Engineered disulfide reveals structural dynamics of locked SARS-CoV-2 spike

Kun Qu[1,#], Qiuluan Chen[2,#], Katarzyna A. Ciazynska[1], Banghui Liu[2], Xixi Zhang[2], Jingjing Wang[2], Yujie He[2], Jiali Guan[2], Jun He[2,4], Tian Liu[5], Xiaofei Zhang[2,4], Andrew P. Carter[1], Xiaoli Xiong[1,2,3,4,*], John A. G. Briggs[1,6,*]

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.

3. The State Key Laboratory of Respiratory Disease (SKLRD), Guangzhou Institute of Biomedicine and Health, the Chinese Academy of Sciences, Guangzhou, China.

4. CAS Key Laboratory of Regenerative Biology, Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Center for Cell Lineage and Development, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, China.

5. Center for Proteomics and Metabolomics, Bioland Laboratory, Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Guangzhou, China.

6. Max Planck Institute of Biochemistry, 82152 Martinsried, Germany.

# equal contribution

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

Abstract

The spike (S) protein of SARS-CoV-2 has been observed in three distinct pre-fusion conformations: locked, closed and open. Of these, the function of the locked conformation 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 confirming that the x3 disulfide bond has the ability to stabilize the otherwise transient locked conformations. Structural analyses reveal that wild-type SARS-CoV-2 spike can adopt two distinct locked-1 and locked-2 conformations. For the D614G spike, based on which all variants of concern were evolved, only the locked-2 conformation was observed. Analysis of the structures suggests that rigidified domain D in the locked conformations interacts with the hinge to domain C and thereby restrains RBD movement. Structural change in domain D correlates with spike conformational change. We propose that the locked-1 and locked-2 conformations of S are present in the acidic high-lipid cellular compartments during virus assembly and egress. In this model, release of the virion into the neutral pH extracellular space would favour transition to the closed or open conformations. The dynamics of this transition can be altered by mutations that modulate domain D structure, as is the case for the D614G mutation, leading to changes in viral fitness. 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.

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 take part 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 (cryo-EM): 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). 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 can sterically hinder 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. Closed SARS-CoV-2 spike usually coexist with spikes exhibiting an open conformation (Walls et al., 2020; Wrapp et al., 2020). In the open

101 conformation, one or more RBDs are raised up to expose the receptor binding loops. ACE2
102 or antibody binding to these loops appears to maintain the spike in an open conformation
103 (Benton et al., 2020; Xu et al., 2021). It is believed that opening of the cleaved spike
104 ultimately leads to structural transition into the postfusion conformation. The virus uses the
105 dramatic structural refolding from the prefusion conformation to the lowest-energy
106 postfusion conformation to drive membrane fusion (Bosch et al., 2003).
107
108 Among the observed conformational states of SARS-CoV-2 spike, the locked conformation is
109 unique in that it has not been observed for the highly related SARS-CoV-1 spike. Other
110 coronavirus spikes do have locked-like conformations featuring FPPR motifs that preventing
111 RBD opening, but for many of them, stochastic formation of an open conformation has not
112 been observed (Yuan et al., 2017). The early cryo-EM studies of purified SARS-CoV-2 spike
113 ectodomain did not identify the locked conformation, and we previously found that only a
114 small fraction of trimeric ectodomains adopt the locked conformation (Xiong et al., 2020),
115 suggesting this conformation is transient. However, the locked conformation continued to
116 be identified in other cryo-EM studies: spikes in locked conformation have been identified in
117 purified full-length spike proteins including the transmembrane region (Bangaru et al., 2020;
118 Cai et al., 2020); insect-cell-expressed spike ectodomains can also adopt the locked
119 conformation, perhaps due to stabilizing lipid in the insect cell media (Toelzer et al., 2020).
120 However, curiously, the locked conformation was not observed in the high-resolution
121 structures of spike proteins on virions, where only closed and open conformation spikes
122 were observed (Ke et al., 2020).
123
124 Considering the above observations, the role of the locked conformation in the SARS-CoV-2
125 life cycle is unclear. Here, based on our previous locked spike structure (Xiong et al., 2020),
126 we engineered a disulfide bond which stabilises the SARS-CoV-2 spike in the locked
127 conformation. This construct provides a tool to characterize the structure and dynamics of
128 the locked form of S and to assess how mutations and other environmental factors could
129 affect its structure and function. Our data reveal multiple distinct locked conformations and
130 suggest that structural dynamics of domain D modulates the conformational state of the
131 SARS-CoV-2 spike. Newly obtained information regarding the "locked" spike allowed us to
132 consider its functional role in the context of SARS-CoV-2 life cycle.
133
134 **Results and Discussion**
135
136 *Design of a spike protein ectodomain stabilized in the locked conformation*
137
138 We did not previously observe a locked conformation for S-R (S with a deletion at the furin
139 cleavage site leaving only a single arginine residue), while for S-R/PP (S-R additionally
140 containing the widely-used, stabilizing double-proline mutation) the locked conformation
141 was rare, <10% of S trimers when imaged by cryo-EM (Xiong et al., 2020). The scarcity of
142 the locked conformation for the expressed ectodomain makes it difficult to study. To
143 overcome this challenge, based on the locked SARS-CoV-2 S-R/PP structure (Xiong et al.,
144 2020), we engineered cysteines residues replacing positions D427 and V987 to generate a
145 new disulfide-stabilised S protein S-R/x3. We predicted that S-R/x3 should be predominantly
146 in the locked conformation (**Fig. 1a**). Under non-reducing conditions, S-R/x3 ran exclusively
147 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 between the 2 engineered cysteines during protein expression. 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**).

Biochemical properties of the x3 spike protein

We tested the sensitivity of the x3 disulfide bond to reduction by dithiothreitol (DTT), and compared this to the sensitivity of the previous described x2 disulfide bond, which stabilises the spike in the closed state (Xiong et al., 2020). After 5 min of incubation with 2.5 mM DTT, the x3 disulfide in both S-R/x3 and S-R/x3/D614G spikes was substantially reduced, and 5 min of incubation with 20 mM DTT was sufficient for almost complete reduction (**Fig. 2a**). In contrast, a 60 min incubation with 20 mM DTT was needed to achieve near-complete reduction of the x2 disulfide bond (**Fig. 2a and Fig. S1a**). The x3 and x2 disulfides both include C987 in S2, therefore, different local chemical environments of the other engineered cysteines at position 427 and 413 respectively, are likely to be responsible for the difference in susceptibility to reduction by DTT.

We tested the receptor binding properties of S-R/x3 and S-R/x3/D614G in the absence and presence of DTT. In the absence of DTT, S-R/x3 and S-R/x3/D614G bind ACE2-Fc with weak responses (<0.4) and low affinities (>100 nM) (**Fig. 2b**), while under similar conditions, S-R and S-R/PP bind ACE2-Fc with maximum responses between 0.9-1.0 and affinities of ~1 nM (Carnell et al., 2021). In the presence of 20 mM DTT, S-R/x3 and S-R/x3/D614G bound ACE2-Fc with maximum responses of 0.8-1.0 and affinities in the low nano-molar range (**Fig. 2b**). In contrast, likely due to the inefficiency of x2 disulfide bond reduction, S-R/x2 and S-R/x2/D614G only moderately increased responses and affinities under reducing conditions (**Fig. 2b**). These observations are consistent with our prediction that the x2 and x3 disulphide bonds prevent RBD opening and thereby prevent strong receptor binding, and that reduction of the x3 disulphide by DTT permits RBD opening and receptor binding.

Entirely in line with the receptor binding data, we found that the antibody CR3022, which binds a cryptic epitope exposed only in the open RBD, is only able to bind S-R/x3 and S-R/x3/D614G spikes in the presence of DTT, and does not bind S-R/x2 and S-R/x2/D614G spikes (**Fig. S1b**).

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 the majority of the particles (76%) are in the locked conformation, while 23% of the particles are in the closed conformation (**Fig. S2**). Consensus refinement imposing C3 symmetry of all the locked spike particles obtained a 2.6 Å resolution map (**Fig. S2**). The x3 disulfide bond appears to be radiation sensitive, and the x3 disulfide density is only visible in the cryo-EM density derived from early exposure frames

(**Fig. S3**). Further 3D classification after symmetry expansion revealed that locked S-R/x3 spike protomers adopt two distinct conformations that differ primarily at domain D (**Fig. S2 and Fig. 3a and Fig. 3b**), we will refer to these conformations as “locked-1” and “locked-2”. Classification revealed that locked S-R/x3 S trimers were formed by all four possible combinations of locked-1 and locked-2 protomers (the combinations 111, 112, 122 and 222 **Fig. S2**).

In the locked-1 conformation, residues 617-641 in domain D form a large loop, of which residues 632-641 can be clearly identified in the cryo-EM density while residues 619-631 are disordered and not resolved (**Fig. 3a**). In the locked-2 conformation, residues 617-641 in the locked-2 protomer are ordered and can be clearly identifiable in cryo-EM density (**Fig. 3b**). Both locked-1 and locked-2 conformations are characterised by: structural rigidity (showing minimal local dynamics in NTD and RBD (**Fig. S4**)); RBDs are tightly clustered around the three-fold symmetry axis (**Fig. S5a and Fig. S5b**); linoleic acid is bound into the described pocket in the RBD (**Fig. 3a and Fig. 3b**); and residues 833-855 are folded into a helix-loop-helix fusion peptide proximal region (FPPR) motif (**Fig. 3a and Fig. 3b**). These resolved features are consistent with our previously observations of the locked structure (Xiong et al., 2020), further confirming that insertion of the x3 disulfide bond successfully stabilizes the locked conformation during protein expression and purification. We observed density consistent with the presence of the pigment biliverdin in the reported NTD binding pocket (Rosa et al., 2021) in all our spike structures (**Fig. 3**), consistent with the greenish color of the purified protein. We confirmed the presence of biliverdin by mass spectrometry (**Fig. S6**).

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

Introduction of the D614G substitution into SARS-CoV-2 S spike ectodomain constructs with the double proline stabilising modification has been reported to lead to a higher fraction of purified S in the open conformation (Benton et al., 2021; Yurkovetskiy et al., 2020). Consistent with this, we found that 63% of purified S-R/D614G spikes adopted an open conformation (the remainder are in a closed conformation) (**Fig. S2**), compared to ~18% of S-R spikes in the open conformation (Xiong et al., 2020). We previously speculated that the D614G substitution alters the structure of the locked conformation, modulating spike structural dynamics (Xiong et al., 2020). Therefore, we determined the structure of the S-R/x3/D614G spike. The proportion of locked conformations is greatly reduced compared to S-R/x3, constituting only ~19% of total particles (**Fig. S2**). Classification did not identify two different locked conformations of the S-R/x3/D614G – all locked spikes were in the locked-2 conformation. As in the S-R/x3 locked-2 structure, residues 617-641 are ordered, and linoleic acid is bound (**Fig. 3e**) confirming that the D614G spike retains the ability to bind linoleic acid, a molecule that has been suggested to be important for regulating conformational state of the spike (Toelzer et al., 2020). Despite the presence of the x3 disulfide bond, 81% of the particles are in a closed conformation. The closed conformation accommodates the disulfide bond by motion of the RBD towards S2 (**Fig. S5h**).

Structural transition of locked spike to closed conformation and the effect of low pH

The S-R/x2 spike can be stored at 4°C for at least 30 days without significant loss of particles by negative stain EM (Xiong et al., 2020). We observed that the S-R/x3 spike can also be stored at 4°C for 40 days retaining well-formed trimers in negative stain EM. We performed cryo-EM imaging of the remaining trimers, finding that they had transitioned into the "closed" conformation (**Fig. S2**), and that linoleic acid was no longer bound in the RBD. This suggests that both the locked-1 and locked-2 conformations are metastable at pH 7.4, and spontaneously transition into the closed conformation. It is consistent with the general observation that most published ectodomain structures (not stabilised by the x3 crosslink) adopt the closed/open conformations rather than the locked conformation. We adjusted the buffer pH of the now "closed" spike to pH 5.0 and incubated overnight before determining its structure by cryo-EM. We found that low pH treatment had converted most of the closed spike to a conformation similar to the locked-2 conformation (**Fig. 3d, Fig. S2**), however, in the acid reverted locked-2 structure the linoleic acid binding pocket is unoccupied (**Fig. 3d**).

