Title: Engineered Immunogens to Expose Conserved Epitopes Targeted by Broad Coronavirus Antibodies

Authors: Brenda Kapingidza[1,2,#], Daniel J. Marston[1,2,#], Caitlin Harris[1,2], Daniel Wrapp[1,2], Kaitlyn Winters[1,4], Brianna Rhodes[1,2], Pranay Vure[1,4], Christopher W. Woods[1,2,5], Elizabeth A. Petzold[5], Emmanuel B Walter[1,6], Rob Parks[1,2], Maggie Barr[1,2], Qi Yin[1], Derek W. Cain[1,2], Kevin Wiehe[1,2], Kevin O. Saunders[1,3,7,8], Barton F. Haynes[1,2,3], Mihai L. Azoitei[1,2,*]

Affiliations:

[1]Duke Human Vaccine Institute, Duke University, Durham, NC

[2]Department of Medicine, Duke University, Durham, NC

[3]Department of Immunology, Duke University, Durham, NC

[4]Duke University, Durham, NC

[5]Center for Infectious Diseases and Diagnostic Innovation, Duke University Medical Center, Durham, NC 27710, USA.

[6] Department of Pediatrics, Duke University, Durham NC

[7]Department of Surgery, Duke University, Durham, NC

[8]Department of Molecular Genetics and Microbiology, Duke University, Durham, NC

[#]These authors contributed equally to this work.

* Corresponding author. Email: mihai.azoitei@duke.edu

Key Words: SARS-CoV-2, vaccine, epitope-scaffolds, stem helix, fusion peptide, pan-coronavirus vaccine

Abstract:

Immune responses to SARS-CoV-2 primarily target the receptor binding domain of the spike protein, which can readily mutate to escape acquired immunity. Other regions in the spike S2 subunit, such as the fusion peptide and the stem helix, are highly conserved across sarbecoviruses and recognized by broadly reactive antibodies, providing hope that targeting these epitopes by vaccination could offer protection against both current and emergent viruses. Here we employed computational modeling to design epitope scaffolds that display the fusion peptide and the stem helix epitopes. The engineered proteins bound both mature and germline versions of multiple broad and protective human antibodies with high affinity. Binding specificity was confirmed both biochemically and via high resolution crystal structures. Finally, the epitope scaffolds showed potent engagement of antibodies and memory B-cells from subjects previously exposed to SARS-CoV2, illustrating their potential to elicit antibodies against the fusion peptide and the stem helix by vaccination.

Main Text:

The majority of immune responses elicited against SARS-CoV-2, by either natural infection, vaccination, or a combination of both, are focused on the receptor binding domain (RBD)[1-4]. However, deep scanning mutagenesis studies[5] and the emergence of novel variants that evade pre-existing immunity have revealed the plastic nature of RBD[2,6], suggesting that next-generation coronavirus vaccines that aim to protect against both current and emergent CoV will likely need to induce antibodies against CoV spike regions beyond the RBD.

Two regions in the S2 subunit of the SARS-CoV-2 spike, the fusion peptide (FP) and the stem helix, show high conservation across diverse coronaviruses. Antibodies isolated against these sites cross-react with alpha and beta human coronaviruses and with diverse animal viruses[7-10]. Fusion peptide antibodies, such as DH1058, DH1294, VN01H1, Cov44-79, and Cov44-62, bind to SARS-CoV-2 spike residues 813-824, cross-react with all seven human CoVs, and protect by either neutralization, Fc mediated mechanisms or both. Two major classes of antibodies with overlapping epitopes have been identified against the stem helix region[9-12]. S2P6 and CC40.8, which are representative of these two classes, bound spike glycoproteins representative of all sarbecovirus clades, and protected against challenge in animals by inhibiting S-mediated membrane fusion. However, humoral responses against the stem helix or the FP are not robustly induced by existing vaccines or by natural infection. This is likely due to a combination of factors, including the occlusion of the stem helix and fusion peptide on the pre-fusion spike, requiring ACE2 binding for exposure[8], as well as the presence of the prominently displayed, immunodominant RBD domain. Given their ability to recognize diverse coronaviruses, immunogens that induce strong responses against the FP and stem helix regions could be a key component of a future “pan-coronavirus” vaccine offering broad protection against both currently circulating and emergent coronaviruses.

For the design of immunogens that expose occluded epitopes in novel molecular contexts, we and others previously described the development of “epitope scaffolds” proteins[13-15]. These molecules are engineered by transplanting the structure of the antibody-bound epitopes from viral molecules onto unrelated protein scaffolds. Unlike peptide-based immunogens that can sample diverse epitope conformations, epitope scaffolds present only the antibody-bound conformation of the target epitope on their surface, which typically leads to higher antibody affinity and the elicitation of antibodies specific for the structure of the target epitope[16,17]. Here, we applied epitope grafting to design stem helix or FP protein scaffolds that strongly interact with broadly cross-reactive antibodies against the fusion peptide or the stem helix regions of spike, with the eventual goal of inducing such antibodies by vaccination with these new immunogens.

Results

Design of epitope scaffolds that bind broadly cross-reactive antibodies against the fusion peptide.

The FP domain of the SARS-CoV-2 encompasses residues 808-833 (Figure 1a). Multiple FP antibodies, such as DH1058, DH1294, VN01H1, Cov44-79, and Cov44-62 target epitopes centered on residues 813-824, with key contacts made with virus residues R815, E819 and F823 (SARS-CoV-2 numbering, **Figure 1a**, **Supp. Fig. 1**)[7,8,18,19]. These residues are occluded in the structure of the prefusion SARS-CoV-2 spike, likely limiting their immune recognition (**Figure 1a**). To improve the accessibility and immune recognition of this epitope, we developed fusion peptide epitope scaffolds (ES) using “side chain” grafting computational methods that we and others previously described (**Supp. Fig. 2**)[13-15]. A large library (~10,000) of small monomeric scaffolding proteins was queried computationally to identify proteins with exposed backbone regions that closely matched (<0.5Å RMS) the structure of the FP epitope

residues 815-823 (**Supp. Fig. 2**) In regions with high structural mimicry to the DH1058-bound epitope, the epitope sequence replaced that of the parent scaffold and additional mutations were introduced to accommodate the grafted sequence (**Supp. Fig. 2**).

Fifteen engineered FP ESs were expressed recombinantly (**Figure 1b** for representative designs) with six out of 15 FP ESs producing soluble and stable proteins that bound DH1058 by ELISA (**Supp. Fig. 3**, **Supp. Table. 1**). DH1058 mAb dissociation constants were determined by surface plasmon resonance (SPR) for all six ESs, and to DH1294 mAb for the three highest DH1058 mAb binders (**Figure 1c; Supp. Fig. 4**). Designs FP-2, FP-10, and FP-15 bound to DH1058 mAb with picomolar affinity and to DH1294 mAb with K_Ds of 1221nM, 0.1nM and 219nM respectively. To further probe the breadth of their recognition, the binding of FP ESs was measured against fusion peptide antibodies VN01H1, COV44-79, and COV44-62, which were not explicitly considered at the design stage (**Figure 1d; Supp. Fig. 5**). Overall, designs FP-2, FP-10, FP-10 and FP-15 showed binding to all three antibodies at levels similar or higher than those exhibited by the synthesized SARS-CoV-2 fusion peptide. All the ESs bound at least two of the three mAbs, underscoring their ability to recognize diverse FP antibodies and providing evidence that these engineered scaffolds are reliably presenting the immunologically relevant confirmation of the fusion peptide.

DH1058 mAb bound four out of the six FP ESs with affinities 100-fold higher than that measured here for the synthesized fusion peptide (K_D=66.1nM), and those previously reported for different SARS-CoV-2 spike proteins (K_Ds of 99nM for WA-2 and 106nM for Omicron)[18]. To ensure that the ES affinity gains are not due to contacts between the scaffold and the mAb that are not present in the native DH1058-epitope complex, interface hotspot residues corresponding to fusion peptide amino acids R815, E819, and F823 were mutated to alanine in designs FP-2, FP-10, and FP-15. The alanine substitutions eliminated DH1058 binding to the synthetic fusion peptide and to designs FP-2 and FP-10. Two of the three substitutions similarly

eliminated DH1058 binding to FP-15, with the F823A equivalent retaining some low-level antibody binding (**Figure 1e**).

A high resolution (2.2Å) crystal structure of FP-15 in complex with DH1058 revealed the binding mode and validated the computational modeling (**Figure 1f, Supp. Table 2**). The structure was in close agreement with that of the computationally generated antibody-ES complex, both overall and over the epitope scaffold only (RMSD<1Å). The epitope conformation engaged by DH1058 on FP-15 was essentially identical to that induced by the antibody on spike (RMSD=0.3Å)[18], confirming that the antibody interacts with the epitope presented by the ES in the same manner it binds to its natural target.

Design of epitope scaffolds to engage S2P6-like antibodies against the stem helix region of the CoV spike.

Next, we engineered epitope scaffolds to engage antibodies against the stem helix region of the CoV spike. Antibodies such as S2P6, DH1057.1, and CC40.8 bind to an epitope located between residues 1144 and 1158, which is almost completely occluded by the trimerization interface in the pre-fusion spike structure (**Figure 2a**)[9,10,20]. While both CC40.8 and S2P6 target the same spike region and make critical contacts with residues Phe1148 and Phe1156, CC40.8 makes additional interactions with Leu1145 and induces a more extended epitope conformation upon binding (**Figure 2a**)[10]. Because the CC40.8 epitope is longer and includes the S2P6 epitope, we first attempted to engineer epitope scaffolds that display the antibody-bound conformation of the CC40.8 epitope, with the expectation that these molecules would also be capable of binding to S2P6-class antibodies. However, the “side chain” grafting protocol used for FP epitope scaffolds above, did not identify any candidate scaffolds with exposed backbones similar to that of the CC40.8 epitope, likely due to the unusual conformation of this epitope that adopts a distorted alpha-helical shape near the N-terminus (**Figure 2a**). In

contrast, multiple candidate epitope scaffolds that displayed the S2P6-bound conformation of the epitope segment 1148-1156 were successfully designed computationally using side chain grafting. This was possible because the S2P6 epitope adopts a canonical alpha-helical structure, with most of the antibody contacting residues contained within three helical turns (**Figure 2a**).

