Check for updates

The SARS-CoV-2 spike reversibly samples an open-trimer conformation exposing novel epitopes

Shawn M. Costello^{1,8}, Sophie R. Shoemaker^{2,8}, Helen T. Hobbs³, Annalee W. Nguyen⁴, Ching-Lin Hsieh⁵, Jennifer A. Maynard⁴, Jason S. McLellan⁵, John E. Pak⁶ and Susan Margusee^{2,3,6,7}

Current COVID-19 vaccines and many clinical diagnostics are based on the structure and function of the SARS-CoV-2 spike ectodomain. Using hydrogen-deuterium exchange monitored by mass spectrometry, we have uncovered that, in addition to the prefusion structure determined by cryo-electron microscopy, this protein adopts an alternative conformation that interconverts slowly with the canonical prefusion structure. This new conformation—an open trimer—contains easily accessible receptor-binding domains. It exposes the conserved trimer interface buried in the prefusion conformation, thus exposing potential epitopes for pan-coronavirus antibody and ligand recognition. The population of this state and kinetics of intercon-version are modulated by temperature, receptor binding, antibody binding, and sequence variants observed in the natural population. Knowledge of the structure and populations of this conformation will help improve existing diagnostics, therapeutics, and vaccines.

he spike protein from SARS-CoV-2 (also called the S-protein) is the primary target for current vaccines against COVID-19 and is the focus of many therapeutic efforts¹⁻⁴. This large, heavily glycosylated, trimeric protein is responsible for cell entry via recognition of the host receptor angiotensin-converting enzyme 2 (ACE2) and membrane fusion⁵⁻⁷. It is also the principal antigenic determinant of neutralizing antibodies⁸. Shortly after the determination and release of the viral genome sequence, a version of the spike ectodomain (termed S-2P) was designed to stabilize the prefusion conformation, and the structure was determined by cryo-electron microscopy (cryo-EM)^{9,10}.

This S-2P ectodomain comprises the first ~1,200 residues of the spike protein (Fig. 1a), with two proline substitutions in the S2 domain designed to stabilize the prefusion conformation, mutations that abolish the furin-cleavage site, and the addition of a C-terminal trimerization motif⁹. This mimic, its structure, and others that followed have been widely used for vaccine development and interpretation of many structure-function and epidemiological studies. To date, there are more than 250 structures of SARS-CoV-2 spike ectodomains in the Protein Data Bank¹¹.

These structural studies, together with other functional studies, demonstrate that, like all class 1 viral fusion proteins, the spike protein is dynamic, sampling several different conformations during its functional lifecycle^{12,13}. The three individual receptor-binding domains (RBDs) sample an 'up' state and a 'down' state; the up state exposes the ACE2-binding motif and therefore is required for infectivity^{7,10,14,15}. After receptor binding and cleavage between the S1 and S2 domains, the protein undergoes a major refolding event to allow fusion and adopts the stable postfusion conformation^{6,7,16-18}.

Despite the wealth of structural information, there are very few experimental studies on the dynamics within the prefusion state. The noted RBD up/down conformational transition has been monitored on the membrane via single-molecule FRET and occurs on the order of seconds¹⁹. Large computational resources have been devoted to molecular simulations of the spike protein, revealing a dynamic prefusion state with a range of accessible conformations, including the potential of a further opening of the RBD and N-terminal domain (NTD) away from the trimer interface^{20,21}. Experimentally, the conformational landscape of spike has not been well interrogated, and the effects of perturbations, such as ligand binding (both receptor and antibodies) or amino acid substitutions found in emerging variants of concern are unknown.

For these reasons, we turned to hydrogen-deuterium exchange monitored by mass spectrometry (HDX-MS) to probe the conformational landscape of the soluble spike prefusion ectodomain as well as the effects of ligand binding and sequence variation. We uncovered a stable alternative conformation that interconverts slowly with the canonical prefusion structure. This conformation is an open trimer, with easily accessible RBDs. It exposes the S2 trimer interface, providing new epitopes in a highly conserved region of the protein.

Results

Continuous-labeling HDX–MS on spike 2P. HDX–MS offers an ideal complement to the ever-growing number of structural studies on the SARS-CoV-2 spike protein, providing information on its conformational ensemble and dynamics. HDX–MS monitors the time course of the exchange of amide hydrogens on the peptide

¹Biophysics Graduate Program, University of California, Berkeley, CA, USA. ²Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. ³Department of Chemisal Engineering, The University of Texas at Austin, Austin, TX, USA. ⁵Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX, USA. ⁶Chan Zuckerberg Biohub, San Francisco, CA, USA. ⁷California Institute for Quantitative Biosciences, University of California, Berkeley, CA, USA. ⁸These authors contributed equally: Shawn M. Costello, Sophie R. Shoemaker. ^{Kag}e-mail: marqusee@berkeley.edu



Fig. 1 | **SARS-CoV-2** spike ectodomain and HDX-MS experimental overview. a, Schematic of the prefusion-stabilized SARS-CoV-2 spike protein and a model of the trimeric prefusion conformation²⁴. ss, signal sequence; NTD, N-terminal domain; RBD, receptor binding domain; FP, fusion peptide; TMD, transmembrane domain. b, Schematic of the HDX-MS experiment and the resulting mass distributions for a peptide that exists in either one (left) or two (right) separable conformations. For the two conformations to result in a bimodal mass distribution, they must not interconvert during the timescale of the HDX experiment (hours). Rapid interconversion would result in a single mass distribution with the ensemble averaged mass profile.

backbone with the hydrogens in the solvent (see Fig. 1b for description). An individual amide's ability to undergo exchange is directly related to its structure and stability^{22,23}.

We first followed the continuous-labeling time course of hydrogen exchange at 25 °C on the entire S-2P ectodomain, over a period of 15 seconds to 4 hours (Methods). Using a combination of porcine pepsin and aspergillopepsin digestion, we obtained 85% peptide coverage, allowing us to interrogate the dynamics of the entire protein (800 peptides, including 9 of the 22 glycosylation sites, average redundancy of 8.6) (Supplementary Fig. 1a). Notably, we have coverage in areas not resolved in the cryo-EM structure, including loops in the N-terminal domain (NTD) and RBD that are recognized by antibodies, loops in the S2 region that include the protease cleavage sites, and C-terminal residues after residue 1145 that include the second heptad repeat (HR2). Control experiments using deuterated protein demonstrate that our HDX protocol results in an average back exchange of 22% (Supplementary Fig. 1b).

The vast majority of peptides show a classic single isotopic envelope whose centroid increases in mass as deuterons are added over time (Fig. 1b). A small minority of the peptides, however, show bimodal behavior, with two isotopic envelopes both increasing in mass over time: one less exchanged distribution and a second more exchanged distribution (these peptides are described in detail below). The HDX profile of all the peptides, with the exception of the more exchanged distributions in the bimodal peptides, is consistent with the known prefusion conformation (Fig. 2 and Supplementary Fig. 2): secondary structure and buried elements within the trimer exchange slower than exposed loops. We also observe protection for residues 1140–1197, which includes HR2, a region not defined in single-particle cryo-EM structures, supporting the predicted helical structure of this region²⁴ and the relative rigidity of the stalk observed by cryo-electron tomography (cryo-ET)²⁵.