The structure of Domain D differs in locked-1 and locked-2 conformations

The locked-1 and locked-2 conformations differ primarily at residues 617-641 within domain D, and this structural perturbation is associated with differences in the FPPR motif and in residues around 318-319 which are located in the junction region connecting domain C and D (**Fig. 3a, b; 4a, b**). In the S-R/x3 locked-1 structure, the long sidechains of R634 and Y636 interdigitate with the sidechains of F318 and Y837 through cation-π and π-π stacking interactions (dashed red lines in **Fig. 4a**). In this interaction, R634 in domain D is sandwiched between F318 of the domain C/D junction region, and Y837 of the FPPR motif, bridging these three structural features. These interactions appear to be further stabilised by hydrophobic contacts between the aromatic sidechains of F318 and Y636 and the domain D hydrophobic core. This arrangement positions loop residue R634 approximately 9 Å above D614 (**Fig. 4a)**. It appears likely that this zip-locking interaction not only maintains residues 617-641 in a loop structure but also restrains motion of domain C relative to domain D.

In the 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. 4c**). We suggest that substitution of negatively charged D614 to neutral G alters the local electrostatic interactions, preventing R634 binding between F318 and Y837, and triggering the 617-641 loop refolding. This would provide an explanation for the absence of the locked-1 conformation in S-R/x3/D614G.

In the locked-2 structure, 617-641 loop refolds into two short alpha helices, and the sidechains of F318 and R319 at the domain C/D junction reorient (**Fig. 4b)**. These structural changes allow formation of new electrostatic cation-π interactions between R319 and aromatic residues W633 and F592 (dashed red lines in **Fig. 4b**) and formation of hydrophobic interactions between residues in the 617-641 loop and the hydrophobic core of domain D formed by the beta-sheet structure (**Fig. 4b**). In the acid-reverted locked-2 structure of S-R/x3, F592 is no longer within cation-π interaction distance of R319 and instead forms cation-π interactions with H625 which is positively charged under low pH (**Fig. 4d**).

Comparison to other 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 cryo-EM studies have captured S trimers in locked conformations, including in purified full-length spike protein with or without the PP mutation (Bangaru et al., 2020; Cai et al., 2020) (**Fig. S8a, b, h**); soluble S-GSAS/PP trimers (where the furin cleavage site is replaced with a GSAS sequence) purified from insect cells (Toelzer et al., 2020) (**Fig. S8c, d**); soluble S-GSAS/PP trimers purified from mammalian cells (Wrobel et al., 2020; Xu et al., 2021) (**Fig. S8e, f, g**), and our previous study of soluble S-R/PP and S-R/PP/x1 (Xiong et al., 2020). In all of these cases the construct used had a D at position 614 and the locked-1 conformation was observed (**Fig. S8a-h**). Among these studies, the percentage of particles adopting the locked conformation and the dynamics of the purified spike vary. The reason why locked conformations are observed under some conditions but not others is not completely clear. Detergent in the purification procedure for the full-length spike (Bangaru et al., 2020; Cai et al., 2020) and high detergent concentration during the EM grid preparation (Wrobel et al., 2020) could prevent certain species of spike from being imaged. Other factors such as 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); and differences in reagents as concluded by Xu et al (Xu et al., 2021), may have all played a part. However, we speculate that the primary reason that different studies captured this protein in varied conformations is the inherent instability and dynamic nature of certain spike conformations combined with differences in sample preparation methods during purification and imaging. This conclusion justifies the need for development of methods to stabilize distinct conformations of SARS-CoV-2 spike.

The structure of the full-length 614G spike protein (PDB:7krq) (Zhang et al., 2021) is in an intermediate conformation between locked-2 and closed conformations (**Fig. S7b, Fig. S8i,j**). The FPPR motif in this structure is ordered, but linoleic acid is missing in the RBD and the 617-641 loop only partially contacts the domain D hydrophobic core. We proposed that the capture of this intermediate form reflects rapid transition of the D614G spike away from the locked-2 conformation, consistent with our observations that the majority of S-R/D614G spike is open (**Fig. S2**) and that even in the presence of x3, only a minority of S-R/x3/D614G spikes adopt the locked-2 conformation.

Zhou and colleagues investigated the effect of low pH on ACE2 and antibody binding to SARS-CoV-2 spike protein (Zhou et al., 2020). The best-resolved protomer in their all RBD "down" structure at pH 5.5 (6bm5) is similar to our acid-reverted S-R/x3 structure: the 617-641 loop is ordered to contact the domain D hydrophobic core and a cation- π interaction is formed between R319 and W633. Unlike in our structure, the FPPR motif is still showing some dynamics and is not retracted back to be below domain C (**Fig S8l**), this affects the positioning of F592 so that it is not interacting with H625. Therefore, 6bm5 is likely to have captured an intermediate state in the low pH promoted structural transition to locked-2 conformation. In line with our conclusion, this observation demonstrates that low pH is not sufficient to revert the 614D spike in the closed conformation back to the locked-1 conformation.

335
336 *The functional role of the locked-1 and locked-2 structures.*
337
338 Despite their differences, locked-1 and locked-2 conformations have in common that the
339 617-641 loop interacts with the domain D hydrophobic core and with residues at the
340 domain C/D junction such as F318, R319 and F592 (**Fig.4a-d**). In contrast, in the closed and
341 open spike structures (independent of whether 614 is D or G) these interactions are lost,
342 and the 617-641 loop density is disordered (**Fig. 3c, e; Fig. 4e, f**). Superposition of locked
343 and closed structures indicates that the loss of interactions at the domain C/D junction
344 leaves the junction unrestrained, allowing domain C and the RBD to move relative to
345 domain D and the NTD (**Fig. 4g, h; Fig. S7b**) and allowing the spike to open. Transitions from
346 the locked towards the closed and open positions of the RBD therefore correlates with the
347 structural changes in domain D.
348
349 Low pH is able to revert the closed spike towards the locked-2 conformation by restoring
350 interactions in domain D and the domain C/D junction (**Fig. 4d**) but cannot revert the S-R/x3
351 closed spike to the locked-1 conformation as the complex interdigitation involving 617-641
352 loop residues (**Fig. 4a**) cannot be reformed. The absence of the locked-1 conformation in S-
353 R/x3/D614G suggests to us that in the absence of the interaction between D614 and
354 R634/K858 the interdigitation interaction was not able to form during the 614G spike
355 synthesis.
356
357 Based on the coexistence of S-R/x3 locked-1, locked-2 and closed conformations in freshly
358 purified protein, and the spontaneous transition of S-R/x3 locked conformations to closed
359 conformation in our storage experiment (**Fig. S2**), we speculate that the S-R/x3 spike protein
360 is synthesized in the locked conformations and transitions to closed conformation due to
361 instability of locked conformations at neutral pH. Similarly, the observations that the locked-
362 1 conformation is more prevalent than the locked-2 conformation in S-R/x3, and that the
363 majority of purified S-R/x3/D614G spikes (which can only adopt the locked-2 conformation)
364 have transitioned away from the locked conformation (**Fig. S2**), suggests that the rigidified
365 domain D in the locked-2 spike is less stable than that the interdigitated arrangement in the
366 locked-1 spike and therefore more prone to transition into closed state under neutral pH.
367 Transition of domain D away from the locked conformations removes the structural
368 restraint on motion of domain C and the RBD. Structural transitions between locked, closed
369 and open conformations are complex, and although we suggest that changes in domain D
370 are a dominant factor in regulating transition between locked and closed conformations,
371 other factors, including linoleic acid binding, are also likely to play a role.
372
373 The absence of the stable, interdigitated locked-1 conformation in constructs containing the
374 D614G substitution is likely due to the change of electrostatic properties caused by the
375 D614G substitution. Instead D614G-containing spikes adopt the less stable locked-2
376 conformation and are more prone to transition away from the locked state towards closed
377 and open conformations (**Fig. 5a**). This phenomenon explains why studies of 614G SARS-
378 CoV-2 spikes without stabilising mutations failed to capture the locked conformations, and it
379 is consistent with observations that the 614G spike tend to adopt an open conformation
380 (Gobeil et al., 2021; Yurkovetskiy et al., 2020).
381

382 Beta-coronaviruses have recently been shown to assemble in the low pH, high lipid
383 environment of the ERGIC, and egress through the acidic environment of endosome (Ghosh
384 et al., 2020) (**Fig 5b**). Using x3 stabilised 614D and 614G spikes, we found both nascent
385 locked-1 and locked-2 conformations bind lipid, and low pH is able to revert the 614D spike
386 towards the locked-2 conformation. Given that both low pH and lipid binding favour locked
387 conformations, we speculate that during virus assembly, both the 614D and 614G spikes are
388 in locked conformations, and that this provides a mechanism to prevent premature
389 transition into open or post-fusion conformations during virus assembly. Once viruses are
390 released into the neutral pH environment outside of the cell, over time the spike transitions
391 to the closed or open conformations (**Fig 5b**). The 614D spike may transition from locked-1
392 to locked-2 to closed over time. The 614G spike can only adopt the locked-2 conformation,
393 exhibits a different structural dynamics and a higher tendency to open. We suggest that the
394 higher tendency to open is a result of a shifted locked to closed equilibrium (**Fig. 5a**). The
395 resulting change in virus life cycle is not fully understood and could potentially affect
396 receptor binding, cell entry and antigen presentation. Nevertheless, the D614G change
397 provided a transmission advantage leading to global dominance of the mutant virus based
398 on which all current variants of concern evolved (Korber et al., 2020). Further study of the
399 structural transitions between locked-1, locked-2 and closed conformations of the spike is
400 required, and can be facilitated by the x3 construct described here.
401
402 **Acknowledgements**
403 We thank the staff of the MRC-LMB EM Facility their support; Jake Grimmett and Toby
404 Darling for scientific computing infrastructure; and Patricia Edwards for supporting cell
405 culture. This study was supported by funding from the European Research Council (ERC)
406 under the European Union's Horizon 2020 research and innovation programme (ERC-CoG-
407 648432 MEMBRANEFUSION to JAGB), the Medical Research Council as part of United
408 Kingdom Research and Innovation (MC_UP_A025_1011 to APC; MC_UP_1201/16 to JAGB),
409 the Max-Planck Society (to JAGB), and the Natural Science Fund of Guangdong Province
410 (2021A1515011289 to XX). XX acknowledges start-up grants from the Chinese Academy of
411 Sciences and Bioland Laboratory (GRMH-GL).
412
413 **Author Contributions**
414 XX and JAGB conceived the study; KQ, XX and JAGB planned the study. XX, APC and JAGB
415 designed constructs which were made by APC. Protein purification was performed by QC,
416 KAC, BL, XZ, JW, JG, XX. QC planned and performed DTT sensitivity tests, BLI experiments
417 and analyzed the data with assistance from XX. TL and XZ performed mass spectrometry.
418 QC, BL, XZ, JW, YH, JG, JH, and XX performed negative stain EM and/or checked protein
419 quality by negative stain EM. XX prepared cryo-EM grids. KQ performed cryo-EM and
420 processed cryo-EM data. KQ and XX built molecular models which KQ, XX, and JAGB
421 interpreted. XX and JAGB managed the project. KQ, QC, XX and JAGB made figures. KQ, XX
422 and JAGB wrote the paper with input from all authors.
423
424 **References**
425
426 Afonine, P. V., Poon, B.K., Read, R.J., Sobolev, O. V., Terwilliger, T.C., Urzhumtsev, A., and
427 Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta
428 Crystallogr. Sect. D Struct. Biol. *74*, 531–544.

