Antibodies utilizing VL6-57 light chains target a convergent cryptic epitope on SARS-CoV-2 spike protein driving the genesis of Omicron variants

Qihong Yan[1,*], Xijie Gao[2,*], Banghui Liu[3,*], Ruitian Hou[4], Ping He[5], Yong Ma[3], Yudi Zhang[3,6], Yanjun Zhang[1], Zimu Li[3], Qiuluan Chen[7], Jingjing Wang[3], Xiaohan Huang[1], Huan Liang[1], Xianying Chen[1], Xuefeng Niu[1], Jun He[3,#], Ling Chen[1,3,5,#], Jincun Zhao[1,5,#], Xiaoli Xiong[3,#]

[1]State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China
[2]Key Laboratory of Biological Targeting Diagnosis, Therapy and Rehabilitation of Guangdong Higher Education Institutes, The Fifth Affiliated Hospital of Guangzhou Medical University, Guangzhou, China
[3]State Key Laboratory of Respiratory Disease, CAS Key Laboratory of Regenerative Biology, Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangdong Provincial Key Laboratory of Biocomputing, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, China
[4]Guangzhou Eighth People's Hospital, Guangzhou Medical University, Guangzhou, China
[5]Guangzhou National Laboratory, Guangzhou, China
[6]University of Chinese Academy of Sciences, Beijing, China
[7]Bioland Laboratory (Guangzhou Regenerative Medicine and Health - Guangdong Laboratory), Guangzhou, China
[*]These authors contributed equally: Qihong Yan, Xijie Gao, Banghui Liu
[#]Correspondence: he_jun@gibh.ac.cn; chen_ling@gibh.ac.cn; zhaojincun@gird.cn; xiong_xiaoli@gibh.ac.cn

Abstract

Continued evolution of SARS-CoV-2 generates variants to challenge antibody immunity established by infection and vaccination. A connection between population immunity and genesis of virus variants has long been suggested but its molecular basis remains poorly understood. Here, we identify a class of SARS-CoV-2 neutralising public antibodies defined by their shared usage of VL6-57 light chains. Although heavy chains of diverse genotypes are utilized, convergent HCDR3 rearrangements have been observed among these public antibodies to cooperate with germline VL6-57 LCDRs to target a convergent epitope defined by RBD residues S371-S373-S375. Antibody repertoire analysis identifies that this class of VL6-57 antibodies is present in SARS-CoV-2-naïve individuals and is clonally expanded in most COVID-19 patients. We confirm that Omicron specific substitutions at S371, S373 and S375 mediate escape of antibodies of the VL6-57 class. These findings support that this class of public antibodies constitutes immune pressure promoting the introduction of S371L/F-S373P-S375F in Omicron variants. The results provide further molecular evidences to support that antigenic evolution of SARS-CoV-2 is driven by antibody mediated population immunity.

Introduction

The SARS-CoV-2 pandemic offered an opportunity to interrogate immune response towards an emergent virus infecting an immunologically naïve population. During the first 3 years of the pandemic, the spike (S) protein on the virus surface has been actively undergoing antigenic drift generating a large number of variants (Carabelli et al., 2023; Harvey et al., 2021; Markov et al., 2023; Volz, 2023). In late 2021, the Omicron BA.1 variant emerged, its highly mutated S protein renders most antibodies isolated earlier in the pandemic ineffective, conferring the Omicron BA.1 variant with an extraordinary immune evasion capability (Dejnirattisai et al., 2022; Liu et al., 2022). Further antigenic drift of S-protein had been observed and numerous Omicron subvariants emerged, including BA.2-BA.5 (Cao et al., 2022; Tuekprakhon et al., 2022; Wang et al., 2022a). Since late 2022, XBB and BQ.1 subvariant series have emerged based on BA.2 and BA.5 respectively with further antigenic changes on the S-protein (Qu et al., 2023; Wang et al., 2023). As the pandemic progresses, S-protein is drifting antigenically further generating new variants.