Identification of an alternative conformation. Bimodal mass envelopes can indicate the presence of two two different conformations that interconvert slowly on the timescale of our hydrogen-exchange experiment: one where the amides are more accessible to exchange compared to the other. However, it can also be a result of the kinetics of the hydrogen-exchange process itself, so-called EX1 exchange (when the rates of hydrogen-bond closing are much slower than the intrinsic chemistry of the exchange process). In this rare scenario, the heavier mass distribution will increase in intensity at the expense of the lighter one over the observed time period. This is not what we observe for the spike protein: the bimodal mass distributions

retain their relative intensities, increasing in average mass over time (Supplementary Fig. 3). The observed ratio is the same for every bimodal peptide under any given condition. The fact that every peptide shows the same ratio indicates that these bimodal peptides reflect two conformations; they report on the regions of the protein that show differences in hydrogen exchange in each conformation.

The bimodal peptides we observe are predominantly in the most conserved region of the protein-the S2 region²⁶ (Fig. 3a). When mapped onto the canonical prefusion conformation, many come from the helices at the trimer interface (residues 962-1024, 1146-1166, 1187–1196), indicating a second conformation with less stable hydrogen bonding for these helices. This is consistent with a loss of interprotomer contacts and increased solvent accessibility. We also observe bimodal peptides in other areas of the interprotomer interface, such as residues 870-916 in S2 and residues 553-574 and 662-673 in S1, again suggesting a change in trimer contacts. Finally, we see bimodal peptides in two regions that do not form interprotomer contacts (residues 291-305, 626-636); instead, these residues form the interface between the NTD and second S1 subdomain (SD2), suggesting that this subdomain interface is also structurally different in this second conformation. All the other peptides (the majority) fit to a classic unimodal distribution in the mass spectrum. The fact that the majority of the peptides are unimodal and that they behave the same in both conformations indicates that they exist in similar structures in each conformation and suggests that the individual domains have the same structure in each conformation. Previous HDX studies involving the spike protein did not note this behavior, which can be attributed to differences in experimental conditions and protocols^{27,28} (Supplementary Note).

Interconversion between the two conformations. These data suggest a model wherein the spike protein populates two conformations within the prefusion state—the classical prefusion structure seen in cryo-EM (herein referred to as state A)—and one in which each domain has a similar protomer topology and a more flexible or exposed open-trimer interface (herein referred to as state B). The above data do not, however, provide evidence that these states interconvert; any potential interconversion must be slower than the 4-hour hydrogen-exchange experiment. Since the transition of the RBD between the up and down conformation occurs on the order of seconds, this conformational heterogeneity is not the source of the bimodal distributions and the observed hydrogen-exchange reports on the weighted average of the two RBD conformations. There are several irreversible situations that could account for conformational

ARTICLES



Fig. 2 | Peptide-level HDX is consistent with the known prefusion conformation. Percent deuteration after 1 minute of deuterium labeling for every peptide in the S-2P continuous-labeling dataset for the S1 (top) and S2 (bottom) domains. Each line represents an individual peptide spanning the residues indicated on the *x* axis, with percent deuteration after 1 minute of exchange indicated on the *y* axis (for bimodal peptides, only the less exchanged centroid is shown). Secondary structures in the prefusion conformation are shaded in blue (alpha helices) and green (beta strands). A measure of solvent accessible surface area (SASA) is shown above in Å² (calculated as a three-residue sliding average using the default get_area function in pymol), using a model of the full-length prefusion trimer with all three RBDs in a down position²⁴. These data are consistent with the SARS-CoV-2 spike trimer secondary structures, notably regions buried in the trimer interface, such as the central helix, showing increased protection relative to more exposed regions lacking secondary structure. Important sequence features are indicated above the plot including the NTD (green), RBD (blue), fusion peptide (FP, cyan), heptad repeat 1 (HR1, yellow), central helix (CH, orange), core domain (CD, purple) and HR2 (yellow). Locations of glycans are noted with stars, and these belong to one of three categories: glycans detected in at least one peptide of our dataset (black, 9/22), glycans known to be on the spike protein where we lack coverage (white, 12/22), and glycans known to be on the spike protein but for which non-glycosylated peptides were observed (pink, 1/22).

heterogeneity, such as differences in glycosylation, proteolytic degradation, irreversible misfolding or aggregation. To rule these out, we tested whether the two conformations interconvert reversibly. We used bimodal peptides to quantify the population of each conformation under differing conditions (such as temperature, time, ligands). Under each condition, we carried out a 1-minute pulse of hydrogen exchange (25 °C) and integrated the area under the two mass envelopes for a single bimodal peptide to ascertain the fraction of each conformation under that condition or moment in time (Methods). We monitored two bimodal peptides, one from region 878-903 and one from region 978-1001. Using these two peptides, we quantified the population of each conformation under different conditions²⁹. We chose these bimodal peptides because of their high signal-to-noise ratio and because they report on two distinct regions, thus providing information from both the top and bottom of the S2 interface. We selected a pulse length of 1 minute, as this provides clearly separable bimodal distributions for these two peptides. For every condition tested, irrespective of the state A:state B ratio, both peptides report the same fractional population for the two conformations, indicating that all these data can be best described as a variable mixture of just two conformations: the canonical prefusion conformation and an unexpected alternative conformation.

Long-term incubation (4 days, 25 °C, pH 7.4) demonstrated a slow shift in population from a majority in the canonical prefusion

state (state A) to a majority in the alternative conformation (state B) (Fig. 3b and Supplementary Fig. 4a). Thus, the prefusion state can transform into the alternative state and the bimodal behavior cannot be due to sample heterogeneity, such as differential glycosylation. This observed $A \rightarrow B$ conversion, however, does not rule out an irreversible process, such as degradation or misfolding.

Postulating that the bimodal peaks represent a reversible structural transition, we used temperature to perturb the system and then investigated the protein's ability to interconvert. Indeed, the conformations do interconvert reversibly, with a preference for B at 4 °C and A at 37 °C (Fig. 3b). The observed kinetics of interconversion are extremely slow: for A→B, the halftime ($t_{1/2}$) was ~17 hours at 4 °C, but for the same sample at ~37 °C, $t_{1/2}$ for B→A was ~9 hours (Fig. 3c and Table 1). Notably the final, and presumably equilibrium, distribution at either temperature shows an observable population of both states (>5%), suggesting an energy difference of less than 2 kcal/mol between the two conformations under both conditions.

Structural studies on the prefusion conformation of spike have noted a temperature-dependent behavior. Cryo-EM experiments of spike incubated at 4°C for 5–7 days showed less than 10% of the definable prefusion spike particles seen on grids of freshly prepared spike. Incubating spike at 37 °C for three hours after storage at 4 °C recovered particle density to the level seen using freshly prepared protein³⁰. Failure to detect particles also correlated with a loss in

recognition by an antibody known to recognize quaternary structure. These studies are consistent with our findings—long-term incubation of spike at 37 °C biases toward the prefusion conformation, whereas long-term incubation at 4 °C prefer an expanded conformation, which is apparently not well visualized on cryo-EM grids. Our results suggest that, although there is a loss of quaternary structure, the HDX protection and therefore secondary structures for each domain are similar for both states, indicating that each protomer is still structured and not denatured—a feature uniquely addressable by HDX.

Effect of sequence changes—HexaPro. The apparently small energetic difference between these states indicates that small sequence changes may affect the relative populations and/or rates of interconversion between them. Indeed, the S-2P variant was designed to stabilize the prefusion conformation, avoiding spontaneous conversion to the postfusion form. This S-2P construct is the basis for most currently employed vaccines. Shortly after the determination of the S-2P structure by cryo-EM, a new version was constructed, termed HexaPro or S-6P, which contains four additional proline mutations designed to increase the apparent stability of the prefusion state and improve cellular expression³¹.