429 Bangaru, S., Ozorowski, G., Turner, H.L., Antanasijevic, A., Huang, D., Wang, X., Torres, J.L.,
430 Diedrich, J.K., Tian, J.-H., Portnoff, A.D., et al. (2020). Structural analysis of full-length SARS-
431 CoV-2 spike protein from an advanced vaccine candidate. Science *370*, 1089–1094.
432 Barnes, C.O., Jette, C.A., Abernathy, M.E., Dam, K.-M.A., Esswein, S.R., Gristick, H.B.,
433 Malyutin, A.G., Sharaf, N.G., Huey-Tubman, K.E., Lee, Y.E., et al. (2020). SARS-CoV-2
434 neutralizing antibody structures inform therapeutic strategies. Nature *588*, 682–687.
435 Benton, D.J., Wrobel, A.G., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and
436 Gamblin, S.J. (2020). Receptor binding and priming of the spike protein of SARS-CoV-2 for
437 membrane fusion. Nature *588*, 327–330.
438 Benton, D.J., Wrobel, A.G., Roustan, C., Borg, A., Xu, P., Martin, S.R., Rosenthal, P.B., Skehel,
439 J.J., and Gamblin, S.J. (2021). The effect of the D614G substitution on the structure of the
440 spike glycoprotein of SARS-CoV-2. Proc. Natl. Acad. Sci. *118*, e2022586118.
441 Bosch, B.J., van der Zee, R., de Haan, C.A.M., and Rottier, P.J.M. (2003). The Coronavirus
442 Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of
443 the Fusion Core Complex. J. Virol. *77*, 8801–8811.
444 Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S.M., Walsh, R.M., Rawson, S., Rits-Volloch, S.,
445 and Chen, B. (2020). Distinct conformational states of SARS-CoV-2 spike protein. Science
446 *369*, 1586–1592.
447 Carnell, G.W., Ciazynska, K.A., Wells, D.A., Xiong, X., Aguinam, E.T., McLaughlin, S.H.,
448 Mallery, D., Ebrahimi, S., Ceron-Gutierrez, L., Asbach, B., et al. (2021). SARS-CoV-2 Spike
449 Protein Stabilized in the Closed State Induces Potent Neutralizing Responses. J. Virol. *95*,
450 e0020321.
451 Emsley, P., and Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. Acta
452 Crystallogr. Sect. D Biol. Crystallogr. *60*, 2126–2132.
453 Gavor, E., Choong, Y.K., Er, S.Y., Sivaraman, H., and Sivaraman, J. (2020). Structural Basis of
454 SARS-CoV-2 and SARS-CoV Antibody Interactions. Trends Immunol. *41*, 1006–1022.
455 Ghosh, S., Dellibovi-Ragheb, T.A., Kerviel, A., Pak, E., Qiu, Q., Fisher, M., Takvorian, P.M.,
456 Bleck, C., Hsu, V.W., Fehr, A.R., et al. (2020). β-Coronaviruses Use Lysosomes for Egress
457 Instead of the Biosynthetic Secretory Pathway. Cell *183*, 1520-1535.e14.
458 Gobeil, S.M.-C., Janowska, K., McDowell, S., Mansouri, K., Parks, R., Manne, K., Stalls, V.,
459 Kopp, M.F., Henderson, R., Edwards, R.J., et al. (2021). D614G Mutation Alters SARS-CoV-2
460 Spike Conformation and Enhances Protease Cleavage at the S1/S2 Junction. Cell Rep. *34*,
461 108630.
462 Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., McKeane, L., Nakane, T., Zivanov, J., Neufeldt,
463 C.J., Cerikan, B., et al. (2020). Structures and distributions of SARS-CoV-2 spike proteins on
464 intact virions. Nature *588*, 498–502.
465 Kidmose, R.T., Juhl, J., Nissen, P., Boesen, T., Karlsen, J.L., and Pedersen, B.P. (2019).
466 Namdinator - Automatic molecular dynamics flexible fitting of structural models into cryo-
467 EM and crystallography experimental maps. IUCrJ *6*, 526–531.
468 Klasse, P., and Moore, J.P. (2020). Antibodies to SARS-CoV-2 and their potential for
469 therapeutic passive immunization. Elife *9*.
470 Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner,
471 N., Giorgi, E.E., Bhattacharya, T., Foley, B., et al. (2020). Tracking Changes in SARS-CoV-2
472 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell *182*, 812-
473 827.e19.
474 Krammer, F. (2020). SARS-CoV-2 vaccines in development. Nature *586*, 516–527.
475 Li, F. (2016). Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu. Rev.

476 Virol. *3*, 237–261.
477 Mastronarde, D.N. (2005). Automated electron microscope tomography using robust
478 prediction of specimen movements. J. Struct. Biol. *152*, 36–51.
479 Raj, V.S., Mou, H., Smits, S.L., Dekkers, D.H.W., Müller, M.A., Dijkman, R., Muth, D.,
480 Demmers, J.A.A., Zaki, A., Fouchier, R.A.M., et al. (2013). Dipeptidyl peptidase 4 is a
481 functional receptor for the emerging human coronavirus-EMC. Nature *495*, 251–254.
482 Rosa, A., Pye, V.E., Graham, C., Muir, L., Seow, J., Ng, K.W., Cook, N.J., Rees-Spear, C.,
483 Parker, E., dos Santos, M.S., et al. (2021). SARS-CoV-2 can recruit a heme metabolite to
484 evade antibody immunity. Sci. Adv. *7*.
485 Toelzer, C., Gupta, K., Yadav, S.K.N., Borucu, U., Davidson, A.D., Kavanagh Williamson, M.,
486 Shoemark, D.K., Garzoni, F., Staufer, O., Milligan, R., et al. (2020). Free fatty acid binding
487 pocket in the locked structure of SARS-CoV-2 spike protein. Science *370*, 725–730.
488 Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veesler, D. (2020).
489 Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell *181*, 281-
490 292 e6.
491 Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.-L., Abiona, O., Graham, B.S.,
492 and McLellan, J.S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion
493 conformation. Science *367*, 1260–1263.
494 Wrobel, A.G., Benton, D.J., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and
495 Gamblin, S.J. (2020). SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on
496 virus evolution and furin-cleavage effects. Nat. Struct. Mol. Biol. *27*, 763–767.
497 Xiong, X., Qu, K., Ciazynska, K.A., Hosmillo, M., Carter, A.P., Ebrahimi, S., Ke, Z., Scheres,
498 S.H.W., Bergamaschi, L., Grice, G.L., et al. (2020). A thermostable, closed SARS-CoV-2 spike
499 protein trimer. Nat. Struct. Mol. Biol. *27*, 934–941.
500 Xu, C., Wang, Y., Liu, C., Zhang, C., Han, W., Hong, X., Wang, Y., Hong, Q., Wang, S., Zhao, Q.,
501 et al. (2021). Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex
502 with receptor ACE2 revealed by cryo-EM. Sci. Adv. *7*, eabe5575.
503 Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Lu, G., Wu, Y., Yan, J., Shi, Y., et al.
504 (2017). Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the
505 dynamic receptor binding domains. Nat. Commun. *8*, 1–9.
506 Yurkovetskiy, L., Wang, X., Pascal, K.E., Tomkins-Tinch, C., Nyalile, T.P., Wang, Y., Baum, A.,
507 Diehl, W.E., Dauphin, A., Carbone, C., et al. (2020). Structural and Functional Analysis of the
508 D614G SARS-CoV-2 Spike Protein Variant. Cell *183*, 739-751.e8.
509 Zhang, J., Cai, Y., Xiao, T., Lu, J., Peng, H., Sterling, S.M., Walsh, R.M., Rits-Volloch, S., Zhu, H.,
510 Woosley, A.N., et al. (2021). Structural impact on SARS-CoV-2 spike protein by D614G
511 substitution. Science *372*, 525–530.
512 Zhou, T., Tsybovsky, Y., Gorman, J., Rapp, M., Cerutti, G., Chuang, G.-Y., Katsamba, P.S.,
513 Sampson, J.M., Schön, A., Bimela, J., et al. (2020). Cryo-EM Structures of SARS-CoV-2 Spike
514 without and with ACE2 Reveal a pH-Dependent Switch to Mediate Endosomal Positioning of
515 Receptor-Binding Domains. Cell Host Microbe *28*, 867-879.e5.
516 Zivanov, J., Nakane, T., and Scheres, S.H.W. (2019). A Bayesian approach to beam-induced
517 motion correction in cryo-EM single-particle analysis. IUCrJ *6*, 5–17.
518 Zivanov, J., Nakane, T., and Scheres, S.H.W. (2020). Estimation of high-order aberrations and
519 anisotropic magnification from cryo-EM data sets in RELION-3.1. IUCrJ *7*, 253–267.
520

521 **Methods**
522
523 *Expression Constructs*
524 Expression constructs were generated essentially as described in Xiong et al. 2020 (Xiong et
525 al., 2020). Starting from the S-R construct (Xiong et al., 2020), introduction of cysteines to
526 form x3 disulfide and introduction of the D614G substitution were carried out using Q5
527 polymerase PCR with primers containing desired substitutions, followed by In-Fusion HD
528 (Takara Bio) assembly. The previously reported construct consisting of the ectodomain of
529 human angiotensin-converting enzyme 2 (ACE2, residues 19–615) fused with a Fc-tag (Raj et
530 al., 2013) was cloned into a pCDNA3.1 vector using In-Fusion assembly.
531
532 *Protein Production and Purification*
533 Spike proteins were expressed in Expi293 cells and purified by metal exchange
534 chromatography exactly as described in Xiong et al. 2020 (Xiong et al., 2020). Purified spike
535 proteins were quality checked by negative stain EM. To purify ACE2-Fc protein, 500 µg of
536 plasmid was incubated with 1350 µg of polyethylenimine for 10 min, and then transfected
537 into 500 mL of 293F cells at 3 million/ml. The transfected cells were cultured at 33°C. Cell
538 culture supernatant was collected and clarified at day 6 post transfection and loaded onto a
539 5 ml Hitrap protein A column (Cytiva). The column was washed with 100 ml PBS and eluted
540 with citric acid buffer (100 mM, pH 3.0). The eluted fractions were collected into 1 M Tris-
541 HCl (pH 8.0) buffer and neutralized. Purified ACE2-Fc was concentrated and buffer
542 exchanged into PBS by a 50-kDa MWCO ultra centrifugal filter (Millipore).
543
544 *Negative staining EM*
545 3 µl of proteins (~0.05 mg/ml) diluted with PBS buffer were deposited onto carbon-coated
546 grids (EMS CF200-Cu) glow-discharged for 15 seconds at 25 mA in air. After 60s incubation,
547 excess proteins were wicked by filter paper. Grids were washed once in buffer, and stained
548 twice in Nano-W stain (Nanoprobes) with blotting in between. The grids were air dried on
549 filter paper and imaged using a Tecnai T12 Spirit operated at 120 kV. Micrographs were
550 taken in Digital Micrograph (Gatan).
551
552 *Disulfide bond reduction under native conditions*
553 S-R/x2, S-R/x2/D614G, S-R/x3, and S-R/x3/D614G spike proteins were incubated with 0, 2.5,
554 5, 10, and 20 mM DTT for 5, 20 or 60 min at room temperature in PBS. Reactions were
555 stopped by addition of 55 mM iodoacetic acid (IAA) for 10 min in the dark at room
556 temperature. Reaction mixtures were mixed with 4× non-reducing loading buffer and were
557 analysed by SDS-PAGE.
558
559 *Biolayer interferometry*
560 Binding affinities of ACE2-Fc or CR3022 IgG towards SARS-CoV-2 spikes were assessed by BLI
561 on an Octet RED96 instrument (FortéBio, USA) following essentially the same protocol as
562 previously described (Carnell et al., 2021). All steps were performed at 25°C and at a shaking
563 speed of 1000 rpm. All reagents were formulated in PBS-TB buffer (PBS with 0.02% v/v
564 Tween-20 and 0.1% w/v BSA). Before the experiments, all biosensors were pre-equilibrated
565 in the PBS-TB buffer for 10 min. ACE2-Fc and CR3022 were diluted to 11 µg/mL to be
566 immobilized on Protein A biosensors (FortéBio, USA) to levels of ~1.2 and ~1.6 nm
567 respectively. After a 60 s baseline step in PBS-TB, the ligand-loaded biosensors were