Fifteen S2P6-targeted designs, named S2hlx-1 to 15, were chosen for recombinant production in *E. coli* (**Figure 2a** for representative designs; **Supp. Table 1**) and seven yielded soluble protein. In an initial ELISA screen, all seven designs bound both S2P6 and DH1057.1 mAbs comparable to or better than a synthesized SARS-CoV-2 stem helix peptide (residues 1147-1161). As expected, neither this peptide nor the ESs bound CC40.8 mAb due to the absence of epitope residue L1145 (**Supp. Fig. 6**). Upon SPR analysis, S2hlx ESs were measured to bind S2P6 mAb with dissociation constants between 0.2nM and 83nM and DH1057.1 mAb with K_Ds between 0.6nM and 4.5uM (**Figure 2b, c; Supp. Fig. 7**). S2P6 mAb was previously reported to bind spike proteins with dissociation constants of 7nM, for those of SARS-CoV-1 and SARS-CoV-2, and 12nM for the MERS-CoV spike[9]. S2hlx-7 and S2hlx-15 bound S2P6 mAb with substantially higher affinities (K_Ds of 0.2nM and 0.7nM), while the other five designs had lower affinities than native spikes (K_Ds between 47nM and 83nM). S2hlx-7 and S2hlx-15 also showed sub-nanomolar binding to DH1057.1 mAb, while other designs had weaker binding, in the micromolar range, for this antibody (**Figure 2b**). In addition to the mature mAbs, binding was also tested to their inferred unmutated common ancestors (UCAs) or germline-reverted forms (iGLs). While all the S2hlx ESs interacted with DH1057UCA, S2hlx-15 displayed the highest affinity (K_Ds=1.1nM) (**Figure 2b**). For S2P6-iGL, S2hlx-4 is the only design that had measurable binding by ELISA, although its interactions were reduced >10-fold compared to those of the affinity-matured antibody, likely due to a significantly faster off-rate as measured by SPR (**Supp. Fig. 7**, **Supp. Fig. 8**).

The binding specificity of S2hlx-4, S2hlx-7 and S2hlx-15 was confirmed by mutating epitope residues equivalent to spike Phe1148, Leu1152 and Phe1156. Alanine mutations limited the binding to both S2P6 and DH1057.1 mAbs in S2hlx-4 and -15 to a similar extent to what as observed for synthesized stem helix peptides (**Figure 2d**)[9]. For S2hlx-7, the individual epitope mutations greatly decreased DH1057.1 mAb interactions but had almost no effect on S2P6 mAb recognition. However, an S2hlx-7 variant that contained alanine substitutions at all three epitope positions completely abrogated binding, thus confirming that antibody interactions were mediated by the grafted epitope residues. In addition, none of the parent scaffolds these designs were based on showed binding to S2P6 or DH1057, further demonstrating epitope specificity (**Supp. Fig. 9**).

Crystallographic analysis of the high resolution (1.9Å) S2hlx-7/DH1057.1 complex revealed for the first time the structure and interaction mode of the DH1057.1 antibody with the stem helix. DH1057.1 approached the epitope at the same angle as S2P6 mAb and similarly interacted with residues Phe1148, Leu1152, and Phe1156 (SARS-CoV-2 spike numbering) (**Figure 2e**), as predicted by the alanine mutations (**Figure 2d**). Unlike S2P6, DH1057.1 interacts with residue K1149 through salt bridges mediated by heavy chain residues Asp95 and Asp100C (**Figure 2e**). The structure of S2hlx-7 in the complex was closely aligned with that of the computational model both overall (RMSD=0.5Å) and, most importantly, over the grafted epitope region (RMSD=0.25Å) (**Figure 2e**). The epitope conformation displayed on S2hlx-7 closely matched that of the antibody-bound S2P6 epitope and that of the corresponding region on the prefusion spike (RMSD<0.2Å, **Figure 2e**). These data confirmed the design of multiple protein antigens that have high affinity and specificity for S2P6-like mAbs by mimicking the conformation of their epitope on native spikes.

Design of epitope scaffolds to engage the two major classes of antibodies against the stem helix.

Given the failure of side chain grafting to produce epitope scaffolds that engage the CC40.8 mAb, we employed an alternative design technique, termed backbone grafting. Here, rather than transplanting epitope side chains onto a preexisting scaffold backbone, the parent scaffold backbone is replaced with the backbone of the desired epitope. This allows the grafting of more structurally complex epitopes that have no existing structural match, such as the CC40.8 S2-helix epitope. Candidate scaffolds were identified by aligning the N- and C-termini of the stem helix epitope in its CC40.8-bound conformation with the N- and C-termini of possible insertion sites on a candidate scaffold (**Figure 3a, b**). For scaffolds where the RMSD of this alignment was below 0.75Å, the epitope replaced the overlapping backbone of the parent scaffold, and additional mutations were introduced to integrate the epitope into the scaffold and to ensure productive interactions with CC40.8 mAb (**Figure 3a**). The structures of candidate designs were predicted using Alphafold2[21], and additional mutations were introduced to ensure proper epitope conformation. Final designs contained between 16 to 32 mutations relative to the parent scaffolds, and the grafted epitope was 3.1 to 6.7Å away by RMSD from the backbone it replaced.

Six ESs, named S2hlx-Ex1 to 6 (**Figure 3b** for representative models; **Supp. Table 1**) were tested for bacterial expression and analyzed for binding to S2P6 mAb in an initial screen. While four of the six ESs, S2hlx-Ex2, -Ex3, -Ex4 and -Ex6, bound to S2P6, three of them had poor expression (S2hlx-Ex2, -Ex4 and -Ex6) (**Supp. Fig. 10**). To improve expression, we tested different protein expression tags and evaluated structural homologs of the parent scaffolds as alternative design templates with limited success (**Supp. Figs. 11, 12**). Next, we optimized the sequence of the S2hlx-Ex designs using ProteinMPNN[22], a recently described deep learning algorithm that takes a protein structure as input and generates amino acid sequences that are predicted to generate the same fold. Three backbone grafting designs, S2hlx-Ex2, -Ex4, and -Ex6 were optimized with this approach, and sequences that expressed with high yield (40mg/L

of culture on average, **Supp. Fig. 13**) were identified for all of them, albeit with variable success rates. For a given structure, ProteinMPNN generated designs that had highly different sequences outside the epitope. For example, S2hlx-Ex2 based designs -Ex15, -Ex17, -Ex19, -Ex20 had between 42% and 55% sequence identity with the S2hlx-Ex2 template (**Figure 3c**), and only 58-75% identity with each other, yet all bound to the stem helix mAbs by ELISA (**Figure 3d**).

The binding affinities for the designs based on the backbone structures of S2hlx-Ex2 (S2hlx-Ex15, -Ex17, -Ex19, and -Ex20), S2hlx-Ex4 (S2hlx-Ex8-Trx and S2hlx-Ex25), and S2hlx-Ex6 (S2hlx-Ex34 and -Ex35) to CC40.8, S2P6, and DH1057.1 mAbs were determined by SPR. S2hlx-Ex2 derived ESs bound S2P6 mAb with dissociation constants between 2nM and 30nM, DH1057.1 mAb with K_Ds between 2nM and 49nM, and CC40.8 mAb with K_Ds between 3nM and 27nM (**Figure 3e; Supp. Fig. 14**). S2hlx-Ex15 and -Ex19 had the highest binding affinities, in a similar range to those reported for the stem helix peptide[9,10]. Although S2hlx-Ex25 and Ex8-Trx showed similar binding to the mature antibody panel by ELISA, the calculated binding affinities were substantially lower, with only S2P6 binding in the sub-micromolar range (611.7nM and 330nM respectively, **Supp. Fig. 15**). Finally, S2hlx-Ex34, derived from S2hlx-Ex6, bound to S2P6, DH1057.1, and CC40.8 at 82nM, 180nM and 486nM respectively (**Supp. Fig. 15**). S2hlx-Ex15 and -Ex19 binding was also tested against broadly reactive stem helix antibodies Cov89-22 and Cov30-14, which were recently discovered and were not explicitly targeted by the design[23]. Both ESs bound tightly by ELISA (**Supp. Fig. 16**), indicative of their ability to recognize diverse antibodies that target the stem helix epitope. In addition to the mature mAbs, S2hlx-Ex15 and -Ex19 binding was also tested to the inferred precursors of S2P6, DH1057.1, CC40.8 and Cov89-22, and compared to the interactions of the stem helix peptide with the same antibodies by ELISA. Interestingly, when the peptide was absorbed directly onto the plate, strong binding of the mature antibodies was detected as expected, but no interactions were

measured with the precursors (**Supp. Fig. 16**). Binding between the peptide and the precursor antibodies was detected when the biotinylated peptide was captured onto streptavidin-coated wells, suggesting that epitope presentation plays a critical role in precursor recognition. S2hlx-Ex15 bound at least an order of magnitude tighter to the germline antibodies than the strep-captured peptide by ELISA. Kinetic analysis revealed that S2hlx-Ex15 bound to the UCA for DH1057, and the iGLs for S2P6 and CC40.8. with low nanomolar affinities (K_Ds of 22.4nM 1.4nM and 9.7nM, respectively), and all the related S2hlx-Ex2 derived designs showed measurable interactions (**Supp. Fig. 14**).