Using the same pulsed-labeling HDX–MS process, HexaPro shows the same bimodal behavior, with the same regions reporting on the two conformations (Fig. 3a and Supplementary Fig. 4b). At 4 °C, HexaPro, like S-2P, converts to state B, but with slower kinetics ($t_{1/2}$ of ~6 days). At 37 °C, HexaPro shifts back to state A, with a $t_{1/2}$ of ~2 hours (Fig. 3c). In sum, as expected on the basis of the design criteria, HexaPro does result in a bias toward the prefusion conformation. Importantly, these changes demonstrate how a small number of mutations can perturb and modulate the conformational landscape of spike, suggesting that the evolving sequence variants may show differences in this conformational exchange (see below).

Effects of sequence changes—an interprotomer disulfide bond. To further probe the structural features of the B conformation, we turned to a variant of HexaPro engineered to contain a disulfide bond. This variant trimer contains three disulfide bonds (S383C/D985C) that reach across protomers and lock the RBDs in the down state^{32–34}. We found that, when probed by continuous-labeling HDX–MS, this disulfide-locked variant remains completely in the A state and does not show any observable population of the B state, even after O/N incubation at 25 °C (Supplementary Fig. 3c). These data are consistent with a model where formation of the B state requires opening of the interprotomer (trimer) interface and exposure of the RBDs, which would be prohibited by the interprotomer crosslinks.

Fig. 3 | The Spike ectodomain reversibly samples two conformations.

a, Left, SARS-CoV-2 spike monomer with all regions that have peptides showing bimodal mass distributions colored in blue. Right, Example mass spectra from two peptides after one minute of deuteration (top, residues 982-1001; bottom, residues 878-903), with overlaid fitted gaussian distributions that describe each protein conformation in blue (light blue, less exchanged A state; dark blue, the more exchanged B state). b, Conformational preference for the S-2P spike construct at 25 °C, 4 °C and 37 °C, monitored by pulsed-labeling. At 25 °C, S-2P converts from primarily state A to ~50:50 A:B after 4 days. At 4 °C, S-2P prefers state B; at 37 °C, it prefers state A. c, The kinetics of interconversion between the A and B states for different spike variants. Starting from an initial prefusion conformation (state A, 37 °C), samples were rapidly transferred to 4 °C and assayed for conversion to state B over time using pulsed-labeling HDX-MS. To estimate fraction state A, peptides from two different regions (residues 982-1001 (circles) and residues 878-903 (triangles)) were fit to two gaussians. Data from both regions were used to determine the rate of interconversion.

Effects of sequence changes—B.1.1.7 (Alpha) variant. Increasingly infectious SARS-CoV-2 variants of concern are being discovered throughout the global population on a regular basis. Most of these variants include mutations in the spike protein, primarily in the S1 domain; some reside in the ACE2-interaction surface, while others do not. Therefore, we asked whether these mutations can influence the biases and kinetics of interconversion of the A and B conformation. We monitored A/B conversion for a variant of HexaPro that includes five S1 mutations in the B.1.1.7 (Alpha) variant (H69–V70del (NTD), Y144del (NTD), N501Y (RBD), A570D (SD1), P681H (SD2)) and none in the S2 domain, termed Alpha S1 HexaPro. Indeed, Alpha S1 HexaPro shows notable differences in both the relative preference for state B and the kinetics of interconversion. At 4 °C, Alpha S1 HexaPro



Table 1 Rates of interconversion between the A (prefusion)
and B (open trimer) conformations of spike ectodomains

Temperature	Protein	k_{observed} (h ⁻¹)	t _{1/2} (h)
37 °C→4 °C	S-2P	0.04	17
	HexaPro	0.005	143
	Alpha S1 HexaPro	0.2	4
	S-2P+3A3	0.1	5
37 °C→10 °C	S-2P	0.05	14
	HexaPro	0.004	171
	Alpha S1 HexaPro	0.1, 0.2*	3, 5*
$4 ^{\circ}\text{C} \rightarrow 37 ^{\circ}\text{C}$	S-2P	0.08	9
	HexaPro	0.3	2
	Alpha S1 HexaPro	0.2	4

*, time course was monitored twice, and the results of each fit are reported. $k_{observed}$ is the observed rate of change in the population of the A state after a temperature jump. This relaxation rate is the sum of the forward and reverse rates, which is dominated by the major conformational change $(A \rightarrow B \text{ at } 4^{\circ}C, 10^{\circ}C \text{ and } B \rightarrow A \text{ at } 3^{\circ}C)$. k_{12} is the half time for that same rate, $\ln 2 / k_{barrend}$.

converts to state B nearly 20 times faster than does HexaPro (Fig. 3c and Table 1). Furthermore, Alpha S1 HexaPro shows no detectable prefusion conformer at 4 °C, while HexaPro shows at least 30% even after several weeks at 4 °C (Supplementary Fig. 4b). At 37 °C, the kinetics and equilibrium distribution appear nearly identical between the two. All of the mutations are at solvent-exposed residues, except residue 570, which contacts the S2 subunit and resides in a region with observed bimodal behavior. Thus, despite their location in the S1 subunit and not at the core trimer interface, these specific B.1.1.7 mutations allosterically affect the interconversion of these two states.

Effects of ACE2 binding. The primary function of the RBD is to recognize the host cell receptor ACE2. In the down conformation, the RBD is occluded from binding to ACE2, and in the up conformation, it is accessible. The entire trimer can exist with zero, one, two, or all three RBDs in the up conformation^{7,15}. In the isolated RBD, the receptor-binding motif (RBM) should always be accessible for ACE2 binding. We used continuous-labeling HDX to monitor the binding of the receptor, both in isolation and in the full-length spike (S-2P), using a soluble dimeric form of ACE2 (ACE2-Fc, herein referred to as ACE2). For isolated RBD (residues 319–541, see Methods), we obtained 141 peptides, including one glycosylated peptide spanning the N-glycosylation site at residue 343 (no peptides are observed for site 331), resulting in 82% sequence coverage with an average redundancy of 8 (Supplementary Fig. 1a).

The effects of ACE2 binding are illustrated in Fig. 4. In the presence of ACE2, the latter half of the RBM (residues 472–513), shows a notable decrease in hydrogen exchange upon binding ACE2 (Fig. 4B), consistent with the known ACE2–RBD binding interface^{35,36}. We also observe small changes for other regions that are near the binding interface. Importantly, we see very similar changes in HDX rates in the RBD both in isolation and in the context of the spike ectodomain, suggesting that all three RBDs in full-length spike interact with ACE2 in our experiment and that both the A and the B state can productively bind ACE2, which for the A state requires that the RBD transition to the up state.

In the context of the spike trimer, we also observe notable changes outside of the RBD, particularly in state A, where a few peptides exchange more rapidly in the presence of ACE2 (in state B, these peptides do not have any notable difference in the presence of ACE2) (Fig. 4c). These peptides are located on the top of S2 (residues 978–1001), a region known to become more exposed when RBD transitions from a down to an up conformation. Since ACE2 binding in the prefusion state requires the RBDs to be in the up conformation,

this increased exchange reflects the known biases in the RBD conformation—a prefusion state whose RBDs are primarily in the down conformation and must transition to an up conformation to bind ACE2. We also see changes in the interconversion between state A and state B in the presence of ACE2, such that state B is more preferred (Fig. 4d).

RBD dynamics are similar in isolation and in intact spike. The isolated RBD has been used for many biochemical studies and is the main component of many clinical diagnostics. It is therefore important to ask whether there are large differences in the RBD when it is isolated versus in the context of the spike trimer; our experiments allow us to directly compare the two. Very few peptides in the RBD show substantial changes in HDX behavior (Fig. 5) and support the use of approaches such as deep sequence mutagenesis on the isolated RBD to gain information on the potential effects of variants, such as escape mutations³⁷.