568 submerged in spike protein solutions (3 fold serial diluted from 1500 to 2.06 nM for ACE2-Fc
569 binding or 2 fold serial diluted from 200 to 3.13 nM for CR3022 binding) for 300s and then
570 transferred into PBS-TB for 600 s to measure spike association and dissociation kinetics. To
571 assess the effect of DTT, equilibration, association and dissociation steps were performed in
572 PBS-TB buffer supplemented with 20 mM DTT. The biosensors were regenerated by 10 mM
573 glycine (pH 2.0) between subsequent experiments. Data were aligned, inter step corrected
574 to the association step and further analyzed using the Data Analysis HT12 software
575 (FortéBio) and results were plotted in GraphPad Prism7. Experiments were repeated at least
576 3 times to assess consistency and control experiments with S-R, S-R/PP were performed to
577 confirm consistency with our previous study (Carnell et al., 2021).
578
579 *Cryo-EM*
580 Grid preparation and image collection were performed for S-R/x3 and S-R/x3/D614G spike
581 proteins essentially as described in Xiong et al. 2020 (Xiong et al., 2020). C-Flat 2/2 3C grids
582 (Protochips) were glow-discharged for 45 seconds at 25 mA. 3 µl of freshly purified protein
583 at 0.6 mg/ml supplemented with 0.01% octyl-glucoside (OG) was applied to the grids, which
584 were plunge-frozen in liquid ethane using a Vitrobot (Thermo Fisher Scientific).
585
586 An aliquot of S-R/x3 freshly purified spike at 1.0 mg/ml was stored at 4˚C for 40 days. 10 µl
587 of the stored protein was subjected to plunge-freezing at 1.0 mg/ml following the same
588 procedure as for the freshly purified S trimers. Another 10 µl of the stored protein was
589 incubated with 1 µl of citrate acid (pH 4.8) overnight and then plunge-frozen.
590
591 Grids were stored in liquid nitrogen and loaded into a Titan Krios electron microscope
592 (Thermo Fisher Scientific) operated at 300kV. Movies with 48 frames were collected with a
593 Gatan K3 BioQuantum direct electron detector with the slit retracted. Three shots per hole
594 were achieved with beam-image shift controlled in SerialEM-3.7.0 (Mastronarde, 2005). An
595 accumulated dose of 50 electrons/$Å^2$ were acquired in counting mode at the magnification
596 of 81,000 X, corresponding to a calibrated pixel size of 1.061 Å/pixel. Detailed data
597 acquisition parameters are summarized in Extended Data Table 1.
598
599 *Cryo-EM image processing*
600 Real-time data processing was performed in RELION-3.1's Scheduler as described in (Xiong,
601 2020). Motion correction, contrast transfer function (CTF) estimation, template particle
602 picking and initial 3D classification were carried out while micrographs were being collected.
603 An EM structure of the SARS-CoV-2 S protein in open form was filtered to 20 Å resolution as
604 a 3D reference for template picking. Initial 3D classification with an open S model was
605 accomplished at bin4 in batches of 500,000 particles to identify S protein trimers. Subsets of
606 S trimers in the 3D classes which displayed clear secondary structures were pooled and
607 subjected to one round of 2D classification cleaning. Subsequently, a second round of 3D
608 classification was used to assess the ratio of closed and locked states (**Fig. S1**). Auto
609 refinement, Bayesian polishing and CTF refinement were performed iteratively on classified
610 closed and locked subsets, respectively (Zivanov et al., 2019, 2020). Following the final
611 round of 3D auto-refinement, map resolutions were estimated at the 0.143 criterion of the
612 phase-randomization-corrected Fourier shell correlation (FSC) curve calculated between
613 two independently refined half-maps multiplied by a soft-edged solvent mask. Final

614 reconstructions were sharpened and locally filtered in RELION *post-processing* (**Fig. S2**). The
615 estimated B-factors of each map are listed in Extended Data Table 1.
616
617 The radiation damage caused fading of the disulphide bond between the two engineered
618 cysteines in S-R/x3 and S-R/x3/D614G. In order to recover the bond density, EM maps of
619 closed and locked S trimers were reconstructed from the first 4 frames in the movies of S-
620 R/x3 and the first 5 frames of S-R/x3/D614G, respectively. Distinct densities of the
621 disulphide bond were observed in the EM structures reconstructed from early exposed
622 frames and are shown in **Fig. 2**.
623
624 *Model building and refinement*
625 For the closed conformation structures, the SARS-CoV-2 S protein ectodomain structure
626 (PDBID: 6ZOX (Xiong et al., 2020)) was fitted into the EM density as a starting model. For the
627 locked conformations, S structures from our previous study (PDBID: 6ZP2, 6ZOZ (Xiong et al.,
628 2020)) were used as starting models. Model building and adjustment were performed
629 manually in Coot-0.9 (Emsley and Cowtan, 2004). Steric clash and sidechain rotamer
630 conformations were improved using the Namdinator web server (Kidmose et al., 2019).
631 After further manual adjustment, the locked and closed structures were refined and
632 validated in PHENIX-1.18.261 (Afonine et al., 2018) to good geometry. Refinement statistics
633 are given in Extended Data Table 1.
634
635 *Mass spectrometry*
636 10 µl of purified spike protein at ~ 2 mg/ml was extracted with 80 µL acetonitrile, 3 µL of
637 organic supernatant was injected into liquid chromatography-quadrupole time of flight
638 mass spectrometry (LC-Q-TOF) for further analysis. The sample was separated using an
639 Agilent 1290 Infinity II LC system (Agilent, Singapore) with a BEH C18 column (2.1 × 100 mm,
640 1.7 µm, Waters). The mobile phase consisted of (A) 0.1% formic acid in water and (B) 0.1%
641 formic acid in acetonitrile (Optima LC/MS Grade, Fisher Chemical). The gradient was as
642 follows: 0-1min, 5% B; 1-15 min, 5% to 95% B; 15-20 min, 95% B. The flow rate was 0.3
643 mL/min; column temperature, 30°C. Metabolites were detected across a mass range of m/z
644 100 to 1700 using a 6546 Q-TOF mass spectrometer (Agilent, Singapore). MS parameters
645 were as follows: gas temperature, 325°C; gas flow, 8 L/min; nebulizer, 45 psig; sheath gas
646 temperature, 325°C; sheath gas flow, 8 L/min; VCAP, 3500 V; nozzle voltage, 1500 V.
647 AutoMS2 scan was used at 3 spectra/s, collision energy was set at 10, 20, 40 eV in collision
648 induced dissociation (CID). Data were collected and analyzed by MassHunter 10.0 software
649 (Agilent).

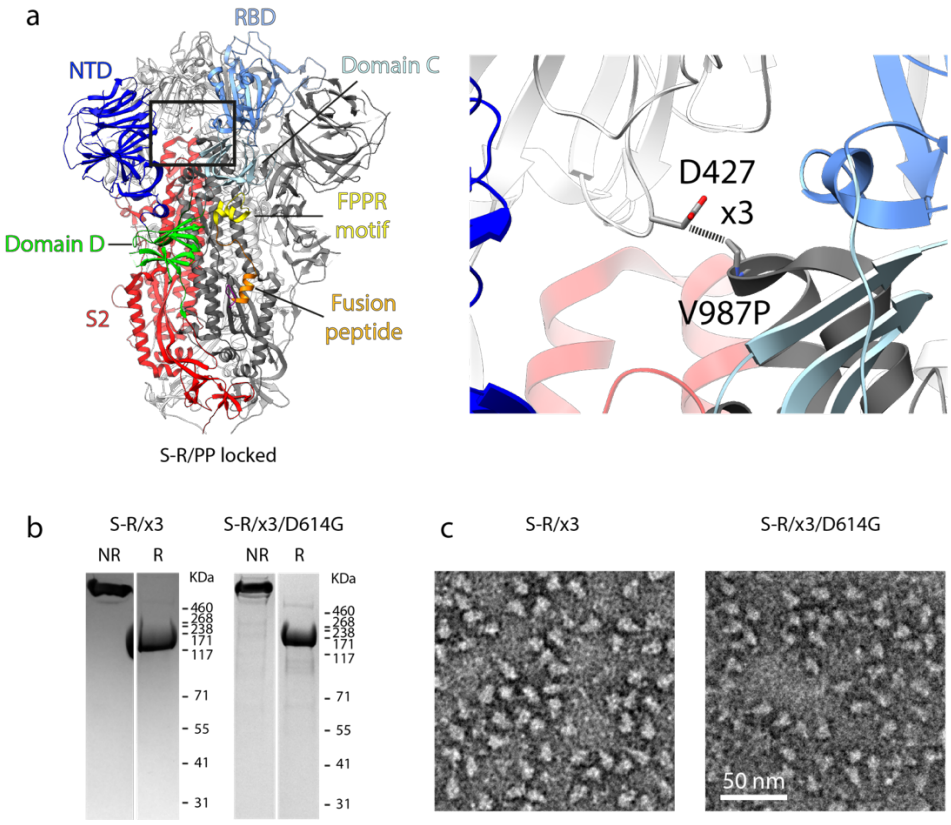
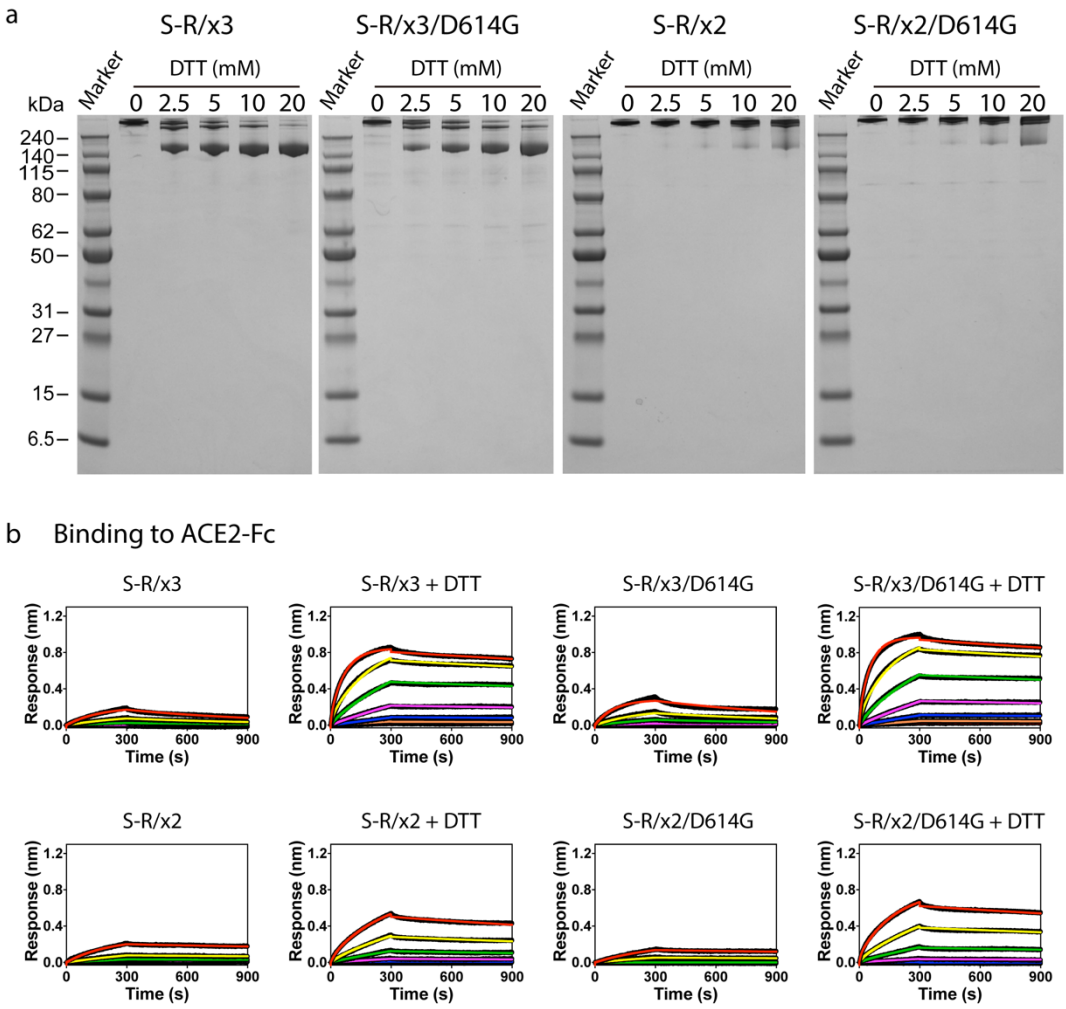


Figure 1. Design of the x3 disulfide bond to stabilise the SARS-CoV-2 spike in the "locked" state. a, Structure of the S-R/PP spike in the locked state. Structural domains are coloured and labelled. The box indicates the location of the zoomed-in view in the right panel. Residues 427 and 987 that were 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.