The emergence of Omicron variants marked great shifts in S protein antigenicity and virological behaviours, sparking several theories on origins of Omicron variants (Mallapaty, 2022). Several recurrent mutations observed in earlier variants became fixed in the Omicron S-protein, most of these recurrent mutations are located within the receptor binding domain (RBD) of the S-protein, including K417N, L452R, T478K, E484A and N501Y. Earlier studies have identified that some of these mutations are located in the epitopes of several classes of germline antibodies that are widely induced within the population. K417N is located in the epitope of VH3-53/3-66 encoded class 1 antibodies inhibiting ACE2 binding (Yan et al., 2021; Yuan et al., 2020a; Zhang et al., 2021b); L452R is located in the epitope of VH1-69 encoded population antibodies (He et al., 2022; Huo et al., 2022; Yan et al., 2022); E484A is located in the epitope of VH1-2 encoded class 2 population antibodies (Rapp et al., 2021; Yuan et al., 2021). These mutations often confer viruses with considerable antibody evading ability. While L452R and N501Y have also been proposed to enhance or compensate fusogenic or receptor binding activity of S-protein (Deng et al., 2021; Harvey et al., 2021; Starr et al., 2022; Tchesnokova

et al., 2021), association of recurrent mutations to population antibody responses strongly implies that antigenic drift of SARS-CoV-2 S-protein may be driven by immune pressure at population level. Population immunity (or herd immunity) has long been proposed to drive antigenic drift of influenza virus haemagglutinin (Boni, 2008), despite molecular basis behind such driving force remains poorly characterized. Apart from the above identified recurrent mutations, several RBD mutations were first emerged with the highly mutated Omicron BA.1 S protein, including S371L, S373P, S375F, N440K, G446S, S477N, Q493R, G496S and Q498R, most of them have been fixed in the S-protein of current circulating viruses. Among these, S371L (S371F in BA.2 - BA.5), S373P and S375F are particularly interesting that they are close together in sequence while located relatively far away from the receptor binding site. The S371L/F-S373P-S375F mutations are considered as an important feature of Omicron variants and they have been shown to alter S-protein functions likely changing virus tissue tropism and disease severity (Kimura et al., 2022a; Pastorio et al., 2022; Willett et al., 2022). However, the driving force behind the genesis of the S371L/F-S373P-S375F mutations remains enigmatic. In this study, we report that diverse antibodies utilizing IGLV6-57 (VL6-57) light chains are widely induced within the population and they target a common epitope defined by S371, S373 and S375 in the ancestral SARS-CoV-2 S-protein. We provide evidences to suggest that the identified class of antibodies is associated with immune pressure driving the genesis of S371L/F-S373P-S375F mutations in Omicron variants.

Identification of SARS-CoV-2 S-specific antibodies with VL6-57 encoded light chains

In an exercise to isolate SARS-CoV-2 S-specific antibodies from convalescent patients infected at the beginning of the pandemic (early 2020), we identified 5 mAbs, namely R1-26, R1-30, R2-3, R2-6, and R2-7 by phage display using SARS-CoV-2 RBD as the bait (He et al., 2022). The identified antibodies either have zero or very low somatic hypermutation rates (**Fig. 1a**), suggesting that they are germline antibodies. Interestingly, all these 5 antibodies utilize light chains encoded by IGLV6-57, while their heavy chain genes are of different genetic origins (**Fig. 1a**). 4 mAbs, namely R1-30, R2-3, R2-6, and R2-7 have a

shared "WLRG" motif in the middle of their HCDR3 (**Fig. 1a**). Different from the other 4 isolated mAbs, R1-26 has a hydrophobic "LGPWV" motif in the middle of HCDR3 (**Fig. 1a**). The five isolated antibodies with IGLV6-57 light chains have affinities in the range of 3.8 - 62.2 nM towards the wildtype SARS-CoV-2 RBD (**Fig. 1a**), with R1-26 being the strongest (K_D = 3.8 nM). In binding competition assays, we found that all the 5 mAbs compete with ACE2 and with each other to bind RBD (**Fig. 1b**). These results suggest that all the 5 mAbs target overlapping epitopes. In pseudovirus neutralization assays, their neutralization activities largely correlate with their affinities, the strongest binder R1-26 has an IC_{50} as low as 2.7 nM (**Fig. 1a**). We studied the most potent mAb - R1-26 in detail and we found that R1-26 is able to bind RBD of SARS-CoV-2 variants emerged before the Omicron BA.1 variant (**Fig. 1c, Table S1**). Consistently, R1-26 is able to bind the corresponding S-trimers tightly without dissociating, likely through avidity (**Fig. 1d, Table S1**). We further confirmed that R1-26 has neutralization activity towards wildtype, Alpha, Beta and Delta SARS-CoV-2 authentic viruses in cell culture (**Fig. 1e**).