We do, however, see some key differences in the RBD—mostly at the termini of the isolated domain and in the expected interactions with the rest of spike and across the protomer interface. The C-terminal region of the RBD (residues 516–537) is notably less protected in the isolated domain. This region is not part of the RBD globular domain and in full-length spike forms part of subdomain 1, so an increase in flexibility when this region is isolated from the rest of the subdomain would not be surprising. Future studies with the isolated RBD may benefit from removal of both C-terminal and N-terminal (no peptide coverage observed for this region) regions, as they are likely disordered and may interfere with crystallization or lead to increases in aggregation.

3A3-an antibody that binds specifically to the B state. Recently, an antibody, 3A3, that binds to MERS-CoV, SARS-CoV-1, and SARS-CoV-2 was developed, with an apparent epitope in a region where we observe bimodal behavior (residues ~980-1000)²⁸. This region, however, is inaccessible in the prefusion structure-it is buried in the prefusion structure when all RBDs are down, and highly occluded when the RBDs are up. Our HDX data indicate that this region is exposed in state B. To confirm the epitope, we repeated the continuous-labeling HDX studies in the presence of 3A3; indeed, we see strong increased protection in the 978-1001 region. Moreover, this protection is directly associated with state B, as evidenced by the now unimodal distribution for peptides in region 978-1001 (Fig. 6a). This unimodal distribution can be explained by a decrease in hydrogen exchange in state B due to direct binding of 3A3 to this region such that the exchange is similar in both the A and B state. These data support a model in which the 3A3 antibody binds uniquely to the B state.

To confirm this hypothesis, we looked at the effect of 3A3 binding on the temperature-induced conversion between A and B using our pulsed-labeling HDX-MS method. 3A3 increases the rate of conversion from A to B at 4 °C, decreasing the $t_{1/2}$ from ~17 hours to ~5 hours (Fig. 6b and Table 1). This increase in the observed rate implies that 3A3 also affects the transition state for the conversion. Furthermore, returning the sample to 37 °C in the presence of 3A3 (state B saturated with 3A3) prohibits any transition back to the prefusion state, indicating that the binding of 3A3 prevents formation of the prefusion state, most likely owing to steric hindrance of the antibody being bound to the trimer interface. Since 3A3 binds wild-type spike and spike with the D614G mutation when expressed on the surface of cells and neutralizes pseudovirus expressing these spikes²⁸, the data collectively suggest state B exposes broadly neutralization-sensitive epitopes that may be of interest for future therapeutics and vaccines.

Discussion

The above data allow us to create a structural model for state B (Fig. 7). The overall fold, or topology, of each domain is likely similar to the

NATURE STRUCTURAL & MOLECULAR BIOLOGY



Fig. 4 | ACE2-binding effects are similar on isolated RBD and RBD in the context of the full-length ectodomain. a, Diagram of the spike structure, with three regions of interest (I,II,III) highlighted. **b**, Left, Heat map showing the difference in RBD peptide deuteration (from continuous-labeling HDX-MS) in the presence and absence of ACE2 on the isolated RBD. Middle, The deuterium-uptake plots are illustrated for three peptides of interest with error bars representing the s.d. of three replicates (some are smaller than marker size). The uptake plot for the RBM (residues 487-510) is shown for both the isolated RBD and HexaPro. Right, Schematic representation of the heat map data on the structure of the RBD-ACE2 complex (PDB: 6MOJ). The structure of the RBD is colored on the basis of the maximum change shown in the heat map for that residue in any peptide. **c**, Changes to peptides from HexaPro upon binding of ACE2 outside of the RBD during continuous-labeling HDX-MS. When ACE2 binds the canonical prefusion structure, state A, peptide 982-1001 (region II) loses intersubunit contacts with the RBD and thus exchanges faster; however, when ACE2 binds the open trimer (state B), it does not, presumably because it is already maximally exposed. For peptide 878-903 (region III), there is no change in the exchange rate to either state A or state B, indicating that this region is not affected by ACE2 binding. In the schematic for region II, one NTD has been removed to visualize the peptide of interest. **d**, Time course of interconversion in the presence of ACE2. Top, pulsed-labeling HDX-MS example spectra of S-2P peptide 878-902 with and without ACE2 before and after 24 hours of incubation at 25 °C. Bottom, time versus fraction state A as determined by monitoring peptide 878-902 in S-2P with and without ACE2 over 24 hours, monitored by pulsed-labeling. After 24 hours, ACE2-bound S-2P prefers state B.

prefusion structure as, with the exception of the bimodal peptides, their hydrogen exchange patterns are similar. The bimodal peptides, which report on the two different conformations, cluster in the trimer interface, suggesting that this interface is less protected in state B. State B is not a monomer. Size-exclusion chromatography, combined with multiangle light scattering (SEC–MALS) and the hydrogen-exchange data at the trimerization motif, confirm that both conformations are trimeric (Supplementary Fig. 5). In these soluble ectodomain constructs, the trimer is held together by the appended C-terminal trimerization domain, while in the full-length native spike trimer, the transmembrane helical segment likely serves this function. Therefore, state B is best modeled as an opened-up trimer, with three protomers whose domains are structurally uncoupled. Our data do not let us address the relative orientation

ARTICLES



Fig. 5 | Comparison of HDX on RBD in isolation versus in S-2P. a, Heat map showing the difference in peptide deuteration for isolated RBD compared with the RBD in S-2P. b, Selected uptake plots of isolated RBD and S-2P RBD. c, Structure of the RBD (model of a single RBD taken from a full-length spike trimer model from ref.²⁴), colored on the basis of the maximum change shown in the heat map for that residue in any peptide. For reference, spheres are shown denoting the beginning and end of the peptides displayed in the uptake plots.

of the protomers within individual trimers; however, an ensemble of opened-up trimers with heterogeneous positioning of the protomers would best explain the lack of cryo-EM data. An opened-up class 1 viral fusion protein has been reported for respiratory syncytial virus (RSV) and visualized by a low-resolution structure³⁸. These structural data from RSV and reports of an opening up of other viral fusion proteins^{39–41} support our model of an ensemble of open-trimers with various degrees of openness.

We propose that this conformational change occurs because the trimer interface of the spike protein consists of weak, hydrophobic

NATURE STRUCTURAL & MOLECULAR BIOLOGY



Fig. 6 | The antibody 3A3 binds selectively to state B in the 978-1001 region. a, Example mass spectra for HexaPro with and without 3A3 for two peptides that have bimodal mass distributions. The bottom peptide (878-902) shows no change in the presence of 3A3, which indicates that the amount of state A and state B has not changed 13 minutes after the addition of 3A3. The top peptide (978-1001), however, shows protection in the presence of 3A3, and the state B distribution is shifted to a deuteration amount indistinguishable from that of state A. These data are at the 3-minute time point from a continuous-labeling HDX-MS time course. **b**, The kinetics of interconversion of S-2P in the presence of 3A3, monitored by pulsed-labeling HDX-MS. The addition of 3A3 accelerated the rate of conversion to state B at 4 °C. The binding of 3A3 prevents the return to state A at 37 °C. Dotted lines indicate the conversion in the absence of 3A3.

interactions that are further weakened at low temperatures. This can be thought of as a form of cold denaturation in which the open trimer has a higher heat capacity than the canonical prefusion structure, consistent with solvation of the hydrophobic trimer interface. As for what is responsible for the slow rate of interconversion, this is still unknown. We believe that in order for state B to be formed, all three RBDs need to be in the up position, which



Fig. 7 | Schematic of the energy landscape for the SARS-CoV-2 spike ectodomain. Reaction coordinate illustrating the different conformations of spike. Three conformational states are depicted: the canonical prefusion ensemble, the expanded open trimer, and the postfusion conformation. The prefusion conformation contains all four RBD states (0, 1, 2, or 3 up). The relative energies and barrier heights as well as the placement of the open trimer along the reaction coordinate are drawn for illustration only.

could be the rate-limiting step for the formation of state B, thus the slow rate may be a result of the intertwined nature of the trimer. Although it is known that RBDs moving up and down occurs on the order of seconds, having all three RBDs open simultaneously long enough for the trimer interface to disassociate may be rare enough to result in a high kinetic barrier. Similarly, the role of the designed prolines (in S-2P and HexaPro) is unclear. The rates of interconversion are not consistent with a model in which the isomerization of prolines plays a role. Rather, our data are consistent with the model in which the introduction of prolines successfully stabilizes the prefusion conformation.