Ligand	Analyte	*kon* (M^{-1} s^{-1})	*koff* (s^{-1})	K_{Dkin} (nM)
ACE2-Fc	S-R/x3	3.23×10^3	12.48×10^{-4}	385.9
	S-R/x3+DTT	113.40×10^3 (fast)	1.68×10^{-4}	1.48
		9.47×10^3 (slow)		17.73
	S-R/x3/D614G	6.52×10^3	9.86×10^{-4}	151.23
	S-R/x3/D614G+DTT	136.60×10^3 (fast)	1.61×10^{-4}	1.18
		11.84×10^3 (slow)		13.61
	S-R/x2	3.06×10^3	2.21×10^{-4}	72.32
	S-R/x2+DTT	18.63×10^3 (fast)	3.22×10^{-4}	17.28
		2.20×10^3 (slow)		146.56
	S-R/x2/D614G	2.06×10^3	2.33×10^{-4}	112.84
	S-R/x2/D614G+DTT	75.61×10^3 (fast)	2.34×10^{-4}	3.09
		3.70×10^3 (slow)		63.23

Figure 2. Biochemical properties of purified x3 spikes comparing to x2 spikes. **a**, Reduction of x3 and x2 disulfide bonds under native conditions by 5 min incubation with indicated concentrations of DTT, reactions were stopped by excess amount of iodoacetamide before SDS-PAGE. **b**, binding of ACE2-Fc to different spike proteins in the absence and presence of 20 mM DTT. Spike proteins were serial diluted to 1500, 500, 166.7, 55.6, 18.5, 6.17, 2.06 and 0 nM and used as analytes in the assays. Kinetic parameters (k_{on}, k_{off}) and dissociation constants derived from kinetic analyses (K_{Dkin}) are summarised in the table.

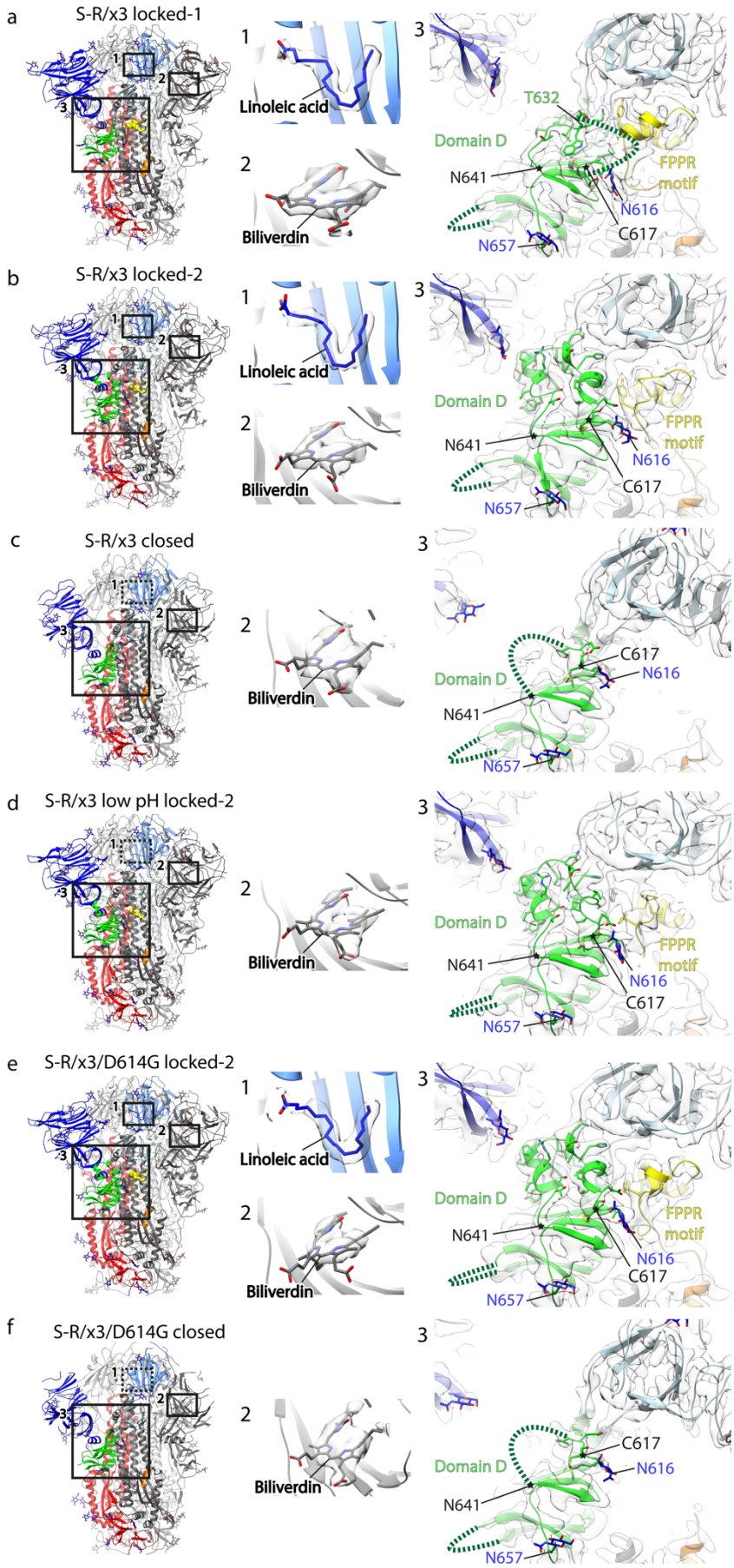

a
S-R/x3 locked-1
1
Linoleic acid
2
Biliverdin
3
T632
Domain D
N641
FPPR
motif
N616
C617
N657
b
S-R/x3 locked-2
1
Linoleic acid
2
Biliverdin
3
Domain D
N641
FPPR
motif
N616
C617
N657
c
S-R/x3 closed
2
Biliverdin
3
C617
N616
Domain D
N641
N657
d
S-R/x3 low pH locked-2
2
Biliverdin
3
Domain D
N641
FPPR
motif
N616
C617
N657
e
S-R/x3/D614G locked-2
1
Linoleic acid
2
Biliverdin
3
Domain D
N641
FPPR
motif
N616
C617
N657
f
S-R/x3/D614G closed
2
Biliverdin
3
C617
N616
Domain D
N641
N657

Figure 3. Structural features of spike proteins in different conformations as determined by cryo-EM. a-f, the 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 middle and right panels, dashed boxes indicate absence of bound factors. The middle panels show cryo-EM densities of bound factors; the right panels show cryo-EM densities of structural regions around domain D. Disordered loops are indicated by dashed lines, glycans are shown as stick representations in blue. The region between C617 and N641 (marked with asterisks) changes conformation between the different structures.

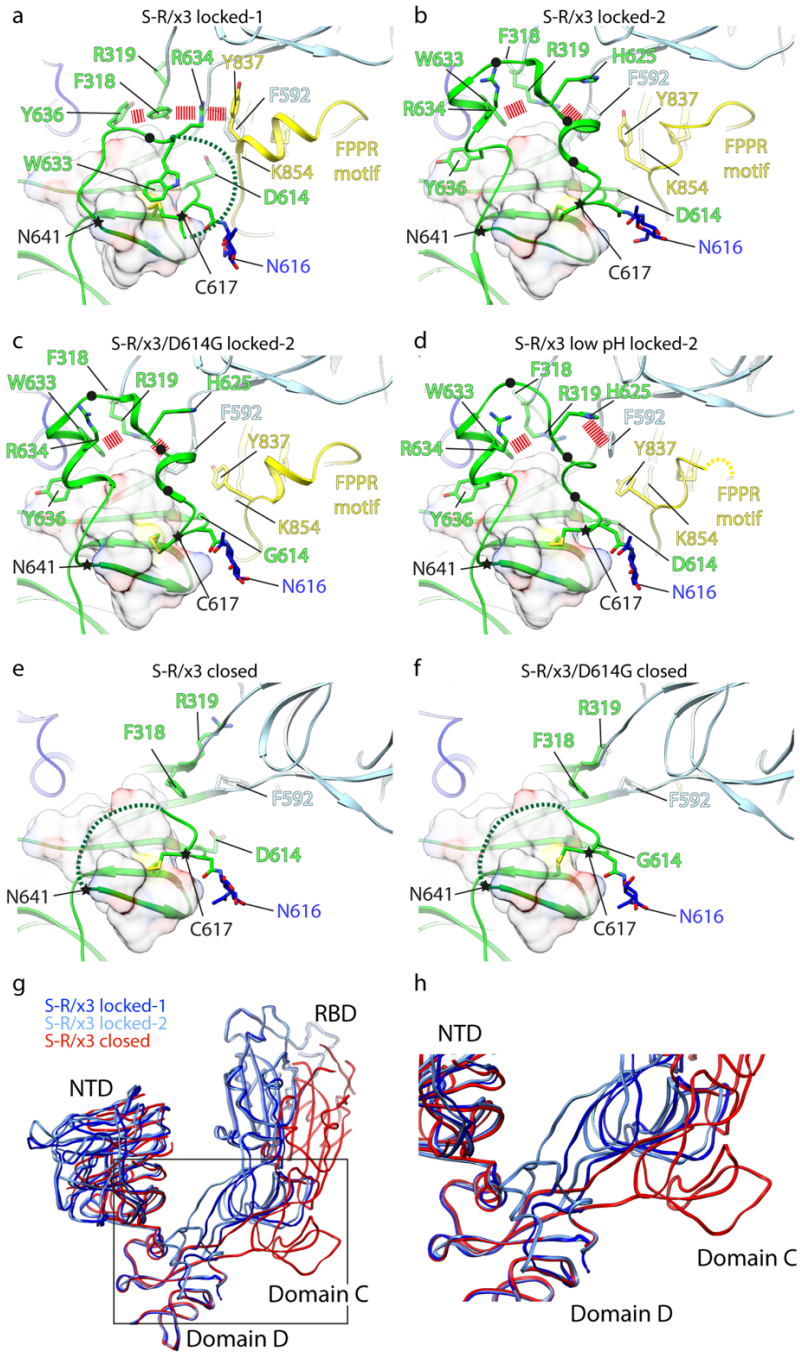


Figure 4. Structural changes in domain D between spikes in locked-1, locked-2 and closed conformations. a-f, structures of domain D in S-R/x3 and S-R/x3/D614G spikes of different conformations. Domain C is light blue, Domain D green, FPPR yellow. Selected amino acid sidechains are shown and key amino acid sidechains are marked. The disordered loops are represented by dashed lines. C617 and N641 are marked with stars to highlight the dynamic region in between. π-π and cation-π interactions stabilising locked conformations are highlighted with dashed lines. Positions of hydrophobic residues with omitted side chains interacting with the domain D hydrophobic core are marked by black dots. The hydrophobic core formed by the domain D beta sheet is shown as a transparent molecular surface. **g,** overlay of the S1 backbone from different conformations of the S-R/x3 spike. **h**, the region within the box in panel **g** is zoomed to highlight the movements of the domain C/D junction and domain C between locked and closed conformations.