To ensure that the S2hlx ESs interacted with the target mAbs in a similar fashion to the epitope on CoV spike, we analyzed the binding of S2hlx-Ex15 and -Ex19; -Ex8 and -Ex25; and -Ex34 and -Ex35 variants containing alanine mutations at key epitope residues corresponding to spike sites Leu1145 and Phe1148 in the stem helix (**Figure 3f**). As predicted based on existing structural and mutagenesis data, all the ESs lost binding to CC40.8 when the Leu1145 equivalent residue was mutated, while binding to S2P6 and DH1057.1 was not affected. When the epitope residue equivalent to spike Phe1148 was mutated, binding was lost to DH1057.1 and S2P6 in all designs.

Analysis of a high resolution (2.3Å) crystal structure of unbound S2hlx-Ex19 confirmed the successful transplantation of the CC40.8 mAb epitope (**Supplemental Table 2**). As expected, the grafted epitope region in S2hlx-Ex19 diverged considerably from the backbone region it replaced in the parent scaffold (RMSD=5.8Å), with the epitope adopting a more exposed conformation that allows for antibody engagement (**Figure 3g**). This was likely facilitated by the design of two aromatic residues (Tyr62 and Phe66) under the epitope helix, which effectively pushed the epitope segment away from the rest of the protein core. Interestingly, the epitope conformation on S2hlx-Ex19 was almost identical to that present on the prefusion stem helix, rather than the more extended one induced by CC40.8 (**Figure 3g**).

Nevertheless, given the high affinity of S2hlx-Ex19 for CC40.8 (K_D=6.7nM), it is likely that the antibody can readily induce the necessary conformation upon binding. Thus, by combining epitope backbone grafting and ProteinMPNN, we engineered antigens that bind with high affinity and specificity to the major classes of stem helix broadly reactive antibodies as well as to their inferred precursors.

Cross-reactivity of engineered epitope scaffolds with pre-existing immune responses induced by SARS-CoV-2 spike.

Next, we tested the ability of the engineered epitope scaffolds to cross-react with pre-existing immune responses elicited by SARS-CoV-2 spike in order to evaluate their potential to preferentially amplify fusion peptide or stem helix directed responses by vaccination. Sera or plasma samples isolated from three groups of 12 study participants who acquired SARS-CoV-2 immunity either through vaccination, infection, or vaccination followed by infection respectively, were first tested for binding to WA-2 spike, RBD and synthesized peptides encoding the fusion peptide and stem helix domains (see **Supp. Table 3** for full patient information). This analysis revealed that only a small fraction of spike-directed humoral responses, regardless of their acquisition route, targeted the fusion peptide or stem helix regions (**Figure 4a**). On average, fusion and stem helix peptide binding levels were 23.3 and 8.4-fold lower respectively than RBD-focused responses across all groups, underscoring that the fusion peptide and the stem helix domains are sub-dominant targets of antibodies elicited by spike.

Both fusion peptide and S2hlx-Ex epitope scaffolds showed significantly higher recognition of human sera than the corresponding spike peptide domains (**Figure 4a, Supp. Fig. 17)**. From the FP ES designs, FP-15 had the best sera recognition, showing 11.3, 7.7, and 5.8-fold higher IgG binding than the synthesized fusion peptide in vaccinated, infected and hybrid immunity samples respectively (**Figure 4b)**. Importantly, these interactions were

mediated by the engineered epitope contacts, as binding to the parent scaffold of FP-15 that does not contain the grafted epitope, was significantly reduced (3.5 -fold in vaccinated, 2.5-fold in infected and 5.5-fold in hybrid, **Figure 4b**). A similar trend was observed for FP-10 and FP-2, but their overall binding levels were lower than FP-15, although still above those of the synthesized fusion peptide.

From the stem helix designs, S2hlx-Ex19 and S2hlx-Ex15 had the best recognition of spike-induced immune responses **(Supp. Fig 17)**. Compared to the synthesized stem helix peptide, S2hlx-Ex19 bound on average 2.4 times better across all three groups (**Figure 4c)**. The binding was epitope specific, as evidenced by the decrease in interactions with S2hlx-Ex19 variants that contained alanine mutations at hotspot epitope residues Leu1145 (LA) and Phe1148 (FA), or that had the whole epitope replaced by the sequence of the parent scaffold (S2hlx-Ex19 PShlx, **Figure 4c**). As shown by the larger decrease in binding to the FA than LA mutants, cross-reactive spike responses were more dependent on stem helix residue Phe1148 which is recognized by both S2P6 and CC40.8 classes of antibodies (**Figure 4d, Supp. Fig 17).** Compared to S2hlx designs, the ability of the S2hlx-Ex epitope scaffolds to recognize CC40.8 mAb translated into a broader engagement of spike elicited humoral responses for these types of molecules (**Figure 4a, Supp. Fig 17)**. We further analyzed by Fluorescence Activated Cell Sorting (FACS) the B cell memory responses of the two subjects with hybrid immunity and that showed the highest titers of S2hlx-Ex19 reactive antibodies (**Supp. Fig 18**). Memory B cells that were specific for the grafted epitope were isolated at a frequency of ~1:3000 and ~1:2100 respectively from the two samples (**Figure 4e, f**). Taken together, the robust sera binding and memory B cell engagement by epitope scaffolds demonstrate their strong recognition of pre-existing SARS-CoV-2 immune response and support their use for boosting fusion peptide and stem helix antibodies by vaccination.

Discussion

While currently approved coronavirus vaccines and multiple next-generation vaccine candidates are based on spike or RBD proteins[24], here we describe new types of immunogens that aim to elicit different humoral responses, focused on the conserved stem helix or fusion peptide epitopes. The ability to elicit high titers of antibodies against these regions will likely be critical to the development of a “pan-coronavirus vaccine”[25], a goal the molecules designed here aim to contribute towards. The epitope scaffold proteins were engineered using a combination of previously developed grafting methods[13,26] and more recently described machine learning techniques[27] that proved critical for improving both immunogen stability and antigenicity. Structural and mutational analysis confirmed that the epitope scaffolds engage multiple antibodies, as well as their precursors, with high affinity and specificity. For each class of epitope scaffolds, some designs showed tighter antibody binding than corresponding epitope peptides, especially as it relates to the recognition of antibody precursors.

Future studies will assess the ability of these immunogens to elicit fusion peptide or stem helix antibodies by vaccination either as a “prime” or as a “boost” to pre-existing spike-induced immunity. The ability of stem helix ESs to recognize multiple antibody precursors with high affinity recommends their use as a “prime”. As “boosting” immunogens, it is promising that both fusion peptide and stem helix ESs showed robust engagement of humoral responses elicited by spike, significantly higher than what was measured with the synthetic peptides encoding the target domains. It will be critical to minimize the immune responses that target the scaffold framework while focusing antibody elicitation on the target epitope. At least two strategies can be used to achieve this. Different epitope scaffolds presenting the same target epitope can be used in “prime:boost” combinations or displayed together as a “mosaic” on nanoparticles.

Immunosilencing by glycosylation of scaffold reactive regions can also be employed as successfully demonstrated for other immunogens[28].

Acknowledgments.

We thank Whitney Edwards Beck, Deborah Murray, Micah McClain, Thomas Burke, Jack Anderson, Lynn Harrington, Christopher Todd and Kathlene Chmielewski for help with acquiring and managing the human samples. We thank Aria Arus-Altuz for developing the fluorescent probes used for FACS. We thank Brian Watts, Ken Cronin and Alam Munir for their assistance at the DHVI Biomolecular Interaction Analysis (BIA) Core Facility. We thank Wes Rountree and Yunfei Wang for the statistical analysis. We thank Christopher Beverly, Advaita Sing and Christopher O'Donnell from the Azoitei lab for their help. We thank Priyamvada Acharya and Jared Lindenberger for their help with crystallization experiments. Structural results shown in this report are derived from data collected at Southeast Regional Collaborative Access Team (SER-CAT; 22 ID and 19 BM) beamlines at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. **Funding**: This work was supported by grants P01AI158571 (B.F.H) and R01AI155804 (M.L.A) from the National Institute of Allergy and Infectious Diseases of the National Institute of Health and by a Duke School of Medicine COVID Award (M.L.A). **Author Contribution.** MLA, BK and DJM: conceptualization and methodology. MLA, BK, DJM, CH, DW, KW, BR, MB, QY, DWC: investigation; MLA, CH, DJM and BK: visualization; CW, EQP, EBW, KW, KOS, BFH: resources; MLA, DJM, BK: writing, original draft; MLA, DJM, BK, BFH, DW, KW, KOS: writing, review and editing; MLA, BFH, KW, KOS: funding acquisition; MLA: project administration and supervision. **Competing interests**: MLA, BFH, DJM and BK submitted a patent application that covers the engineered epitope scaffolds. The other authors declare no competing interests. **Data and materials availability**: Crystal structures have been deposited in the PDB under accession numbers 8F5I, 8FDO and 8F5H. The epitope scaffold

can be obtained through an MTA from the corresponding author. Other reagents are available upon requests.

Materials and methods

Computational design of epitope scaffolds.