A loss of interprotomer contacts in state B implies that the RBDs no longer contact adjacent protomers and thus do occupy distinct 'down' and 'up' conformations; rather, they are likely always in a binding-competent state, perhaps even more accessible than the canonical 'up' state. This increased availability of the RBD may drive a preference for the B state in the presence of ACE2. Furthermore, in the prefusion conformation, having all three RBDs bound to ACE2 could lead to steric hindrances, but in state B, all three RBDs should be able to bind ACE2 with high affinity. Interestingly, introduction of variants of concern, such as in the alpha S1 HexaPro variant, greatly increases the rate of conversion to state B. Whether this plays a role in the noted increased infectivity remains to be determined.

Molecular dynamics has shown a smaller opening of the spike protein where an RBD and adjacent NTD twist and peel away from the center of the spike protein, revealing a cryptic epitope at the top of the S2 domain²⁰. This rapidly sampled conformation is not state B, as it does not involve the S2 trimer interface, and the timescale of conversion to state B is unlikely to be sampled during a molecular dynamics simulation. This partial opening, however, could be on pathway to state B.

A potential role for state B in spike function. The increase in the formation of state B upon binding ACE2 suggests that state B may be a functional intermediate. We cannot say whether this transition occurs in wild-type spike on virions, but it is possible that state B represents an intermediate along the pathway to S1 shedding during the transition from the prefusion conformation to the postfusion conformation. This irreversible transition is not possible in the soluble ectodomain, which removes the proteolytic cleavage site. If state B is an on-pathway intermediate, antibodies or ligands that trap

ARTICLES

the protein in this state may block the protein along the pathway to fusion⁴². Those that act on the transition state and increase the rate of formation of the intermediate could promote the premature formation of the postfusion conformation and thus aid in neutralization. Instead, if formation of state B is off-pathway, antibodies or other ligands that favor state B would essentially trap the protein in an inactive conformation.

The alternative conformation presents new druggable sites. On the basis of our model, the newly identified state B contains a large and unique accessible surface area that is buried in the canonical prefusion conformation, thus exposing different epitopes for antibody and ligand recognition. Moreover, these regions arise from the most highly conserved part of the protein, the S2 trimer interface, and therefore may present an ideal target for vaccines that would provide protection across coronaviruses (pan-coronavirus). In terms of therapeutics, ligands directed towards this region may also be broadly efficacious against variants of concern as well as other coronaviruses. Indeed, the antibody 3A3 represents one such potential therapeutic.

State B will affect measured ligand-binding affinities. Finally, independent of whether this additional conformation is an on-pathway intermediate in the coronavirus functional lifecycle, it is ubiquitous among in vitro preparations of the spike protein. We see evidence of this conformation in every variant examined, excluding the disulfide-locked sample. Many biochemical and diagnostic assays use these isolated spike constructs, and many laboratories store this protein at 4°C, where the alternative, expanded conformation is favored. Given that state A and B have differing affinities for the receptor and some antibodies, the temperature and time-dependent changes in the population of state B complicates quantitative analysis of binding affinities and needs to be further evaluated.

Conclusion

In sum, we have found that the SARS-CoV-2 spike ectodomain reversibly samples an open-trimer conformation, potentially allowing for the development of pan-coronavirus vaccines and therapeutics. This open trimer is folded and exposes a highly conserved region of the protein. It is similar in energy to the well-characterized prefusion conformation determined by cryo-EM. The fraction of spike in each conformation depends on temperature, ligands, and sequence. Mutations, receptor binding, antibodies, and temperature all affect the kinetics and energetics of this conformational change. Thus, quantitative measurements, such as in vitro binding assays, need to be re-evaluated for possible effects due to this mixed population. How easily this conformation is sampled in natural membrane-bound spike and its position in the viral lifecycle are still unknown; however, an antibody specific to this conformation can bind and neutralize in in vitro SARS-CoV-2 cell fusion and pseudovirus assays, suggesting an important role. Determining which conformation elicits more robust protective immunity against both SARS-CoV-2 variants and other coronaviruses will be important for future vaccine development.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/ s41594-022-00735-5.

Received: 4 August 2021; Accepted: 21 January 2022; Published online: 2 March 2022

References

- Baum, A. et al. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* 369, 1014–1018 (2020).
- Taylor, P. C. et al. Neutralizing monoclonal antibodies for treatment of COVID-19. Nat. Rev. Immunol. 21, 382–393 (2021).
- Hansen, J. et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* 369, 1010–1014 (2020).
- 4. Shi, R. et al. A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120–124 (2020).
- Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S. & Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* 91, eabb9983 (2020).
- Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181, 271–280.e8 (2020).
- 7. Benton, D. J. et al. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **588**, 327-330 (2020).
- Jiang, S., Hillyer, C. & Du, L. Neutralizing antibodies against SARS-CoV-2 and other human coronaviruses. *Trends Immunol.* 41, 355–359 (2020).
- Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367, 1260–1263 (2020).
- Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 181, 281–292.e6 (2020).
- 11. Berman, H. M. et al. The Protein Data Bank. Nucleic Acids Res. 28, 235-242 (2000).
- White, J. M., Delos, S. E., Brecher, M. & Schornberg, K. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit. Rev. Biochem. Mol. Biol.* 43, 189–219 (2008).
- 13. Li, F. Structure, function, and evolution of coronavirus spike proteins. *Annu Rev. Virol.* **3**, 237–261 (2016).
- Zhou, T. et al. 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 (2020).
- 15. Xiao, T. et al. A trimeric human angiotensin-converting enzyme 2 as an anti-SARS-CoV-2 agent. *Nat. Struct. Mol. Biol.* **28**, 202–209 (2021).
- Belouzard, S., Chu, V. C. & Whittaker, G. R. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc. Natl Acad. Sci. USA* **106**, 5871–5876 (2009).
- 17. Cai, Y. et al. Distinct conformational states of SARS-CoV-2 spike protein. *Science* **369**, 1586–1592 (2020).
- Walls, A. C. et al. Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc. Natl Acad. Sci. USA* 114, 11157–11162 (2017).
- 19. Lu, M. et al. Real-time conformational dynamics of SARS-CoV-2 spikes on virus particles. *Cell Host Microbe* 28, 880–891 (2020).
- Zimmerman, M. I. et al. SARS-CoV-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome. *Nat. Chem.* 13, 651–659 (2021).
- Sztain, T. et al. A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat. Chem.* 13, 963–968 (2021).
- Englander, S. W. Hydrogen exchange and mass spectrometry: a historical perspective. J. Am. Soc. Mass Spectrom. 17, 1481–1489 (2006).
- Zheng, J., Strutzenberg, T., Pascal, B. D. & Griffin, P. R. Protein dynamics and conformational changes explored by hydrogen/deuterium exchange mass spectrometry. *Curr. Opin. Struct. Biol.* 58, 305–313 (2019).
- 24. Casalino, L. et al. Beyond shielding: the roles of glycans in the SARS-CoV-2 spike protein. ACS Cent. Sci. 6, 1722-1734 (2020).
- Turoňová, B. et al. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* 370, 203–208 (2020).
- Barnes, C. O. et al. Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell* 182, 828–842(2020).
- Raghuvamsi, P. V. et al. SARS-CoV-2 S protein:ACE2 interaction reveals novel allosteric targets. *eLife* 10, e63646 (2021).
- Huang, Y. et al. Identification of a conserved neutralizing epitope present on spike proteins from all highly pathogenic coronaviruses. Preprint at bioRxiv https://doi.org/10.1101/2021.01.31.428824 (2021).
- Benhaim, M. A., Mangala Prasad, V., Garcia, N. K., Guttman, M. & Lee, K. K. Structural monitoring of a transient intermediate in the hemagglutinin fusion machinery on influenza virions. *Sci. Adv.* 6, eaaz8822 (2020).
- Edwards, R. J. et al. Cold sensitivity of the SARS-CoV-2 spike ectodomain. Nat. Struct. Mol. Biol. 28, 128–131 (2021).
- Hsieh, C.-L. et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. Science 369, 1501–1505 (2020).
- Henderson, R. et al. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* 27, 925–933 (2020).