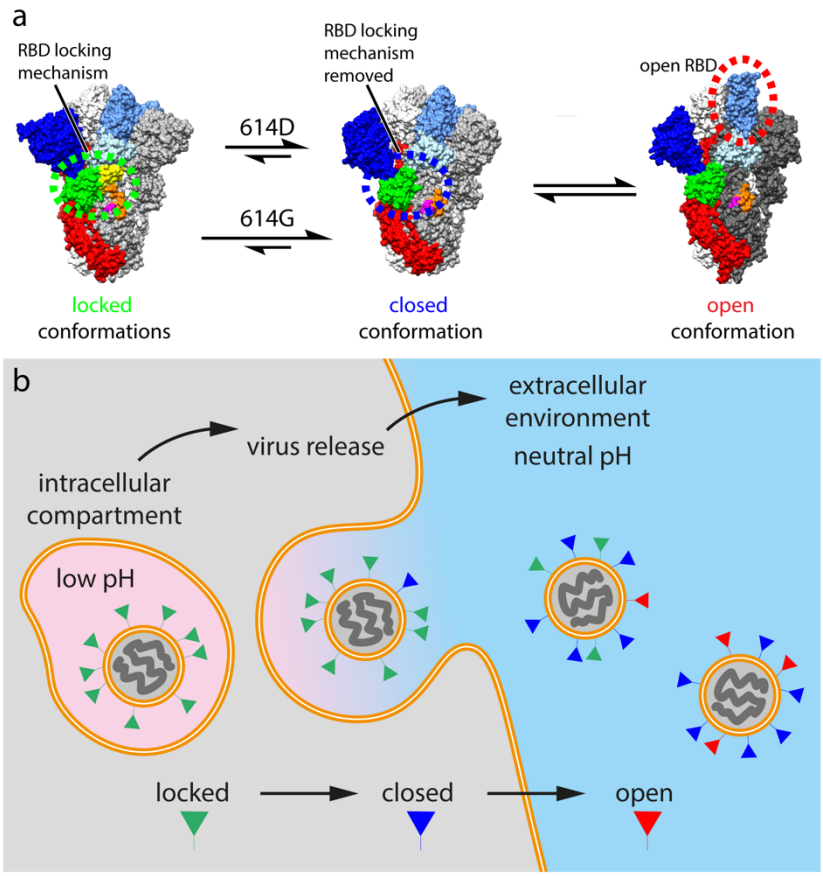
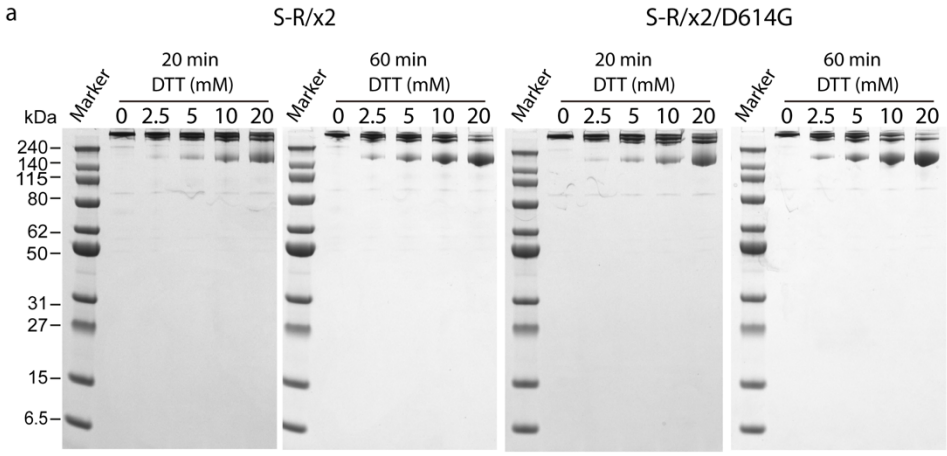
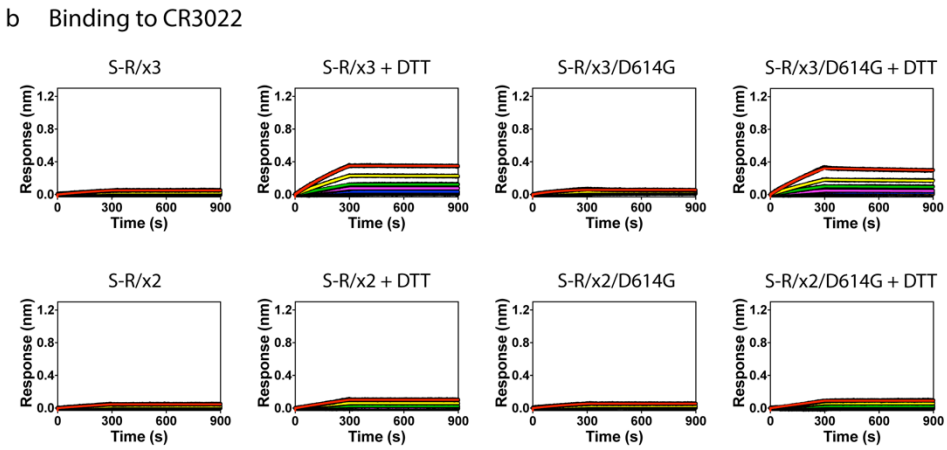


Figure 5. SARS-CoV-2 virus particle release and the accompanied structural transitions of surface spike protein. **a,** the three prefusion conformational states observed for the SARS-CoV-2 spike protein and factors that influence the structural transitions between the conformational states. **b,** schematic diagram illustrating the release of nascent SARS-CoV-2 virus particles from the cell. Green, blue or red triangles indicate the predicted conformational states of spike proteins as the virus travel through the acidic (pink) intracellular compartment before it was released into the neutral (blue) extracellular environment.





Ligand	Analyte	*kon* (M^{-1} s^{-1})	*koff* (s^{-1})	K_{Dkin} (nM)
CR3022	S-R/x3	-	-	weak binding
	S-R/x3+DTT	10.1×10^{3}	0.151×10^{-4}	1.5
	S-R/x3/D614G	-	-	weak binding
	S-R/x3/D614G+DTT	10.3×10^{3}	0.115×10^{-4}	11.2
	S-R/x2	-	-	weak binding
	S-R/x2+DTT	-	-	weak binding
	S-R/x2/D614G	-	-	weak binding
	S-R/x2/D614G+DTT	-	-	weak binding

Figure S1. Biochemical properties of purified spike proteins. a, Reduction of x2 disulfide bonds under native conditions by 20 and 60 min incubation with indicated concentrations of DTT, reactions were stopped by excess amount of iodoacetamide before SDS-PAGE. **b**, binding of CR3022, an antibody targeting a cryptic epitope on RBD, to different spike proteins in the absence and presence of 20 mM DTT. Spike proteins were serial diluted to 200, 100, 50, 25, 12.5, 6.25, 3.13 and 0 nM and used as analytes in the assay.

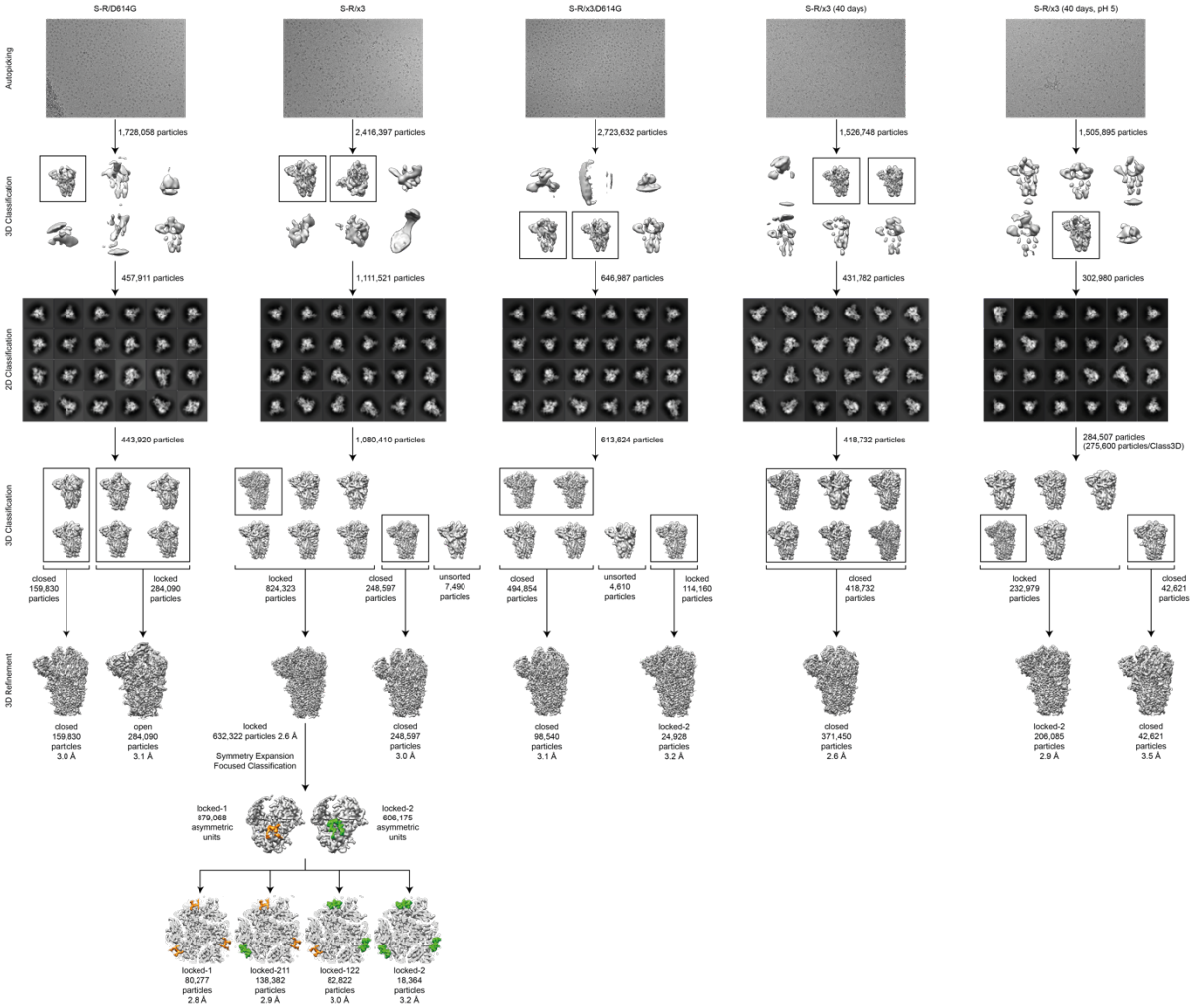


Figure S2. Pipeline used for picking and classification of cryo-EM data. Pipeline is illustrated for S-R/D614G, S-R/x3, S-R/x3/D614G, S-R/x3 after 40 days, and the latter after transfer to pH5.0 buffer. After automated picking, 3D and 2D classification steps were used to remove contaminating objects. 3D classification was then used to sort the data into locked and closed conformations.

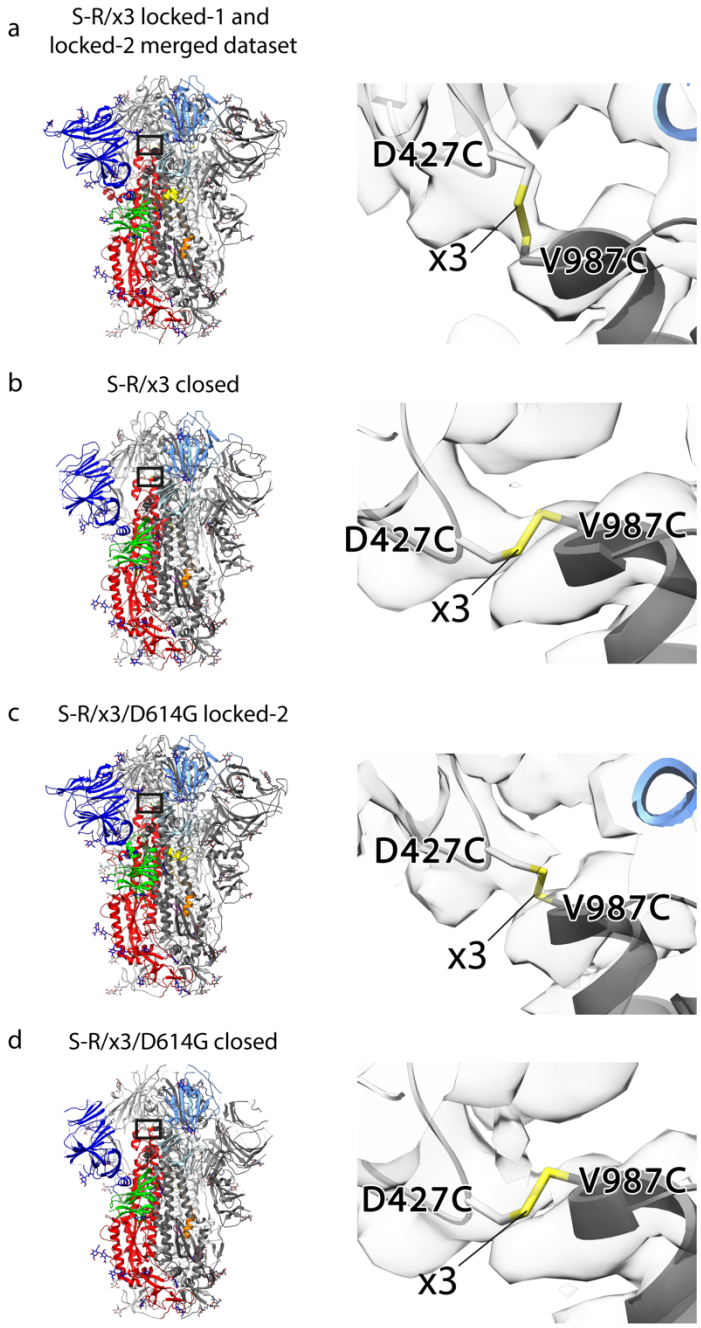


Figure S3. Cryo-EM densities of x3 disulfide bonds. a-d, left panels show the locations of the x3 disulfide bonds within the corresponding spike trimer structures; right panels show molecular models and densities of the x3 disulfide bonds. The radiation damage caused fading of the disulphide bond between the two engineered cysteines in the S-R/x3 and S-R/x3/D614G final maps. In order to recover the disulfide bond densities, 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 disulfide bond were observed in the EM structures reconstructed from early exposed frames and are shown the right panels.