Design of epitope scaffolds by side chain grafting. A database of 9884 candidate scaffolds was created by selecting from the Protein Data Bank (PDB) structures that were: 1) determined by x-ray crystallography; 2) high resolution (<2.8 Å); 3) monomeric; 4) expressed in *E. Coli*; 5) not of human origin and 6) that did not contain ligands. Epitope scaffolds were designed with Rosetta as previously reported, using the RosettaScripts for side chain grafting described in *Silva et. al*., with the following parameters in the *MotifGraft* mover: RMSD_tolerance="0.3"; NC_points_RMSD_tolerance="0.5"; clash_test_residue="ALA"; clash_score_cutoff="5". For the design of fusion peptide (FP) epitope scaffolds, the structure of fusion peptide fragment ^{815}RSFIEDLLF^{823} as described in the crystal structure of the DH1058-FP peptide complex (PDBid: 7tow) was used as the target epitope to graft. With the exception of residues F817 and L821 which were allowed to change in the designed epitope scaffold because they did not contribute to antibody binding, the identity of all the other epitope residues was maintained. For the design of epitope scaffolds that bind S2P6 and DH1057.1 mAbs, the structure of the stem helix fragment ^{1148}FKEELDKYF^{1156} from PDBid:7rnj was grafted. The identity of residues E1150 and K1154 was allowed to change during the design process while the other epitope residues were kept fixed. From the epitope scaffold generated by the automated protocol, the top 100 models by Rosetta ddG that were also smaller than 150 amino acids were visually examined to ensure appropriate epitope transplantation and antibody-scaffold interaction. At this stage, the suitability of a given parent scaffold in terms of function, conformational flexibility, and expression protocol was also investigated by referencing the publication it originated from. If necessary, additional changes were introduced in a candidate epitope scaffold using Rosetta fixed backbone design to remove antibody-epitope scaffold contacts that were not due to the

target epitope and to ensure proper interactions between the epitope and the rest of the scaffold. The best 15 designs for each of the target epitopes were chosen for experimental characterization.

Design of stem helix epitope scaffolds by backbone grafting and MPNN optimization.

Epitope scaffolds that display the stem helix epitope recognized by the CC40.8 antibody (PDBid: 7sjs) were designed by backbone grafting as previously reported and using the RosettaScripts described by Silva et. al. The same set of curated parent scaffolds as the one used for side chain grafting above was searched here. The structure of the stem helix fragment ^{1144}ELDSFKEELDKYFK^{1157} from PDBid:7sjs was grafted. The identity of residues E1144, D1146, S1147, E1150, K1154, and K1157 were allowed to change during the design process while the other epitope residues were kept fixed. The following parameters were used:

```
RMSD_tolerance="5.0" NC_points_RMSD_tolerance="0.75"
clash_score_cutoff="5" clash_test_residue="ALA"
hotspots="2:5:6:8:9:10:12:13"
combinatory_fragment_size_delta="1:1”
max_fragment_replacement_size_delta="-4:4"
full_motif_bb_alignment="0"
allow_independent_alignment_per_fragment="0"
graft_only_hotspots_by_replacement="0"
only_allow_if_N_point_match_aa_identity="0"
only_allow_if_C_point_match_aa_identity="0"
revert_graft_to_native_sequence="0"
allow_repeat_same_graft_output="0" />
```

The initial hits were filtered to remove scaffolds with (1) an average clash score greater than 5 across the epitope, and (2) a Rosetta calculated ddg greater than 0. The top 100 models by Rosetta ddG that were also smaller than 150 amino acids were then visually examined as

above and then Alphafold2 was used to predict the structure of the 10 best designs using the following parameters-

```
--model_preset=monomer
--db_preset=full_dbs
--max_template_date=2021-11-01
--uniref90_database_path=/data/uniref90/uniref90.fasta
--mgnify_database_path=/data/mgnify/mgy_clusters.fa
--uniclust30_database_path=/data/uniclust30/uniclust30_2018_08/uniclust30_2018_08
--bfd_database_path=/data/bfd/bfd_metaclust_clu_complete_id30_c90_final_seq.sorted_opt
--template_mmcif_dir=/data/pdb_mmcif/mmcif_files
--obsolete_pdbs_path=/data/pdb_mmcif/obsolete.dat
--pdb70_database_path=/data/pdb70/pdb70
```

The scaffolds were then aligned via the grafted epitope to the stem helix fragment in the CC40.8 structure PDBid:7sjs and inspected for clashes between the scaffold and antibody. Scaffolds with clashes, such as where the antibody binding epitope was occluded or buried, were modified such that key epitope residues above remained fixed while other residues in the epitope and in contacting parts of the scaffold were allowed to vary through iterative rounds of fixed-backbone rotamer-based sequence design in Rosetta. After each round, we checked the Alphafold2 prediction and repeated until the epitope was predicted to be presented in a suitable orientation for antibody binding (**Figure 3a**). Expression tags were screened by N-terminally tagging S2hlx-Ex4 with GST, SUMO, Thioredoxin (Trx), Maltose binding protein (MBP), or F8H tags. As Trx gave the highest fraction of full-length protein it was then used to tag S2hlx-Ex2 and -Ex6 (**Supp. Fig. 11**).

The parent structure of S2hlx-Ex4 (PDBid:2QYW) was used to screen the VAST+ database (https://structure.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml) for homologous

structures, which were identified by RMSD (<2Å) and high fraction of alignment (>80%). Two homologs identified were expressed with Trx tags (**Supp. Fig. 12**).

For ESs where protein expression was low the protein structures, either the Rosetta model or Alphafold prediction, were entered into ProteinMPNN at the website, https://huggingface.co/spaces/simonduerr/ProteinMPNN. Initial designs for S2hlx-Ex2 and -4 used sampling temperature of 0.1 and backbone noise of 0.02 and the whole epitope sequence was fixed, For S2hlx-Ex2, this approach produced six related soluble ESs but failed for S2hlx-Ex4. The second set of S2hlx-Ex4 and S2hlx-Ex6 designs used the higher sampling temperature (0.25) and higher backbone noise (0.2) to increase sequence diversity in designs, and only L^{1145}xxFKxELDxYF^{1156} residues were fixed. ES structures were modelled using the integrated Colabfold prediction and scored by RMSD and pLDDT, and the highest scoring were re-predicted using Alphafold2. For the S2hlx-Ex6 designs, the epitope was occluded or buried in several designs. In these cases, residues in the epitope or surrounding parts of the ES were manually modified and the ES structure prediction repeated using Alphafold2 until the epitope was predicted to be suitable to binding to S2P6.

Plasmids and DNA synthesis.

Genes encoding designed epitope scaffolds were commercially synthesized and cloned into pET29b (Genscript). Genes encoding the antibody heavy and light chains were similarly synthesized and cloned into the pcDNA3.1 vector (GenScript). Oligonucleotides were synthesized by IDT. Mutations were introduced in the synthesized plasmids by Q5 mutagenesis (NEB).

Recombinant protein expression and purification.

Epitope-scaffolds.

Plasmids encoding epitope-scaffolds were transformed into E. coli BL-21 (DE3) (New England Biolabs) cells and 5mL starter cultures were grown overnight at 37C in Lysogeny Broth (LB) supplemented with 50µg/mL kanamycin. The cultures were diluted 1:100 in Terrific BrothTB media (RPI) supplemented with kanamycin and grown at 37°C to an OD600 of ~0.6. The temperature was subsequently lowered to 16°C and the cultures were grown to OD600 of ~0.8 and induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5mM. The cultures were shaken for 16–18hours at 200rpm. Cell pellets were collected by centrifugation at 14,500x*g*xg and lysed in B-PER Reagent (ThermoFisher Scientific) according to the manufacturer's protocol. The lysates were centrifuged at 14,000 xg for 30 minutes at 4C. The supernatant was incubated with Ni-NTA beads (Qiagen) equilibrated with native binding buffer (20mM NaH2PO4, 500M NaCl, 2% glycerol, 10mM imidazole, pH=7.5) for an hour at 4C. The beads were settled by centrifugation and the supernatant was removed by pipetting. The beads were washed with wash buffer (50mM NaH2PO4, 500M NaCl, 2% glycerol, 30mM imidazole, pH=7.5) after which the protein was eluted with elution buffer (20mM NaH2PO4, 500M NaCl, 250mM imidazole, pH=7.5). Protein expression and purity was confirmed by SDS-PAGE analysis, and quantified spectrophotochemically at 280nm on a Nanodrop 2000 (ThermoFisher Scientific). The eluted protein was concentrated on 3kDa MW spinMW spin columns (ThermoFisher Scientific) and further purified by size exclusion chromatography in 20mM NaH2PO4, 150mM NaCl, pH=7.5, on an AKTA-Go FPLC (Cytiva) using a Superdex 200 Increase 10/300GL column. Fractions containing monomeric protein were pooled and concentrated as above.

Monoclonal antibodies.

Expi293F cells (ThermoFisher Scientific) were split to a density of 2.5×10^6 cells/mL in Expi293 Expression Medium with GlutaMAX (Gibco) and were transiently transfected with an equimolar plasmid mixture of heavy and light chain using Expifectamine (Invitrogen). For a

typical 100mL size transfection 100µg amount of total DNA and 270µL of lipofectamine were used. After overnight incubation, enhancers were added as per the manufacturer's protocol and the cultures were incubated with shaking for five days at 37C and 5% CO_2. The cell culture was centrifuged to remove the cells and the supernatant was filtered with a 0.8-micron filter. The filtered supernatant was incubated with equilibrated Protein A beads (ThermoFisher) for one hour at 4 °C and washed with 20mM Tris, 350mM NaCl at pH=7. The antibodies were eluted with a 2.5% Glacial Acetic Acid Elution Buffer and were buffer exchanged into 25mM Citric Acid, 125mM NaCl buffer at pH=6. IgG expression was confirmed by reducing SDS-PAGE and quantified by measuring absorbance at 280nmm (Nanodrop 2000).

Antigenic characterization of epitope-scaffolds.

Binding by Surface Plasmon Resonance.

To determine the dissociation constants (K_Ds) between epitope scaffolds and target antibodies, binding was measured by Surface Plasmon Resonance (SPR) on a Biacore T200 instrument using Protein-A coated S series chips. The target antibodies were individually captured for 60 sec at 30 µL/min to a level of 1000-2000RU onto this surface. The epitope scaffolds were subsequently injected as analytes at five concentrations using the single cycle injection method. The association phase was carried out for 180 seconds and the dissociation was done for 900 seconds with HBS-EP+ buffer flowing at 30µL/min. Regeneration of the binding surface was done in 10mM glycine-HC, pH=2 for 30 seconds at 30µL/min with a 30 second baseline stabilization. A 1:1 Langmuir or Heterogenous ligand model was used for data fitting and analysis.

Antibody binding by ELISA.