NATURE STRUCTURAL & MOLECULAR BIOLOGY

- McCallum, M., Walls, A. C., Bowen, J. E., Corti, D., & Veesler, D. Structure-guided covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation. *Nat. Struct. Mol. Biol.* 27, 942–949 (2020).
- The CITIID-NIHR COVID-19 BioResource Collaboration et al. A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct. Mol. Biol.* 27, 934–941 (2020).
- 35. Yan, R. et al. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444–1448 (2020).
- Wang, Q. et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell 181, 894–904(2020).
- Starr, T. N. et al. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* 371, 850–854 (2021).
- Gilman, M. S. A., Furmanova-Hollenstein, P., Pascual, G., B van't Wout, A., Langedijk, J. P. M. & McLellan, J. S. Transient opening of trimeric prefusion RSV F proteins. *Nat. Commun.* 10, 2105 (2019).
- Albertini, A. A. et al. Characterization of monomeric intermediates during VSV glycoprotein structural transition. *PLoS Pathog.* 8, e1002556 (2012).
- Kim, I. S. et al. Mechanism of membrane fusion induced by vesicular stomatitis virus G protein. *Proc. Natl Acad. Sci. USA* 114, E28–E36 (2017).
- 41. Bangaru, S. et al. A site of vulnerability on the influenza virus hemagglutinin head domain trimer interface. *Cell* **177**, 1136–1152 (2019).
- 42. Eckert, D. M. & Kim, P. S. Mechanisms of viral membrane fusion and its inhibition. *Annu. Rev. Biochem.* **70**, 777–810 (2001).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

Methods

Protein expression and purification. SARS-CoV-2 spike (2P) and RBD were expressed and purified from stably transformed Expi293 cells, following methods as described⁴³. Briefly, stable Expi293 suspension cells were maintained in Expi293 medium at 37 °C. Cells were grown for 6 days, then collected by centrifugation, and filtered. Supernatant was adjusted to pH 7.4 and loaded onto a 5 mL HisTrap Excel column and washed with 60 CV of buffer. Captured proteins were eluted with 10 CV of 20 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole, pH 7.4. Eluted proteins were buffer exchanged into PBS using either 3-kDa-MWCO (molecular weight cut-off) (for RBD) or 100-kDa-MWCO (for spike) Amicon concentrators and filtered prior to storage at -80 °C. HexaPro, HexaPro-S383C/ D985C, and Alpha S1 HexaPro were expressed and purified from transiently transfected Freestyle 293-F cells as described³¹. Briefly, cell cultures were harvested 4 days after transfection and centrifuged. Supernatants were filtered and flowed over StrepTactin resin (IBA). Proteins were further purified by size-exclusion chromatography using a Superose 6 increase column (GE Healthcare) in a buffer composed of 2 mM Tris pH 8.0, 200 mM NaCl, and 0.02% NaN3. ACE2-Fc was a gift from the Wells lab44. 3A3 IgG was expressed and purified from ExpiCHO cells as described²⁸. Briefly, antibodies were expressed in ExpiCHO cells according to the high-titer protocol provided and purified on a Protein A HiTrap column (GE Healthcare) and buffer exchanged to PBS.

Continuous-hydrogen-exchange labeling. For all continuous-hydrogen-exchange experiments, deuterated buffer was prepared by lyophilizing PBS (pH 7.4, Sigma-Aldrich P4417) and resuspending it in D₂O (Sigma-Aldrich 151882). To initiate the continuous-labeling experiment, samples were diluted tenfold (final spike trimer concentration of 0.167 μ M) into temperature-equilibrated, deuterated PBS buffer (pH_{read} 7, pD 7.4). Samples were quenched at the time points outlined below, by mixing 30 μ L of the partially exchanged protein with 30 μ L of 2× quench buffer (3.6 M guanidinium chloride, 500 mM TCEP, 200 mM glycine pH 2.4) on ice. Samples were incubated on ice for 1 minute to allow for partial unfolding to assist with proteolytic degradation and then were flash frozen in liquid nitrogen and stored at $-80\,^\circ\text{C}$.

For studies comparing HexaPro \pm ACE2 the RBD in isolation vs in S-2P, purified spike (1.67 μM spike trimer or 5 μM RBD) was incubated in PBS at 25 °C overnight (12-16 hours) before the initiation of hydrogen exchange. For experiments done in the presence of ACE2-Fc, the ligand was added during this incubation at a 1.25:1 molar ratio of ligand to spike monomer (6.25 μM ligand) to ensure saturation. Based on the reported affinity (K_D ~15 nM) for ACE2-Fc, fraction bound can be assumed to be greater than 97%. The hydrogen-exchange time points for these experiments were 15 seconds, 60 seconds, 180 seconds, 600 seconds, 1800 seconds.

For the comparison of HexaPro \pm 3A3, HexaPro was incubated overnight at 37°C (12–16 hours). After incubation the protein was moved to 25°C and diluted to 1.67 μ M spike trimer. In the 3A3-bound condition, 6.25 μ M antibody was added and allowed to bind for 10 minutes at 25°C. Given the affinity of 3A3 for HexaPro (12 nM), the fraction bound can be assumed to be greater than 97%. The quench time points for this experiment were 15 seconds, 180 seconds, 1,800 seconds, and 14,400 seconds.

Back exchange control preparation. S-2P was diluted to 1.67 μ M trimer in PBS pH 7.4. To initiate hydrogen exchange, the sample was diluted tenfold (final spike trimer concentration of 0.167 μ M) into deuterated PBS buffer (pH_{read} 7, pD 7.4) that was supplemented with 3.6 M guanidinium chloride, and then incubated at 37 °C. The addition of denaturant and increased temperature should both promote hydrogen exchange, respectively. Following 2 weeks of exchange, 30 μ L of deuterated spike was mixed with 30 μ L of 2× quench buffer lacking denaturant (500 mM TCEP, 200 mM glycine pH 2.4) and kept on ice for 1 minute prior to flash freezing in liquid nitrogen and storage at -80 °C. The results of this control experiment were used to characterize the back exchange of the system and were not used to adjust deuteration values of continuous-labeling experiments.