Binding of R1-26 antibody to SARS-CoV-2 S-trimer

To further understand neutralization activity of R1-26, we determined cryo-electron microscopy (cryo-EM) structures of S-trimer:R1-26 Fab complexes in two different stoichiometries (3:2, 3:3 S-protomer:R1-26 Fab) (**Figs. 2a, S1a**) with upper resolutions in the range of 3.19 - 3.44 Å (**Fig. S2a-c**). An unusual structure showing a head-to-head aggregate of two S-trimers each bound by 3 R1-26 Fabs was also determined at a lower resolution of 5.3 Å (**Figs. 2a, S1a and S2a-c**), the aggregation is mediated by interactions between Fabs bound to different S-trimers and a similar aggregate was observed for the S-trimer:Fab complex of 6M6 (using IGHV3-91, IGLV3-21 genes) (Wang et al., 2022b). It is not known whether the observed interactions could lead to aggregation of S-trimers on virus surface.

Based on a locally refined RBD:Fab structure at ~ 3.5 Å resolution derived from the 3:3 S-protomer:R1-26 Fab complex dataset (**Figs. S1a and S2a**), the detailed R1-26 epitope is resolved (**Fig. 2b**). Buried surface area (BSA) analysis reveals that HCDRs and LCDRs of R1-26 bury comparable RBD surface areas -

459.0 Å^2 and 385.7 Å^2 respectively. Among CDR loops, HCDR3 dominates the contact with RBD, the hydrophobic HCDR3 loop, containing residues L101, G102, P103 and W104, probes into a hydrophobic cavity formed by RBD residues Y369, A372, F374, F377 and P384 (**Figs. 2b left panel, S2d**). This hydrophobic contact is further stabilized by surrounding electrostatic and polar interactions, distributed in two different patches: the larger interaction patch is mediated by hydrogen bonds between backbone of RBD residues S375, F377 and LCDR1 residue N32 sidechain; a cation-π interaction between RBD residue K378 and LCDR1 residue Y33; a charged hydrogen bond between RBD residue T385 and LCDR2 residue E51; and finally a salt bridge between RBD residue K378 and LCDR2 residue D52 (**Fig. 2b, right panel**). The second interaction patch is mediated by a hydrogen bond between the glycan attached to RBD residue N343 and HCDR2 residue K52; and a hydrogen bond between RBD residue N370 and HCDR2 residue Q53 (**Fig. 2b, right panel**). Finally, RBD residue A372 is contacted by LCDR3 residue Y94 (**Fig. 2b, right panel**).

The detailed R1-26 epitope structure confirms that the epitope is fully buried in the previously determined RBD "down" closed and locked S-trimer structures (Qu et al., 2022; Xiong et al., 2020; Zhang et al., 2023), therefore, the epitope is fully cryptic in a S-trimer adopting 3-RBD "down" closed or locked conformations (**Fig. S3a-h**). In the 1-RBD "up" S-trimer, this epitope on the "up" RBD is also largely obstructed (**Fig. S3a**). Likely due to incompatibility with 3-RBD "down" and 1-RBD "up" S-trimer conformations, at least 2 RBDs are observed to adopt "up" position when R1-26 is bound to a S-trimer and only 3:2, 3:3 S-protomer:Fab structures are observed when R1-26 Fab was added in excess (**Fig. 2a**). The detailed epitope structure also confirmed that the R1-26 epitope does not overlap with the ACE2 binding interface (**Fig. S3f and h**) indicating that R1-26 Fab is a class 4 antibody (**Fig. S3f-h**). CR3022 is a well-studied class 4 mAb which cross-reacts to RBDs of SARS-CoV-1 and SARS-CoV-2 (Yuan et al., 2020b), we found that although the epitope of R1-26 appears to largely overlap with that of CR3022, R1-26 features a different approach angle towards the RBD compared with CR3022 (**Fig. S3a, g and h**). Notably, the canonical class 4 antibody CR3022 does not inhibit ACE2 binding and is only weakly neutralizing, while strong ACE2 binding inhibition is

observed for R1-26 (**Fig. 1b**). By modelling, we found that the approach angle of R1-26 is more tilted towards the modelled ACE2 bound to RBD (Compare **Fig. 2c** and **Fig. S4**), while the approach angle of CR3022 is more tilted away from it (**Fig. S4**). In addition, R1-26 and CR3022 adopt different orientations when bound to RBD (compare **Fig. 2c** and **Fig. S4**). Due to the differences in RBD binding, simultaneous binding of R1-26 and ACE2 to RBD would result in steric clashes between R1-26 and the glycan chain attached to ACE2 residue N322 (**Fig. 2c**). By contrast, simultaneous binding of CR3022 and ACE2 is possible likely due to much weaker clash (**Fig. S4**), consistent with previous results of competition assays (Tian et al., 2020). Differences in RBD binding by R1-26 and CR3022 likely confer superior ACE2 blocking activity to R1-26.