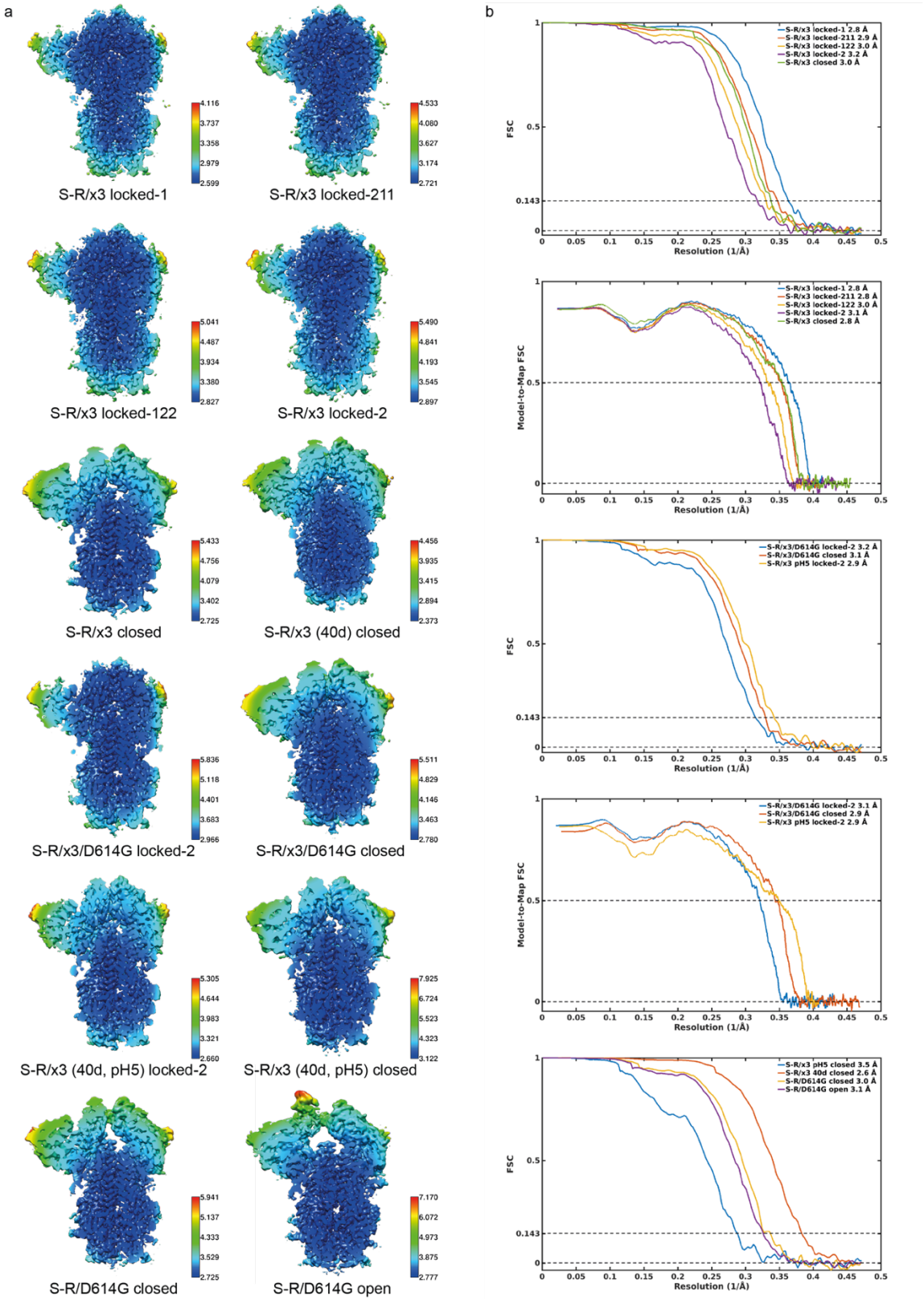


Figure S4. Resolution assessment of cryo-EM structures. a, Local resolution maps for all structures calculated using RELION. **b,** Global resolution assessment by Fourier shell correlation at the 0.143 criterion (left panel), and correlations of model vs map by Fourier shell correlation at the 0.5 criterion (right panel).

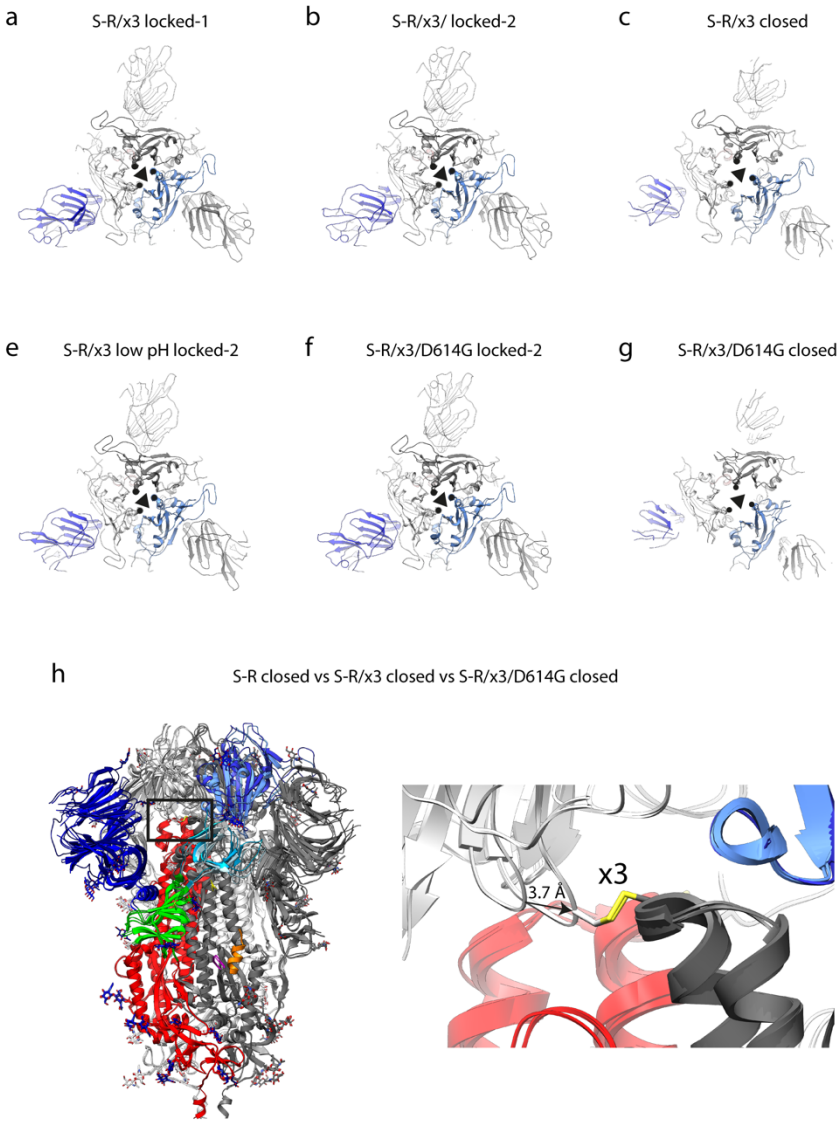
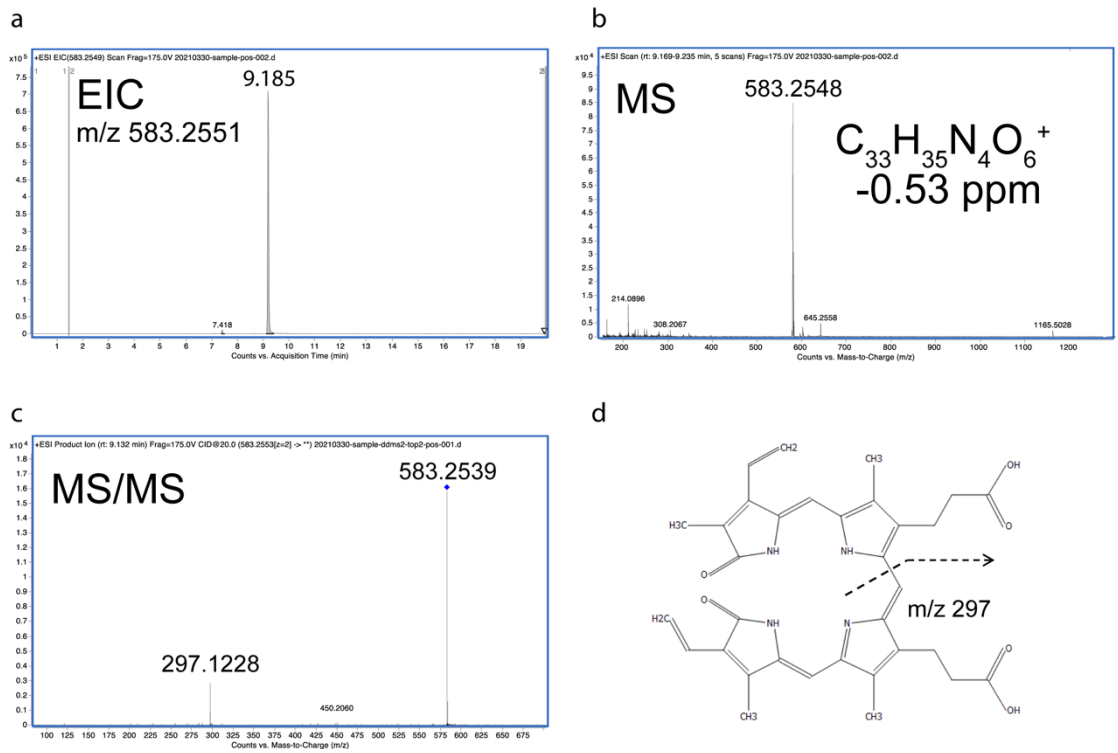
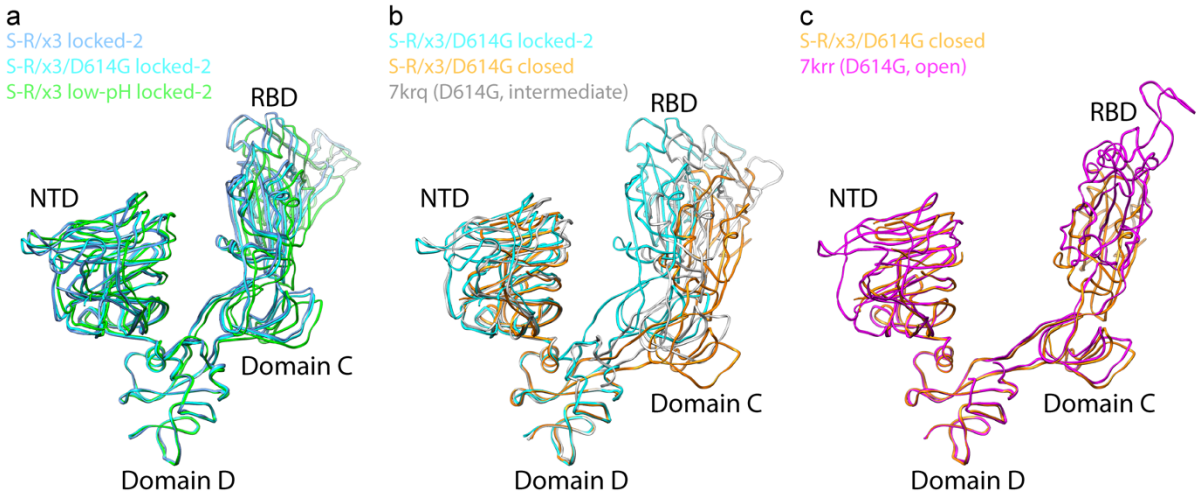


Figure S5. Structural features of S-R/x3 and S-R/x3/D614G S proteins. a-g, top views of S-R/x3 and S-R/x3/D614G spikes in locked-1, locked-2 and closed conformational states. The three-fold axes are indicated by triangles. The structural elements that are closest to the three-fold axes in the locked structures are indicated by dots (panels a, b, e, f). 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 (panels c and g). **h,** comparison of S-R, S-R/x3 and S-R/x3/D614G spikes in closed conformations. Left panel, overlay of the three structures, the box indicates the zoomed-in view in the right 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).