Incubation kinetics and pulse labeling. For evaluating the temperature-dependent kinetics of interconversion, frozen spike samples were thawed, diluted to 5 μ M spike monomer, and incubated at 37 °C overnight. Samples were then moved to a temperature-controlled chamber at 4°C, and the population of each state was evaluated at the specified time points, as described below. After the final 4°C sample was taken (96–526 hours, depending on the spike construct), the sample was returned to a 37°C heat block for further incubation, and the population of each state was again evaluated at the specified time points as described below. To evaluate the relative population of the A and B conformations at each time point, 3 μ L of spike sample was removed from the incubation tube and mixed with 27 μ L of room-temperature deuterated buffer. After a 1-minute labeling pulse, 30 μ L of quench buffer kept on ice was mixed with the 30 μ L of labeled protein. Quenched samples were kept on ice for 1 minute to allow for partial unfolding, and then were flash frozen in liquid nitrogen.

For kinetics carried out in the presence of ACE2 or 3A3, after the initial 37 °C incubation, the sample was brought to 25 °C and ligand was added (6.25 μ M).

To monitor the population of state A and B as a function of time at 25 °C in the presence of ACE2 the sample was kept in a temperature-controlled chamber at 25 °C, and aliquots were removed for pulse labeling as described above at 0 hours, 30 minutes, 3 hours, 6 hours, and 24 hours.

Protease digestion and LC-MS. All samples were thawed immediately before injection into a cooled valve system (Trajan LEAP) coupled to a LC (Thermo UltiMate 3000). Sample time points were injected in random order. The temperature of the valve chamber, trap column, and analytical column were maintained at 2 °C. The temperature of the protease column was maintained at 10 °C. The quenched sample was subjected to inline digestion by two immobilized acid proteases in order, aspergillopepsin (Sigma-Aldrich P2143) and porcine pepsin (Sigma-Aldrich P6887), at a flow rate of 200 µL/minute of buffer A (0.1% formic acid). Protease columns were prepared in house by coupling protease to beads (Thermo Scientific POROS 20 Ål aldehyde activated resin 1602906) and packed by hand into a column (2 mm ID × 2 cm, IDEX C-130B). Following digestion, peptides were desalted for 4 minutes on a hand-packed trap column (Thermo Scientific POROS R2 reversed-phase resin 1112906, 1 mm ID × 2 cm, IDEX C-128). Peptides were then separated with a C8 analytical column (Thermo Scientific BioBasic-8 5- μ m particle size 0.5 mm ID × 50 mm 72205-050565) and a gradient of 5-40% buffer B (100% acetonitrile, 0.1% formic acid) at a flow rate of 40 μ L/min over 14 minutes, and then of 40–90% buffer B over 30 seconds. The analytical and trap columns were then subjected to a sawtooth wash and equilibrated at 5% buffer B prior to the next injection. Protease columns were washed with two injections of 100 µL 1.6 M GdmCl, 0.1% formic acid prior to the next injection. Peptides were eluted directly into a Q Exactive Orbitrap Mass Spectrometer operating in positive mode (resolution 70000, AGC target 3×10^6 , maximum IT 50 ms, scan range 300-1,500 m/z). For each spike construct, a tandem mass spectrometry experiment was performed (Full MS settings the same as above, dd-MS² settings as follows: resolution 17,500, AGC target 2 × 10⁵, maximum IT 100 ms, loop count 10, isolation window 2.0 m/z, NCE 28, charge state 1 and ≥ 7 excluded, dynamic exclusion of 15 seconds) on undeuterated samples. LC and MS methods were run using Xcalibur 4.1 (Thermo Scientific)

Peptide identification. Byonic (Protein Metrics) was used to identify unmodified and glycosylated peptides in the tandem MS data. The sequence of the expressed construct, including signal sequence and the trimerization domain, was used as the search library. Sample digestion parameters were set to non-specific. Precursor mass tolerance and fragment mass tolerance were set to 6 and 10 ppm, respectively. Variable N-linked glycosylation was allowed, with a library of 132 human N-glycans used in the search. Peptide lists (sequence, charge state, and retention time) were exported from Byonic and imported into HDExaminer 3 (Sierra Analytics). When multiple peptide lists were obtained, all were imported and combined in HDExaminer 3.

HDExaminer 3 analysis. Peptide isotope distributions at each exchange time point were fit in HDExaminer 3. For glycosylated peptides, only the highest-confidence modification was included in the mass spectra search and analysis. For unimodal peptides, deuteration levels were determined by subtracting mass centroids of deuterated peptides from undeuterated peptides. For bimodal peaks, extracted peptide isotope spectra were exported from HDExaminer 3 and analyzed separately (see below for details).

Bimodal fitting and conformation quantification. Peptide mass spectra for bimodal peptides were exported from HDExaminer 3. All quantitative analyses of the exported peptide mass spectra were performed using Python scripts in Jupyter notebooks. After importing a peptide mass spectra, the *m*/z range containing all possible deuteration states of the selected peptide was isolated and the find_peaks method from the scipy.signal package was used to identify each isotope in the mass envelope, and the height of each peak was used as its intensity. The area of the total mass envelope was normalized to account for run-to-run differences in intensity. The bimodal mass envelopes for all time points under the same condition were globally fit to a sum of two gaussians, keeping the center and width of each gaussian constant across all incubation time points. Fitting was done using the curve_fit function from the scipy.optimize package. After fitting, the area under each individual gaussian was determined to approximate the relative population of each state.

Size-exclusion chromatography–multiangle light scattering. To generate variable populations of states A and B, S-2P (100 μ L, 0.18 mg/mL) was incubated at either 4 °C or 37 °C for 4 days. Post-incubation, and prior to size-exclusion chromatography–multiangle light scattering (SEC–MALS) injection, an aliquot from each sample was taken and a 1-minute pulse-labeling experiment was performed (see above). The resulting bimodal peptide distributions were used to calculate the fraction of state A and B. Samples were filtered (0.22 μ m hydrophilic polyvinylidene fluoride Ultrafree-MC GV centrifugal filter) prior to SEC–MALS injection. Then, 90 μ L of sample (0.18 mg/mL) was injected onto a Superose 6 increase 10/300 (GE Healthcare) preequilibrated in filtered (0.1- μ m polyethersulfone Nalgene Rapid-Flow vacuum filter) PBS flowing at 0.5 mL/min at

NATURE STRUCTURAL & MOLECULAR BIOLOGY

4 °C. The ÄKTA Pure 25 M1 (Cytiva) chromatography system was coupled to an 18-angle light scattering Wyatt Dawn detector and Wyatt Optilab refractive index detector (Wyatt Technology). Data analysis was carried out using the program ASTRA 7.1.4.8.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Figures using protein structures were generated using either PDB accession no. 6M0J or a full-length structural model of the spike protein created by the Amaro lab and accessible at amarolab.ucsd.edu/covid19.php (PSF-PDB_spike_open_prot_ glyc_memb_wat_ions_amarolab.tar.gz). Centroid values for continuous-labeling hydrogen-exchange experiments are included as supplementary .csv files. Select peptide spectra used for bimodal analysis are available in raw and extracted form as supplementary .csv files. Raw mass spectrometry data are available upon request. Source data are provided with this paper.

References

- Byrum, J. R. et al. multiSero: open multiplex-ELISA platform for analyzing antibody responses to SARS-CoV-2 infection. Preprint at medRxiv https://doi.org/10.1101/2021.05.07.21249238 (2021).
- Glasgow, A. et al. Engineered ACE2 receptor traps potently neutralize SARS-CoV-2. Proc. Natl Acad. Sci. USA 117, 28046–28055 (2020).