Fusion peptide (residues 808-833) and stem helix (1140-1163) peptides and mutated variants were commercially synthesized (Genscript). Peptides were diluted to 0.2µg/mL in Superblock buffer supplemented with azide (80g Whey Protein, 300mL Goat Serum, 20mL aqueous 5% Sodium Azide, 10mL Tween20, 80mL of 25X PBS, diluted to 2L). Epitope scaffolds were diluted in 0.1M Sodium Bicarbonate to 100µg/mL. Antigens were coated onto 96 or 384 well high binding enzyme-linked immunosorbent assay (ELISA) plates (Corning) by overnight incubation at 4C. Plates were washed with Superwash buffer (1X PBS supplemented with 0.1% Tween-20) and blocked for 1 hour at room temperature with Superblock buffer supplemented with sodium azide. Antibodies starting at 100µg/mL were serially diluted in Superblock with azide using a threefold serial dilution, added to the plates, and incubated at room temperature for one hour. Plates were then washed twice with Superwash and goat anti Human IgG-HRP secondary antibody (Jackson Immunoresearch) diluted in SuperBlock (without Sodium azide) was subsequently added. After incubating at room temperature for one hour, the plates were washed four times in Superwash. Room temperature TMB (Tetramethylbenzidine) substrate was added and after 5 minutes elapsed, the reaction was stopped by acid stop solution (0.33N HCl). Absorption was measured at 405nm on a Cytation 1 plate reader (BioTek). Data was analyzed and plotted with Prism version 9.3.1 (Graphpad).

Structural analysis by x-ray crystallography.

Crystallography experiments were performed using the sitting drop vapor diffusion technique. 96-well crystallization plates were set, in which the protein of interest was mixed in a 1:1 ratio with reservoir solution. For *S2hlx-Ex*19, crystals were obtained when a 9.5 mg/mL sample was mixed with 0.1 M citric acid, 3.0 M sodium chloride, pH 3.5 (Index Screen) at 4℃. *S2hlx*-7 and antibody DH1057.1 were mixed in 1:1 molar ratio (300uM) and incubated for one hour at 4C. Crystals for the complex were obtained in 25% PEG 3350 at room temperature.

Prior to data collection, crystals were cryoprotected in mother liquor supplemented with 20% glycerol before being plunge-frozen into liquid nitrogen. A full description of the crystallographic data collection and refinement statistics can be found in **Supplementary Table 2**. Diffraction data were indexed in iMOSFLM[29] and scaled in AIMLESS[30]. Molecular replacement solutions for both *S2hlx-Ex*19 and the *S2hlx*-7-DH1057.1 complex were found in PhaserMR[31], using PDB ID: 3N1B as a search ensemble for *S2hlx-Ex*19, PDB ID: 3LMO as a search ensemble for *S2hlx*-7 and PDB ID: 6UOE as a search ensemble for DH1057.1. Coordinates for *S2hlx-Ex*19 and *S2hlx*-7-DH1057.1 were iteratively built and refined using Coot[32], ISOLDE[33] and Phenix[34] to Rwork/Rfree values of 25.4/ 27.5 and 21.1/24.5, respectively. The presence of translational non-crystallographic symmetry in the *S2hlx*-7-DH1057.1 crystal likely accounts for its relatively high Rfree value and the L-test did not indicate the presence of any twinning.

Reactivity of designed epitope scaffolds with human samples.

Human subjects studies were approved by the Duke University Health System Institutional Review Board (IRB) and conducted in agreement with the policies and protocols approved by the Duke IRB, consistent with the Declaration of Helsinki. Written informed consent was obtained from all research subjects or their legally authorized representatives. Study participants with SARS-CoV-2 acute infection were enrolled in the Molecular and Epidemiological Study of Suspected Infection protocol (MESSI, IRB Pro00100241) at Duke University, and were followed longitudinally. Samples were selected for this study from participants who had seroconverted and had symptom onset more than 10 days prior. From patient 25 and patient 27, PBMCs were collected and analyzed by FACS at the same time point used for sera analysis for patient 25 and 7 days later for patient 27. Samples from vaccinated only participants were obtained from subjects enrolled in the Study of Immune Response to COVID-19 Vaccines (IRB Pro00107929) at Duke University. Supplementary Table 2 describes

the vaccination protocol, SARS-CoV-2 variant detected and the time of collection for the samples used in this analysis.

.

384-well ELISA plates were coated with 2 μg/mL of antigens in 0.1 M sodium bicarbonate overnight at 4 °C. Plates were washed with PBS +0.1% Tween 20 and blocked with blocked with assay diluent (PBS containing 4% (w/v) whey protein, 15% Normal Goat Serum, 0.5% Tween-20, and 0.05% Sodium Azide) at room temperature for 1 h. Human sera were serially diluted threefold in superblock starting at a 1:30 dilution. Serially diluted samples were added and incubated for 1.5 h, followed by washing with PBS-0.1% Tween 20. HRP-conjugated goat anti-human IgG secondary Ab (SouthernBiotech, catalog# 2040-05) was diluted to 1:15,000 and incubated at room temperature for 1 h. These plates were washed four times and developed with tetramethylbenzidine substrate (SureBlue Reserve- KPL). The reaction was stopped with 1 M HCl, and optical density at 450 nm (OD450) was determined. Statistical differences were tested using the Wilcoxon signed rank test.

Flow Cytometry

We prepared B cell tetramers of biotinylated S2hlx-Ex19 by mixing with Streptavidin-VB515 (Miltenyi Biotec) and Streptavidin-AlexaFluor 647 (ThermoFisher Scientific) at 4:1 molar ratio. A scaffold-only tetramer was made with biotinylated S2hlx-Ex19-PShlx mixed with Streptavidin-BV421 (Biolegend).

Cryopreserved PBMCs were thawed in warm RPMI-1640 containing 10% FBS, then counted. Cells were stained with pre-optimized concentrations of the following antibodies: PE anti-human IgD (clone IA6-2, BD Biosciences), PE-TXRD anti-human CD10 (clone HI10A, BD Biosciences), PE-Cy5 anti-human CD3 (clone HIT3a, BD Biosciences), PE-Cy7 anti-human C27

(clone O323, ThermoFisher Scientific), AlexaFluor 700 anti-human CD38 (clone LS198-4-3, Beckman Coulter), APC-Cy7 anti-human CD19 (clone SJ25C1, BD Biosciences), BV570 anti-human CD16 (clone 3G8, Biolegend), BV605 anti-human CD14 (clone M5E2, Biolegend), and BV711 anti-human IgM (clone G20-127, BD Biosciences). S2hlx-Ex19-VB515- and -AF647, and S2hlx-Ex19-PShlx-BV421 were added to identify stem helix epitope-specific B cells. Cells were incubated with Aqua Live/Dead (ThermoFisher Scientific) to exclude dead cells. Cells were acquired on a BD S6 cell sorter (BD Biosciences). Data were analyzed using FlowJo v10.8 (BD Biosciences).

References

1 Piccoli, L. *et al.* Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* **183**, 1024-1042 e1021 (2020). https://doi.org:10.1016/j.cell.2020.09.037

2 Wang, Z. *et al.* mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* **592**, 616-622 (2021). https://doi.org:10.1038/s41586-021-03324-6

3 Goel, R. R. *et al.* Distinct antibody and memory B cell responses in SARS-CoV-2 naïve and recovered individuals after mRNA vaccination. *Science Immunology* **6**, eabi6950 (2021). https://doi.org:10.1126/sciimmunol.abi6950

4 Robbiani, D. F. *et al.* Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437-442 (2020). https://doi.org:10.1038/s41586-020-2456-9

5 Starr, T. N. *et al.* Deep mutational scans for ACE2 binding, RBD expression, and antibody escape in the SARS-CoV-2 Omicron BA.1 and BA.2 receptor-binding domains. *PLoS Pathog* **18**, e1010951 (2022). https://doi.org:10.1371/journal.ppat.1010951

6 Qu, P. *et al.* Enhanced neutralization resistance of SARS-CoV-2 Omicron subvariants BQ.1, BQ.1.1, BA.4.6, BF.7, and BA.2.75.2. *Cell Host Microbe* (2022). https://doi.org:10.1016/j.chom.2022.11.012

7 Dacon, C. *et al.* Broadly neutralizing antibodies target the coronavirus fusion peptide. *Science* **377**, 728-735 (2022). https://doi.org:10.1126/science.abq3773

8 Low, J. S. *et al.* ACE2-binding exposes the SARS-CoV-2 fusion peptide to broadly neutralizing coronavirus antibodies. *Science* **377**, 735-742 (2022). https://doi.org:10.1126/science.abq2679

9 Pinto, D. *et al.* Broad betacoronavirus neutralization by a stem helix–specific human antibody. *Science* **373**, 1109-1116 (2021). https://doi.org:10.1126/science.abj3321

10 Zhou, P. *et al.* A human antibody reveals a conserved site on beta-coronavirus spike proteins and confers protection against SARS-CoV-2 infection. *Science Translational Medicine* **14**, eabi9215 (2022). https://doi.org:10.1126/scitranslmed.abi9215

11 Dacon, C. *et al.* Rare, convergent antibodies targeting the stem helix broadly neutralize diverse betacoronaviruses. *Cell Host & Microbe* **0** (2022). https://doi.org:10.1016/j.chom.2022.10.010

12 Zhou, P. *et al.* Broadly neutralizing anti-S2 antibodies protect against all three human betacoronaviruses that cause severe disease. (Immunology, 2022).