FigureS6. **Identification of biliverdin IXα by mass spectrometry. a**, extracted-ion chromatogram (EIC) of organic extract of purified spike protein identifies a prominent peak with an elution time of 9.185 min. **b**, MS analysis of the peak in panel a identifies a major ion species with an m/z of 583.2551 consistent to the structure of biliverdin IXα (panel d) with a mass difference of -0.53 ppm. **c**, MS/MS spectrum of biliverdin IXα in the sample identifies the parent ion (indicated by the blue diamond) and a fragment with an m/z of 297.1228 corresponding to the proposed fragment ion of biliverdin IXα (indicated by the dashed line in panel d). **d**, Chemical structure of biliverdin Ixα and proposed fragmentation pattern in MS/MS (dashed line).



FigureS7. Comparisons of spike proteins in different conformations. a-c, S1 parts of various different spike structures were aligned using domain D to highlight the movement of the RBD and domain C between the different structures.

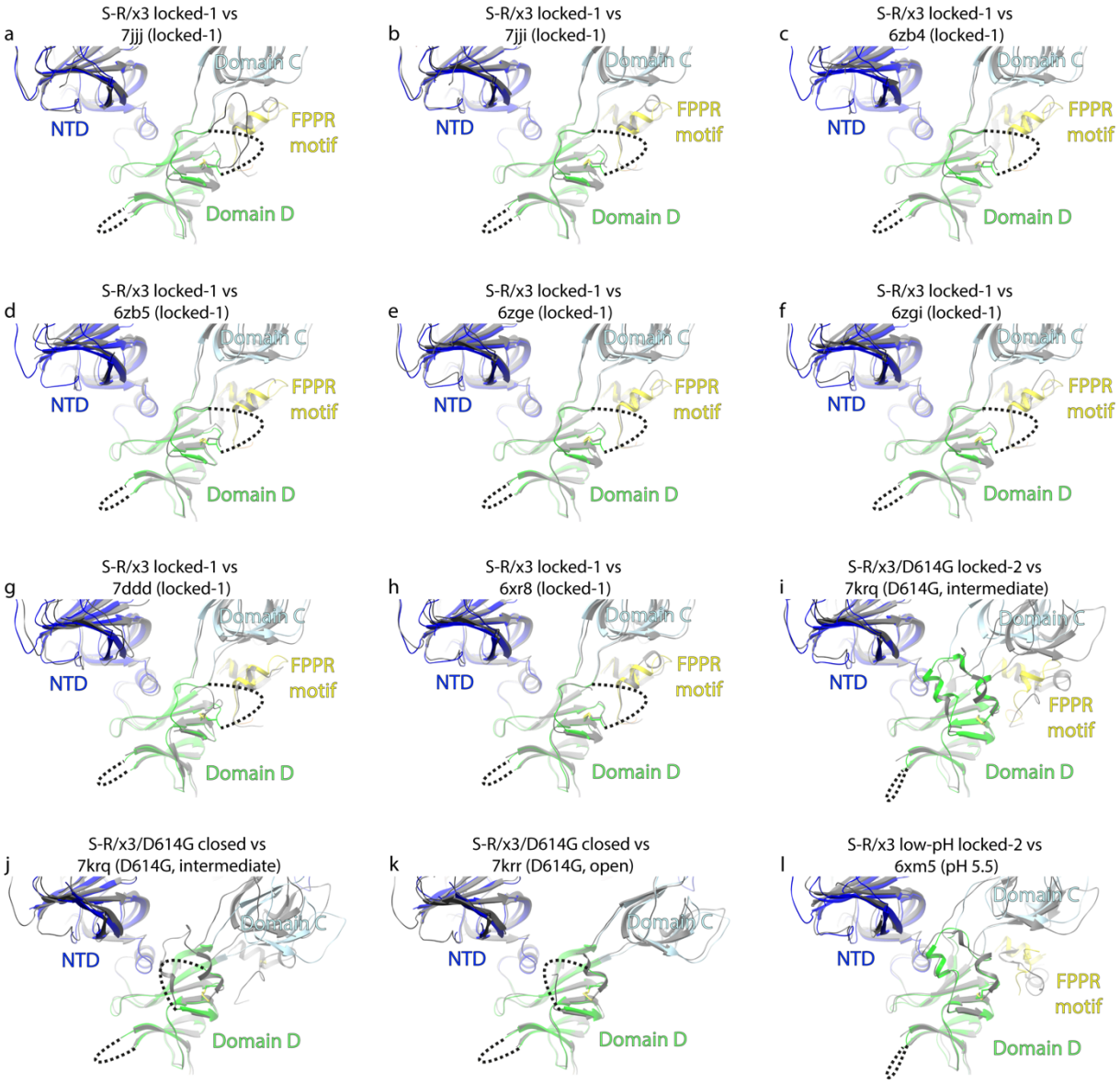


Figure S8. Domain D and surrounding structural features of S-R/x3 locked-1, S-R/x3/D614G locked-2, S-R/x3/D614G closed, S-R/x3 low-pH locked-2 structures compared to different deposited structures. a-h, S-R/x3 locked-1 (colored) was compared to various locked SARS-CoV-2 spike structures (gray), in most of these structures, the domain D and surrounding area adopt a similar structure as in the S-R/x3 locked-1 structure. **i-j,** The full-length D614G SARS-CoV-2 spike structure (gray, PDB: 7krq) was compared to S-R/x3/D614G locked-2 and S-R/x3/D614G closed structures (colored) to show the structure is in a intermediate state between locked and closed conformations. **k,** S-R/x3/D614G closed structure compared to the full-length D614G SARS-CoV-2 spike in open conformation. **l,** S-R/x3 low-pH locked-2 structure compared to another deposited low-pH SARS-CoV-2 spike structure (6xm5, chain A), which shows the most well-defined densities for regions around domain D among their low-pH structures (6xm4, 6xm5, 6xlu, 7jwy) (Zhou et al., 2020). PDB-IDs are from the following studies: 7jjj, 7iii (Bangaru et al., 2020); 6zb4, 6zb5 (Toelzer et al., 2020); 6zge, 6zgi (Wrobel et al., 2020); 7ddd (Xu et al., 2021); 6xr8 (Cai et al., 2020); 7krq, 7qrr (Zhang et al., 2021); 6xm5 (Zhou et al., 2020).

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

	S-R/x3 locked-1 (EMD-xxxx, PDB xxxx)	S-R/x3 locked-211 (EMD-xxxx, PDB xxxx)	S-R/x3 locked-122 (EMD-xxxx, PDB xxxx)	S-R/x3 locked-2 (EMD-xxxx, PDB xxxx)	S-R/x3 closed (EMD-xxxx, PDB xxxx)
Data collection and processing					
Magnification			81,000		
Voltage (kV)			300		
Electron exposure (e^{-}/Å^2)			50		
Defocus range (μm)			1.0-2.4		
Pixel size (Å)			1.061		
Movies (no.)			6,642		
Initial particle images (no.)			2,416,397		
Symmetry imposed	*C3*	*C1*	*C1*	*C3*	*C3*
Final particle images (no.)	80,277	138,382	82,822	18,364	248,597
Map resolution (Å)	2.8	2.9	3.0	3.2	3.0
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.60-5.50	2.72-5.86	2.83-6.93	2.90-7.35	2.73-6.69
Refinement					
Initial model used			PDB 6ZP2		PDB 6ZOX
Model resolution (Å)	2.8	2.8	3.0	3.1	2.8
FSC threshold	0.5	0.5	0.5	0.5	0.5
Map sharpening *B* factor (Å^2)	-40	-40	-40	-40	-25.3
Model composition					
Non-hydrogen atoms	26874	26907	26940	26997	24240
Protein residues	3291	3300	3309	3321	3006
Ligands	75	72	69	66	48
B factors (Å^2)					
Protein	48.16	51.57	59.92	68.06	92.77
Ligand	66.34	68.92	81.72	84.37	103.53
R.m.s. deviations					
Bond lengths (Å)	0.008	0.004	0.004	0.006	0.006
Bond angles (°)	0.760	0.754	0.652	0.717	0.836
Validation					
MolProbity score	0.99	1.19	1.31	1.36	1.43
Clashscore	1.28	2.41	2.71	2.26	3.73
Poor rotamers (%)	0.31	0.03	0	0.52	0.34
Ramachandran plot					
Favored (%)	97.24	97.03	96.26	94.76	96.13
Allowed (%)	2.76	2.97	3.74	5.24	3.87
Disallowed (%)	0	0	0	0	0

	S-R/x3/D614G locked-2 (EMD-xxxx, PDB xxxx)	S-R/x3/D614G closed (EMD-xxxx, PDB xxxx)	S-R/D614G closed (EMD-xxxx)	S-R/D614G open (EMD-xxxx)
Data collection and processing				
Magnification	81,000		81,000	
Voltage (kV)	300		300	
Electron exposure (e^-/Å^2)	50		50	
Defocus range (µm)	1.0-2.4		1.0-2.4	
Pixel size (Å)	1.061		1.061	
Movies (no.)	3,036		1,980	
Initial particle images (no.)	2,723,632		1,728,058	
Symmetry imposed	*C3*	*C3*	*C3*	*C1*
Final particle images (no.)	24,928	98,540	159,830	284,090
Map resolution (Å)	3.2	3.1	3.0	3.1
FSC threshold	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.97-7.94	2.78-6.78	2.73-7.28	2.78-9.88
Refinement				
Initial model used	PDB 6ZOZ	PDB 6ZOX		
Model resolution (Å)	3.2	2.9		
FSC threshold	0.5	0.5		
Map sharpening *B* factor (Å^2)	-30.0	-30.0	-40	-40
Model composition				
Non-hydrogen atoms	27069	22755		
Protein residues	3330	2814		
Ligands	69	45		
B factors (Å^2)				
Protein	77.75	76.80		
Ligand	93.88	88.66		
R.m.s. deviations				
Bond lengths (Å)	0.012	0.006		
Bond angles (°)	0.846	0.709		
Validation				
MolProbity score	1.51	1.35		
Clashscore	3.09	2.92		
Poor rotamers (%)	0.10	0		
Ramachandran plot				
Favored (%)	93.89	96.11		
Allowed (%)	6.11	3.89		
Disallowed (%)	0	0		

	S-R/x3 (40d) closed (EMD-xxxx)	S-R/x3 (40d, pH5) locked-2 (EMD-xxxx, PDB-xxxx)	S-R/x3 (40d, pH5) closed (EMD-xxxx)
Data collection and processing			
Magnification	81,000	81,000	
Voltage (kV)	300	300	
Electron exposure ($e^-/Å^2$)	50	50	
Defocus range (µm)	1.0-2.4	1.0-2.4	
Pixel size (Å)	1.061	1.061	
Movies (no.)	1,839	1,698	
Initial particle images (no.)	1,526,748	1,505,895	
Symmetry imposed	*C3*	*C3*	*C3*
Final particle images (no.)	371,450	232,979	42,621
Map resolution (Å)	2.6	2.9	3.5
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	2.37-5.47	2.66-6.73	3.12-9.79
Refinement			
Initial model used		PDB 6ZOZ	
Model resolution (Å)		2.9	
FSC threshold		0.5	
Map sharpening *B* factor ($Å^2$)	-31.9	-21.8	-50
Model composition			
Non-hydrogen atoms		26571	
Protein residues		3267	
Ligands		66	
B factors ($Å^2$)			
Protein		92.99	
Ligand		110.52	
R.m.s. deviations			
Bond lengths (Å)		0.009	
Bond angles (°)		0.914	
Validation			
MolProbity score		1.69	
Clashscore		4.75	
Poor rotamers (%)		0.42	
Ramachandran plot			
Favored (%)		93.12	
Allowed (%)		6.79	
Disallowed (%)		0.09	