Acknowledgements

We thank the entire Marqusee Lab for advice and critical reading of the manuscript, particularly J. Lindner for help initiating these experiments. We thank P. S. Kim, A. Powell, P. Weidenbacher, and N. Freedland (Stanford University) for advice and discussion; J. Wells, S. Lim, and I. Lui (University of California, San Francisco) for discussion and providing ACE2-Fc; Y. Cheng, A. Mancino, and M. Jin (University of California, San Francisco) for assistance with SEC–MALS; and A. Manglik (University of California, San Francisco) for advice and discussion. We thank J. Morrow and Sierra Analytics for making HDExaminer freely available for at-home use during shelter-in-place orders. S. M. is a Chan Zuckerberg Biohub Investigator. Some figures were created with the use of BioRender.com. This work was funded by NSF Graduate Research Fellowship DGE1752814 (S. M. C.); NIH grant no. GM050945 (S. M.); and NSF grant no. MCB1616591 (S. M.); NIH grant no. AI127521 (J. S. M.); Welch foundation grant no. F-0003-19620604 (J. S. M.); NIH grant no. AI122753 (J. A. M.); Welch foundation grant no. F-1767 (J. A. M.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

Conceptualization: S. M. C., S. R. S., H. T. H., S. M. Methodology: S. M. C., S. R. S., H. T. H., J. E. P., S. M. Investigation: S. M. C., S. R. S., H. T. H., J. E. P., A. W. N., C. L. H. Writing: S. M. C., S. R. S., H. T. H., A. W. N., C. L. H., J. A. M., J. C. M., J. E. P., S. M.

Competing interests

S. M. C., S. R. S., and S. M. are inventors on U.S. patent application no. 63/220,388, (Methods related to an alternative conformation of the SARS-CoV-2 Spike Protein). A. W. N., C.-L. H., J. S. M., and J. A. M. are inventors on U.S. patent application no. 63/135,913 (Cross-reactive antibodies recognizing the coronavirus spike S2 domain). J. S. M. is an inventor on U.S. patent application no. 62/412,703 (Prefusion coronavirus spike proteins and their use). C.-L. H., J. A. M. and J. S. M. are inventors on U.S. patent application no. 63/032,502 (Engineered coronavirus spike (S) protein and methods of use thereof).

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-022-00735-5.

Correspondence and requests for materials should be addressed to Susan Marqusee. **Peer review information** *Nature Structural and Molecular Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Beth Moorefield was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

nature portfolio

Corresponding author(s): Susan Marqusee

Last updated by author(s): Dec 30, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our <u>Editorial Policies</u> and the <u>Editorial Policy Checklist</u>.

Statistics

For	all st	atistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.
n/a	Cor	firmed
	\boxtimes	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
	\square	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
\boxtimes		The statistical test(s) used AND whether they are one- or two-sided Only common tests should be described solely by name; describe more complex techniques in the Methods section.
\boxtimes		A description of all covariates tested
\boxtimes		A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
\boxtimes		A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
\boxtimes		For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted Give P values as exact values whenever suitable.
\boxtimes		For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
\boxtimes		For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
\boxtimes		Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
		Our web collection on <u>statistics for biologists</u> contains articles on many of the points above.

Software and code

Policy information about availability of computer code		
Data collection	All mass spectrometry data collection was performed using Xcalibur 4.1 (Thermo Scientific)	
Data analysis	Peptide identifications were performed using Byonic 3.6.0 (Protein Metrics Inc.) Peptide deuterium uptake was calculated using HDExaminer 3.2 (Sierra Analytics). Bimodal peptide analysis and plotting were performed in Jupyter notebooks using Python.	

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about <u>availability of data</u>

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Figures using protein structures were generated using either PDB accession no. 6M0J or a full-length structural model of the spike protein created by the Amaro lab and accessible at amarolab.ucsd.edu/covid19.php (PSF-PDB_spike_open_prot_glyc_memb_wat_ions_amarolab.tar.gz). Centroid values for continuous labeling hydrogen exchange experiments are included as supplementary csv files. Select peptide spectra used for bimodal analysis are available in raw and extracted form as supplementary csv files. Raw mass spectrometry data is available upon request.

Field-specific reporting

K Life sciences

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see <u>nature.com/documents/nr-reporting-summary-flat.pdf</u>

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Continuous labeling hydrogen exchange experiments were performed in technical triplicate. Pulsed labeling time courses were single replicates except where noted in table 1.
Data exclusions	During manual curation of data in HDExaminer, low quality peptides (peptides with low signal intensity or peptides that have inconsistent monoisotopic masses) are removed. To be included in the manuscript analysis of continuous labeling experiments, peptides must have triplicate data for each time point, and the deuterium uptake at each time point must have a standard deviation below 1 deuteron.
Replication	Continuous labeling hydrogen exchange experiments were performed in technical triplicate with standard deviations reported in the supplementary data and included as error bars where applicable. Pulsed labeling time courses were single replicates except where noted in table 1. The agreement seen between the replicates indicates in table 1 was used as evidence of technical reproducibility for temperature incubation time courses, allowing us to forgo replicates for the remaining time courses and conserve sample. Additionally, for pulsed-labeling experiments the data for two bimodal peptides that were non-overlapping in sequence were combined and the time course was fit to a single exponential, resulting in the rate constants reported in table 1. The agreement between peptides from different regions provided additional confidence that additional time-course replicates were not necessary.
Randomization	For every condition (e.g. S-2P in the presence of ACE2) all time points (i.e. 15 s, 60 s, 180 s, 600 s, 1800 s, 5400 s, and 14400 s of deuteration exposure for continuous labeling experiments, and various incubation times for pulsed labeling experiments) were injected in random order to ensure that observations were not the result of instrument variation.
Blinding	No blinding was performed for this study.

Reporting for specific materials, systems and methods

Methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

, , , , , , , , , , , , , , , , , , , ,		
Involved in the study	n/a	Involved in the study
X Antibodies	\boxtimes	ChIP-seq
Eukaryotic cell lines	\boxtimes	Flow cytometry
Palaeontology and archaeology	\boxtimes	MRI-based neuroimaging
Animals and other organisms		
Human research participants		
Clinical data		
Dual use research of concern		
	Involved in the study Antibodies Eukaryotic cell lines Palaeontology and archaeology Animals and other organisms Human research participants Clinical data Dual use research of concern	Involved in the study n/a Involved in the study n/a Antibodies Image: State St

Antibodies

Antibodies used	3A3 (non-commercial Antibody, produced by Maynard Lab)
Validation	Described in detail in Y. Huang, A. W. Nguyen, CL. Hsieh, R. Silva, O. S. Olaluwoye, R. E. Wilen, T. S. Kaoud, L. R. Azouz, A. N. Qerqez, K. C. Le, A. L. Bohanon, A. M. DiVenere, Y. Liu, A. G. Lee, D. Amengor, K. N. Dalby, S. D'Arcy, J. S. McLellan, J. A. Maynard, Identification of a conserved neutralizing epitope present on spike proteins from all highly pathogenic coronaviruses. bioRxiv (2021), p. 2021.01.31.428824, , doi:10.1101/2021.01.31.428824.

March 2021

nature portfolio | reporting summary

Eukaryotic cell lines

Policy information about <u>cell lines</u>	
Cell line source(s)	Expi293F (ThermoFisher cat. A14527) FreeStyle-293F cells (Invitrogen)
Authentication	Cell lines were purchased commercially and were not further validated.
Mycoplasma contamination	Expi293F cells were not tested. FreeStyle-293F cells (Invitrogen) have tested negative for mycoplasma contamination.
Commonly misidentified lines (See <u>ICLAC</u> register)	Name any commonly misidentified cell lines used in the study and provide a rationale for their use.