13 Azoitei, M. L. *et al.* Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. *Science* **334**, 373-376 (2011). https://doi.org:10.1126/science.1209368

14 Correia, B. E. *et al.* Computational Design of Epitope-Scaffolds Allows Induction of Antibodies Specific for a Poorly Immunogenic HIV Vaccine Epitope. *Structure* **18**, 1116-1126 (2010). https://doi.org:10.1016/j.str.2010.06.010

15 Ofek, G. *et al.* Elicitation of structure-specific antibodies by epitope scaffolds. *Proceedings of the National Academy of Sciences* **107**, 17880-17887 (2010). https://doi.org:10.1073/pnas.1004728107

16 Ofek, G. *et al.* Elicitation of structure-specific antibodies by epitope scaffolds. *Proc Natl Acad Sci U S A* **107**, 17880-17887 (2010). https://doi.org:10.1073/pnas.1004728107

17 Sesterhenn, F. *et al.* Trivalent cocktail of de novo designed immunogens enables the robust induction and focusing of functional antibodies in vivo. *bioRxiv*, 685867 (2019). https://doi.org:10.1101/685867

18 Gobeil, S. M.-C. *et al.* Structural diversity of the SARS-CoV-2 Omicron spike. *Molecular Cell* **82**, 2050-2068.e2056 (2022). https://doi.org:10.1016/j.molcel.2022.03.028

19 Stalls, V. *et al.* Cryo-EM structures of SARS-CoV-2 Omicron BA.2 spike. *Cell Reports* **39**, 111009 (2022). https://doi.org:10.1016/j.celrep.2022.111009

20 Li, D. *et al.* In vitro and in vivo functions of SARS-CoV-2 infection-enhancing and neutralizing antibodies. *Cell* **184**, 4203-4219.e4232 (2021). https://doi.org:10.1016/j.cell.2021.06.021

21 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). https://doi.org:10.1038/s41586-021-03819-2

22 Dauparas, J. *et al.* Robust deep learning–based protein sequence design using ProteinMPNN. *Science* **378**, 49-56 (2022). https://doi.org:10.1126/science.add2187

23 Dacon, C. *et al.* Rare, convergent antibodies targeting the stem helix broadly neutralize diverse betacoronaviruses. *Cell Host & Microbe* https://doi.org:10.1016/j.chom.2022.10.010

24 Saunders, K. O. *et al.* Neutralizing antibody vaccine for pandemic and pre-emergent coronaviruses. *Nature* **594**, 553-559 (2021). https://doi.org:10.1038/s41586-021-03594-0

25 Shah, P., Canziani, G. A., Carter, E. P. & Chaiken, I. The Case for S2: The Potential Benefits of the S2 Subunit of the SARS-CoV-2 Spike Protein as an Immunogen in Fighting the COVID-19 Pandemic. *Frontiers in Immunology* **12** (2021). https://doi.org:10.3389/fimmu.2021.637651

26 Azoitei, M. L. *et al.* Computational design of high-affinity epitope scaffolds by backbone grafting of a linear epitope. *J Mol Biol* **415**, 175-192 (2012). https://doi.org:10.1016/j.jmb.2011.10.003

27 Dauparas, J. *et al.* Robust deep learning–based protein sequence design using ProteinMPNN. *Science* **378**, 49-56 (2022). https://doi.org:doi:10.1126/science.add2187

28 Duan, H. *et al.* Glycan Masking Focuses Immune Responses to the HIV-1 CD4-Binding Site and Enhances Elicitation of VRC01-Class Precursor Antibodies. *Immunity* **49**, 301-311.e305 (2018). https://doi.org:10.1016/j.immuni.2018.07.005

29 Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271-281 (2011). https://doi.org:10.1107/S0907444910048675

30 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204-1214 (2013). https://doi.org:10.1107/S0907444913000061

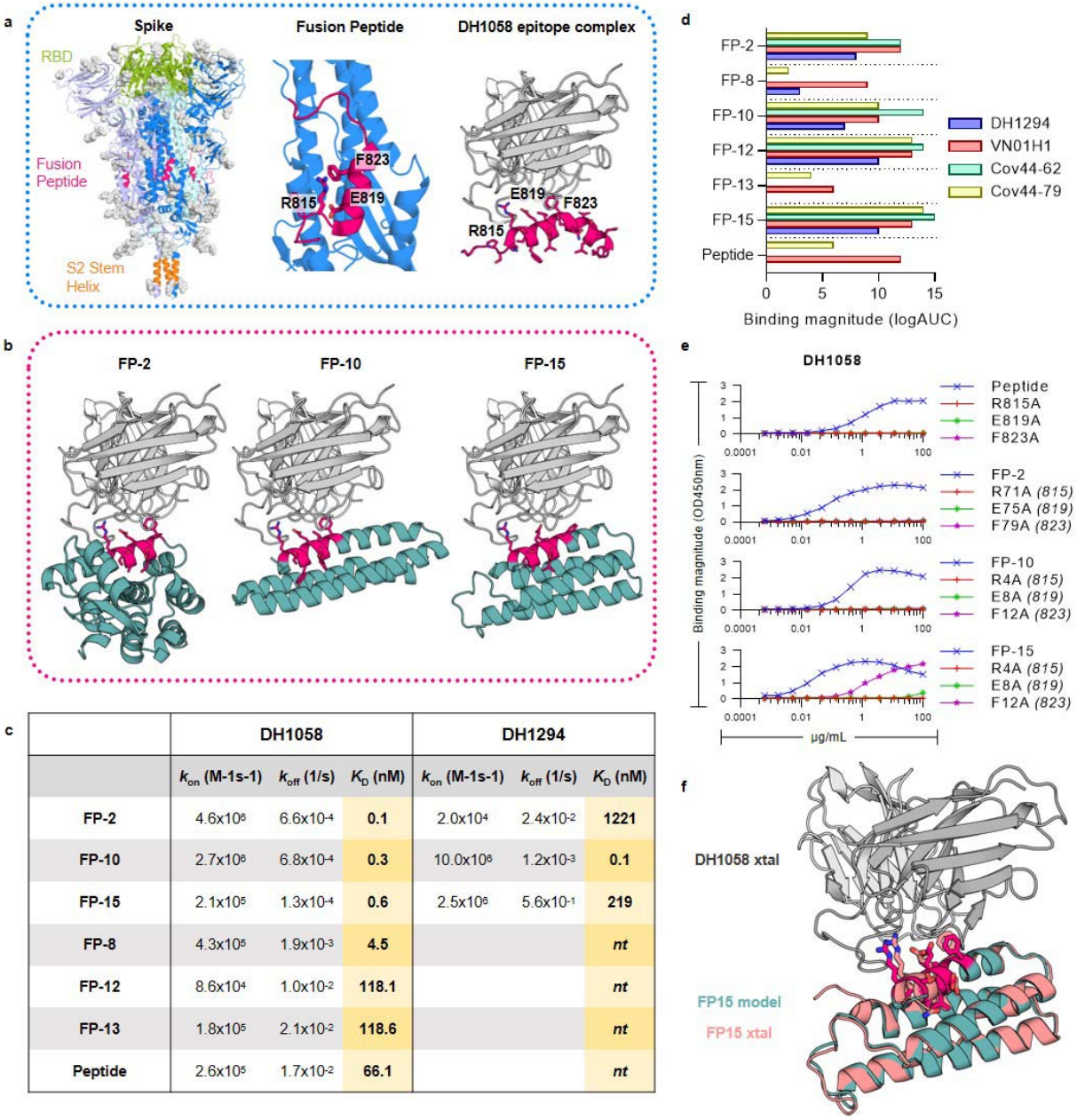
31 McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* **63**, 32-41 (2007). https://doi.org:10.1107/S0907444906045975

32 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004). https://doi.org:10.1107/S0907444904019158

33 Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530 (2018). https://doi.org:10.1107/S2059798318002425

34 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010). https://doi.org:10.1107/S0907444909052925