We further found that fusogenic competent native S-R trimer (without proline stabilization) incubated with R1-26 undergoes structural transition into a post-fusion conformation (**Fig. 2d**). This analysis shows that R1-26 possesses an activity to trigger fusogenic conformational change similar to class 1 mAbs including B38 (He et al., 2022) and S230 (Walls et al., 2019) (**Fig. 2d**), it is interesting that class 4 antibody CR3022 does not have this triggering activity (**Fig. 2d**). Both ACE2 binding inhibition and premature triggering of spike fusogenic change by R1-26 could lead to the inhibition effect observed in a S-protein-ACE2 interaction mediated cell-cell membrane fusion assay by which cell entry inhibition activity of antibody can be assessed (Ma et al., 2023) (**Fig. 2e**). Both activities likely contribute to the neutralization activity of R1-26.

VL6-57 light chain can pair with diverse heavy chains to bind a convergent epitope

To further understand antigen binding by VL6-57 light chain utilizing mAbs, we surveyed the Protein Data Bank (PDB) for structurally characterized VL6-57 light chain utilizing mAbs bound to SARS-CoV-2 S-protein or RBD. Among 376 structurally characterized SARS-CoV-2 S-specific mAbs, 12 mAbs utilize VL6-57 encoded light chains, of note, among them, 10 belong to class 4 antibodies (see **Fig. S5** for a structure gallery of class 4 antibodies), these account for more than a fourth of class 4 antibodies (n=35) found amongst the 376 structurally characterized SARS-CoV-2 S-specific mAbs (**Fig. S5**). Further,

germline usage analysis shows that VL6-57 is the most frequently used light-chain germline gene among the class 4 antibodies (**Fig. S6a, left panel**). A similar germline gene usage profile is observed for the F2 group antibodies (equivalent to class 4 antibodies, as defined by deep mutational scanning) (Cao et al., 2022) (**Fig. S6a, right panel**). We did not find evidence of preferential VL6-57 gene usage in the global B cell repertoire data of SARS-CoV-2-naïve individuals (Robbiani et al., 2020). The above analyses strongly suggest that VL6-57 gene preferentially generates S-specific class 4 mAbs, leading to the hypothesis that VL6-57 light chains may play an important role in antigen binding.

We analyzed buried surface area (BSA) to further understand contribution of VL6-57 light chains in epitope binding. Among the 376 structurally characterized SARS-CoV-2 S-specific mAbs, heavy chains dominate epitope interaction, burying significantly more surface areas than light chains (**Fig. S7a**). Different from typical SARS-CoV-2 S-specific mAbs, the average BSA of VL6-57 class 4 mAbs is comparable between heavy and light chains (**Fig. S7b**). This analysis suggests that among VL6-57 class 4 mAbs, light chains are at least equally important as heavy chains in epitope recognition. Further analysis reveals that LCDR1 and HCDR3 contribute more than 60% of BSA (**Fig. S7c**).

Superposition of the 10 VL6-57 class 4 Fab:RBD complex structures reveals that apart from mAbs - 10-40 and 002-13 (**Fig. S5**), all the other 8 mAbs bind to epitopes that are almost identical to the epitope of R1-26 (**Figs. 3a-k, S5**), demonstrating clear epitope convergence among VL6-57 mAbs. Comparison of light and heavy chain CDR sequences reveals that all these structurally characterized mAbs share similar LCDR1-3 sequences (**Fig. S6c**). Interestingly, structurally characterized S2A4, P5S-3B11, GH12 and 3D11 not only exhibit high similarity in LCDR sequences (**Fig. S6c**) to the isolated R1-30, R2-3, R2-6, and R2-7 mAbs, they also feature 12-residue long HCDR3s with either "WLRG" or "WVRG" motif in the middle (**Fig. S6b**). Therefore, R1-30, R2-3, R2-6, and R2-7 are likely to bind the same epitope as S2A4, P5S-3B11, GH12, and 3D11, which is almost identical to that of R1-26, consistent with the result

that R1-26 competes with R1-30, R2-3, R2-6 and R2-7 in binding to spike RBD (**Fig. 1b**). The sequence analysis also identifies that VL6-57 light chains can also pair with heavy chains with HCDR3 loops rich in hydrophobic, particularly aromatic amino acids. mAbs 553-15, 2-7, FP-12A and IY-2A feature "WYYY", "GYFY", "YYYY", and "LGIFG" motifs in the middle of HCDR3 respectively, notably, the "LGIFG" in IY-2A is reminiscent of the "LGPWV" motif found in R1-26 HCDR3. Analysis of their structures shows that these mAbs bind to the same epitope as for S2A4, P5S-3B11, GH12 and 3D11 with "WLRG" or "WVRG" HCDR3 motifs (**Figs. 3, S5, S6b**).

