0.00	0.09	0.00		

Extended Data Table 1. Cryo-EM data collection, refinement and validation statistics