Figure1

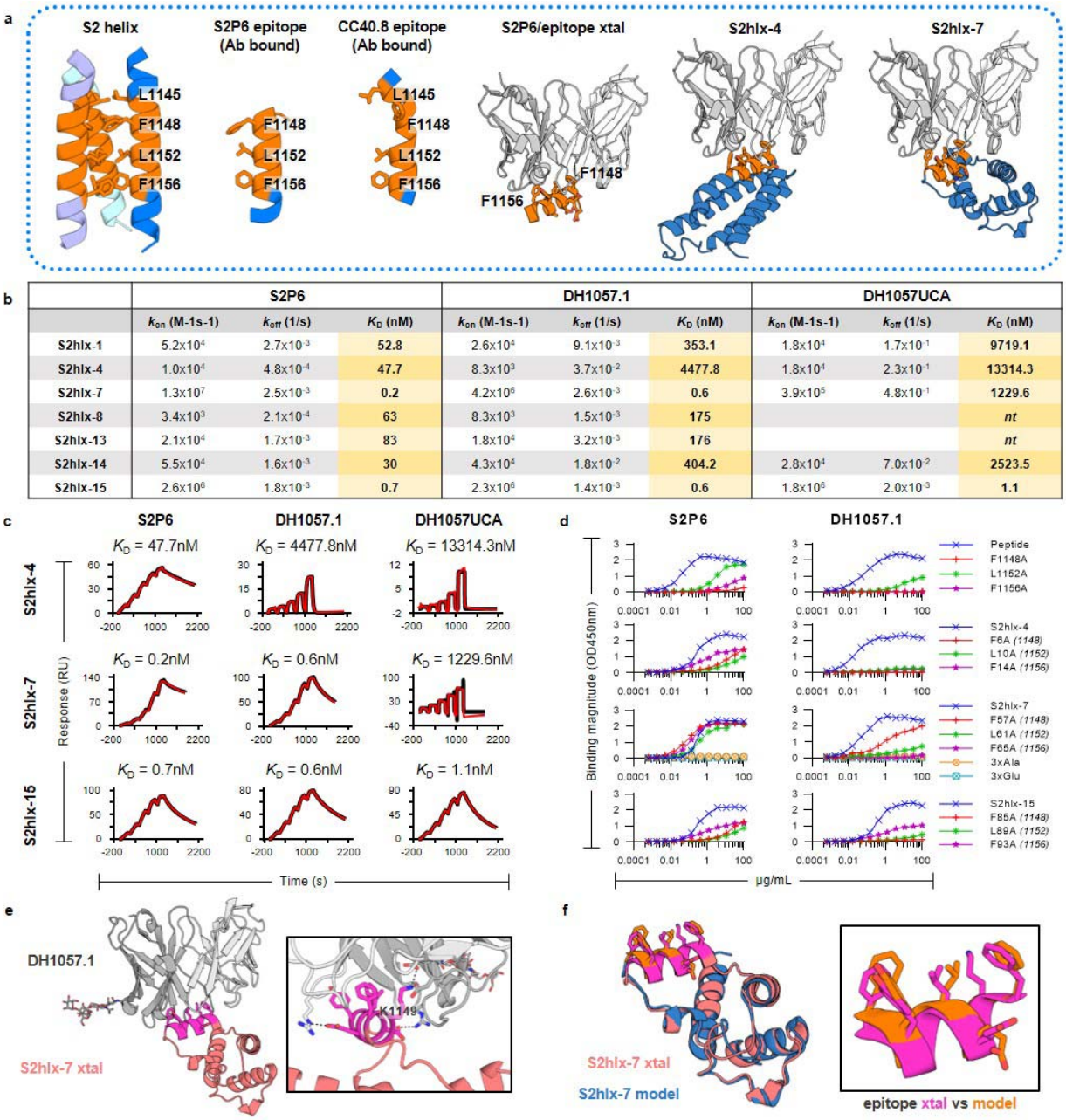


	DH1058			DH1294		
	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)
FP-2	$4.6x10^6$	$6.6x10^{-4}$	**0.1**	$2.0x10^4$	$2.4x10^{-2}$	**1221**
FP-10	$2.7x10^6$	$6.8x10^{-4}$	**0.3**	$10.0x10^6$	$1.2x10^{-3}$	**0.1**
FP-15	$2.1x10^5$	$1.3x10^{-4}$	**0.6**	$2.5x10^6$	$5.6x10^{-1}$	**219**
FP-8	$4.3x10^5$	$1.9x10^{-3}$	**4.5**			***nt***
FP-12	$8.6x10^4$	$1.0x10^{-2}$	**118.1**			***nt***
FP-13	$1.8x10^5$	$2.1x10^{-2}$	**118.6**			***nt***
Peptide	$2.6x10^5$	$1.7x10^{-2}$	**66.1**			***nt***

Figure 1. Design of epitope scaffolds that bind to broadly cross-reactive antibodies against the fusion peptide. (a) *Left*: Structure of the pre-fusion spike trimer (individual monomers: *blue, violet and pale green*; glycans: *grey*; PDBid:6xr8), with the RBD (*green*), fusion peptide (*red*) and stem helix (*orange*) domains highlighted. *Middle*: Zoom of spike monomer with fusion peptide highlighted (*red*) showing key residues engaged by antibodies. *Right*: Structure of

fusion peptide (*red*) bound to DH1058 mAb (*grey*) (PDBid:7tow). **(b)** Computational models of DH1058 (*grey*) bound to ESs FP-2, FP-10 and FP-15 (*green*) with the grafted epitope shown in *red sticks*. **(c)** Binding affinities of FP ESs and the synthetic fusion peptide to DH1058 and DH1294 mAbs as determined by SPR. *nt* = not tested. **(d)** Binding of ES to diverse FP-targeting mAbs by ELISA. **(e)** ELISA binding of DH1058 mAb to synthetic fusion peptides and representative ESs that contain alanine mutations at epitope residues critical for antibody recognition. Numbers in brackets indicate the equivalent position of the mutated epitope residues on spike. **(f)** Crystal structure of DH1058 mAb (*grey*) in complex with FP-15 (*salmon*) overlaid with the computational model of the ESs (*green*). Epitope residues are shown in sticks (model: *red*; crystal structure: *salmon*).

Figure 2

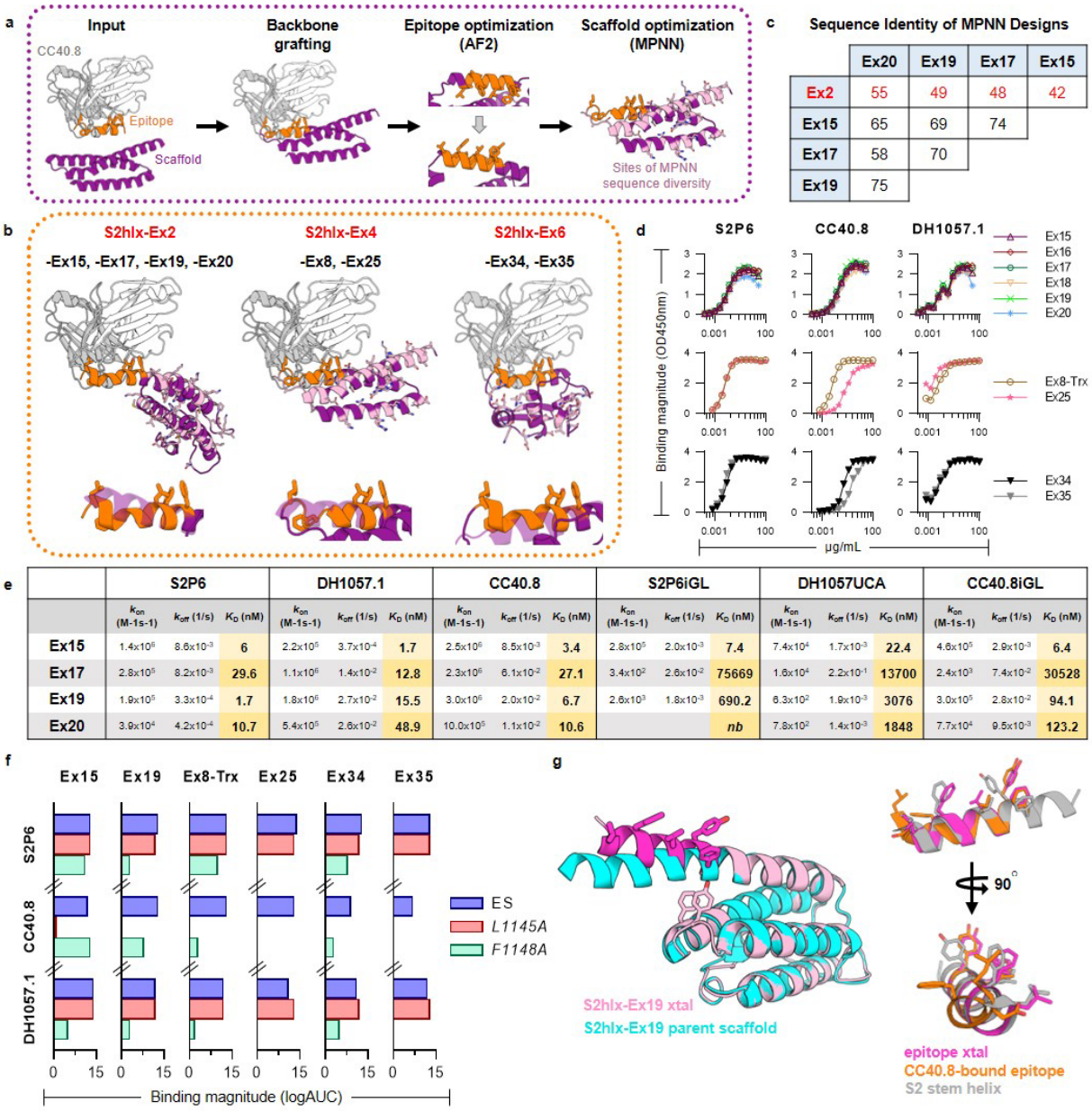


	S2P6			DH1057.1			DH1057UCA		
	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)
S2hlx-1	5.2×10^{4}	2.7×10^{-3}	52.8	2.6×10^{4}	9.1×10^{-3}	353.1	1.8×10^{4}	1.7×10^{-1}	9719.1
S2hlx-4	1.0×10^{4}	4.8×10^{-4}	47.7	8.3×10^{3}	3.7×10^{-2}	4477.8	1.8×10^{4}	2.3×10^{-1}	13314.3
S2hlx-7	1.3×10^{7}	2.5×10^{-3}	0.2	4.2×10^{6}	2.6×10^{-3}	0.6	3.9×10^{5}	4.8×10^{-1}	1229.6
S2hlx-8	3.4×10^{3}	2.1×10^{-4}	63	8.3×10^{3}	1.5×10^{-3}	175			*nt*
S2hlx-13	2.1×10^{4}	1.7×10^{-3}	83	1.8×10^{4}	3.2×10^{-3}	176			*nt*
S2hlx-14	5.5×10^{4}	1.6×10^{-3}	30	4.3×10^{4}	1.8×10^{-2}	404.2	2.8×10^{4}	7.0×10^{-2}	2523.5
S2hlx-15	2.6×10^{6}	1.8×10^{-3}	0.7	2.3×10^{6}	1.4×10^{-3}	0.6	1.8×10^{6}	2.0×10^{-3}	1.1

Figure 2. Design of epitope scaffolds that bind to S2P6-like antibodies against the stem helix. (a) The conformation of the stem helix epitope (*orange*) on the pre-fusion trimer, bound by S2P6 and CC40.8mAbs (PDBids:7rnj and 7sjs), and displayed on representative epitope scaffolds S2hlx-4 and -7. Key epitope residues are labeled and shown in sticks. **(b)** Binding affinities of S2hlx ESs to S2P6, DH1057.1 and DH1057UCA mAbs. *nt*=not tested. **(c)**

Representative single-cycle binding SPR traces used to determine the dissociation constants in **(b)**. **(d)** The effect on S2P6 and DH1057 mAbs binding to epitope mutations in synthesized stem helix peptides (1140-1164) and representative S2-hlx ESs. **(e)** Crystal structure of the DH1057.1 mAb (*grey*) bound to ES S2hlx-7 (*orange*) with the epitope in *pink sticks*. Inset shows details of the epitope/mAb interaction, with epitope residue K1149 highlighted. **(f)** Comparison of the S2hlx-7 conformation in the computational model (*pink*) and the crystal structure (*blue*), both overall and in the epitope region.

Figure 3



	Ex20	Ex19	Ex17	Ex15
Ex2	55	49	48	42
Ex15	65	69	74	
Ex17	58	70		
Ex19	75			

	S2P6			DH1057.1			CC40.8			S2P6iGL			DH1057UCA			CC40.8iGL		
	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)	k_{on} (M-1s-1)	k_{off} (1/s)	K_D (nM)
Ex15	1.4×10^{6}	8.6×10^{-3}	6	2.2×10^{5}	3.7×10^{-4}	1.7	2.5×10^{6}	8.5×10^{-3}	3.4	2.8×10^{5}	2.0×10^{-3}	7.4	7.4×10^{4}	1.7×10^{-3}	22.4	4.6×10^{5}	2.9×10^{-3}	6.4
Ex17	2.8×10^{5}	8.2×10^{-3}	29.6	1.1×10^{6}	1.4×10^{-2}	12.8	2.3×10^{6}	6.1×10^{-2}	27.1	3.4×10^{2}	2.6×10^{-2}	75669	1.6×10^{4}	2.2×10^{-1}	13700	2.4×10^{3}	7.4×10^{-2}	30528
Ex19	1.9×10^{5}	3.3×10^{-4}	1.7	1.8×10^{6}	2.7×10^{-2}	15.5	3.0×10^{6}	2.0×10^{-2}	6.7	2.6×10^{3}	1.8×10^{-3}	690.2	6.3×10^{2}	1.9×10^{-3}	3076	3.0×10^{5}	2.8×10^{-2}	94.1
Ex20	3.9×10^{4}	4.2×10^{-4}	10.7	5.4×10^{5}	2.6×10^{-2}	48.9	10.0×10^{6}	1.1×10^{-2}	10.6			*nb*	7.8×10^{2}	1.4×10^{-3}	1848	7.7×10^{4}	9.5×10^{-3}	123.2

Figure 3. Design of epitope scaffolds that engage the two major classes of antibodies against the stem helix. (a) Overview of the computational design process to graft the antibody-bound CC40.8 epitope onto candidate scaffolds. ESs were designed by transplanting the epitope backbone with Rosetta, validated with Alphafold2 and optimized for expression with ProteinMPNN. **(b)** *Top*: Computational models of CC40.8 mAb (*gray and white*) bound to ESs (*purple*) from the S2hlx-Ex2, S2hlx-Ex4, and S2hlx-Ex6 families of designs. The epitope is

highlighted in *orange sticks*. Residues designed by ProteinMPNN and that can differ between designs from the same family are shown in *pink sticks*. *Bottom*: Overlay of the grafted epitope conformation (*orange*) and the parent scaffold structure it replaced (*purple*). **(c)** Sequence identity between S2hlx-Ex2 family of ESs designed by ProteinMPNN. **(d)** Binding of S2hlx-Ex ESs to S2P6, CC40.8 and DH1057.1 mAbs. ESs are grouped by the parent backbone grafting design they are based on: S2hlx-Ex2 (*upper*), S2hlx-Ex4 (*middle*), and S2hlx-Ex6 (*lower*). **(e)** Dissociation constants of S2hlx-Ex2 based ESs to S2P6, DH1057 and CC40.8 mAbs and their precursors. *nb*=not binding. **(f)** ELISA binding to CC40.8, S2P6 and DH1057 mAbs by ESs based on different parent scaffolds and that contain alanine mutations at indicated epitope residues, with their corresponding position on spike shown in brackets. **(g)** *Left*: alignment of the S2hlx-Ex19 parent scaffold (*cyan*) with the crystal structure of the unbound epitope scaffold (*pink*, with the epitope in *magenta sticks*). Residues Tyr62 and Phe66 that stabilize the grafted epitope conformation are highlighted. *Right*: Comparison of the stem helix epitope conformation in the pre-fusion spike (*grey*), induced by CC40.8 binding (*orange*) and displayed on S2hlx-Ex19 (*magenta*).

Figure 4

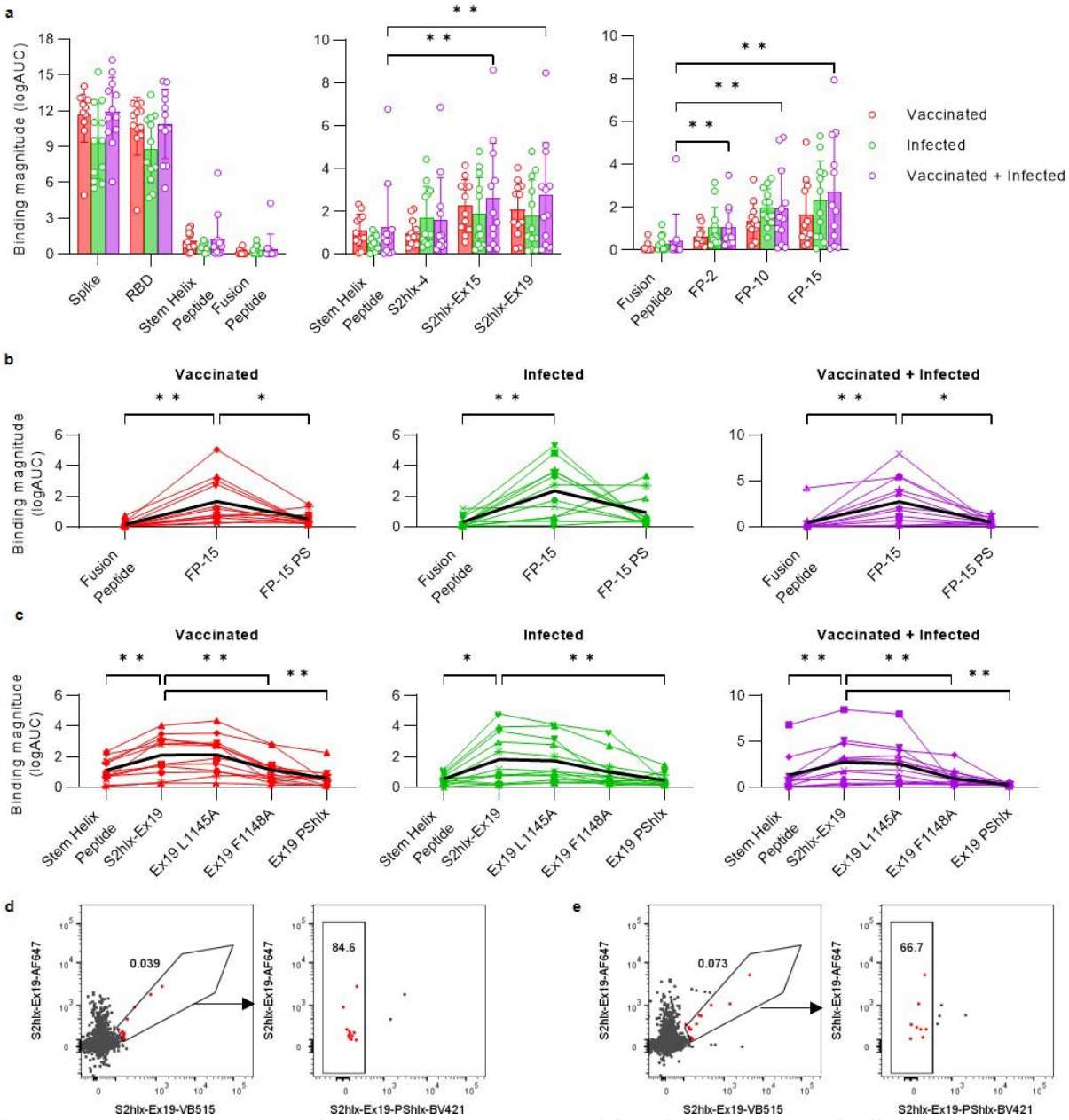


Figure 4. Reactivity of engineered epitope scaffolds with sera from individuals with pre-existing immune responses to SARS-CoV-2 spike. (a) *Left*: ELISA binding of sera or plasma from patients with pre-existing SARS-CoV-2 immunity acquired by vaccination (*red*), infection (*green*), or vaccination followed by infection (*purple*) to recombinant WA-2 **spike** and RBD or to synthetic peptides encoding the **fusion peptide** (residues 808-833) or the **stem helix peptide**

(1140-1164). *Middle* and *right:* Binding of the same samples in (a) to select **stem helix peptide** and **fusion peptide** epitope scaffolds respectively. (b) Comparison between sera binding to the synthesized **fusion peptide** with FP-15 ES, and the parent scaffold from which FP-15 was derived and that lacks the grafted epitope (FP-15 PS). (c) Comparison between sera binding to the synthesized **stem helix peptide** with S2hlx-Ex19 ES, and with versions of this design where the epitope is mutated to reduce binding to either CC40.8 class antibodies (L1145A), to both CC40.8 and S2P6 classes of antibodies (F1148A), or where the epitope is replaced with the sequence from the parent scaffold it replaced (PShlx). Single lines represent measurements for individual samples in the respective cohort. For all graphs *=$p<0.05$ and **=$p<0.01$. (d) Bivariate plot of labeling of memory B cells from Patient 25 with S2hlx-Ex19 conjugated to VB15 (x-axis) and AF647 (y-axis). A third bait comprising S2hlx-Ex19 PShlx conjugated to BV421 was used to exclude scaffold-specific memory B cells. The overall frequency of stem helix epitope-specific cells was 0.033% of memory B cells. (e) Gating of stem helix epitope-specific memory B cells from Patient 27, comprising 0.047% of total memory B cells.