Detailed structural comparison of the identified VL6-57 mAb:RBD complexes reveals that HCDR3s in R1-26, S2A4, 3D11, GH12, P5S-3B11, 553-15, 2-7, FP-12A and IY-2A interact with almost identical RBD residues within the convergent epitope with several modes of interactions (**Fig. 3a-j**). R1-26, FP-12A and IY-2A (**Fig. 3a-c**) appear to bind the epitope using hydrophobic contacts primarily from aromatic residues. 553-15 and 2-7 (**Fig. 3d,e**) bind the epitope by π-π interactions (from the HCDR3 tyrosine) combined with hydrophobic contacts. Finally, S2A4, P5S-3B11, GH12 and 3D11 (**Fig. 3f-i**) binds the epitope with a combination of cation-π interactions (by the arginine in HCDR3) and hydrophobic contacts (by the tryptophan in HCDR3). Different binding modes give rise to variability among HCDR3 structures (**Fig. 3k**), however, hydrophobic interactions are shared among these different interaction modes reflecting the convergent epitope being hydrophobic, containing multiple aromatic residues including Y369, F374 and F377 (**Fig. 3a-i**).

The comparison also identifies that the germline-encoded N32 and Y33 from LCDR1 engage in conserved interaction to spike RBD among the 8 VL6-57 light chain utilizing R1-26-like mAbs (**Figs. 3a-i, S6c**) and LCDR1s show little variability when binding to RBD (**Fig. 3k**). BSA analysis also confirms that LCDR1s bury the largest surface areas among LCDRs (**Fig. S7c**), suggesting that this germline encoded "NY" motif facilitates efficient RBD binding. Collectively, our results demonstrate that the VL6-57 light chain likely serves as an efficient framework for generation of class 4 SARS-CoV-2 S-specific mAbs.

VL6-57 antibodies paired with HCDR3 "WLRG" motif are widely induced in the population and clonally expanded upon SARS-CoV-2 infection

The presence of considerable amount of VL6-57 antibodies among the structurally characterized class 4 mAbs prompted us to further search CoV-AbDab (the Coronavirus Antibody Database) (Raybould et al., 2021) for additional VL6-57 mAbs to investigate their features (**Fig 4a, see method**). The search obtained 290 mAbs utilizing VL6-57 light chains from 49 independent studies. A total of 36 different heavy-chain germline genes are observed to pair with VL6-57 to generate S-specific mAbs (**Fig. S8a**). Notably, we found that there is an enrichment of mAbs with 12-AA long HCDR3s and 9~10-AA long LCDR3s, accounting for 34% (99/290) of the curated VL6-57 mAbs (**Fig. S8b-d**). Moreover, those VL6-57 mAbs with 12-AA long HCDR3s appear to be preferentially derived from heavy chain genes IGHV3-7, IGHV4-39, and IGHV4-59 (**Fig. 4b**). HCDR3 sequence analysis shows that there is a strong preference for a 12-AA long HCDR3 containing the "WLRG" motif (68/99 mAbs) as observed in S2A4, P5S-3B11, 3D11, R1-30, R2-3, R2-6 and R2-7 (**Figs. 4b, S6b**). We also found a small percentage of mAbs (7/99) containing a 12-AA long HCDR3 with the "YYY" motif observed in 553-15 (**Figs. 4b logo plot, S6b**). However, we failed to identify consensus HCDR3 sequences among VL6-57 mAbs with 11-, 13-, 14-, or 15-AA long HCDR3s (**Fig. S8e**). V(D)J rearrangement analysis suggests that the W, L, and R residues within the HCDR3 "WLRG" motif are most likely encoded by IGHD5-12 gene when its 2nd reading frame is used (**Fig. S8f**), the last AA residue G within the motif is located at the DJ junction and is most likely a result of an insertion at the C-terminal side of the D segment (**Fig. S8f**), these findings explain the genetic origin of the HCDR3 "WLRG" motif. LCDR3 sequence analysis shows that a "QSYDSS" motif is enriched (**Fig. S8d**), although BSA analysis shows that LCDR3s in VL6-57 mAbs play an auxiliary role in antigen binding burying smaller areas compared with LCDR1s (**Fig. S7c**), we found that the tyrosine (Y94 in R1-26, **Fig. 2b**) residue within the "QSYDSS" motif is engaging in specific interactions with critical antigen binding residues in HCDR3. In R1-26, FP-12A, IY-2A, 553-15 and 2-7 the LCDR3 tyrosine is interacting with antigen contacting HCDR3 residue W104, Y106, V104, Y103 or Y105, respectively, by hydrophobic contacts (**Fig. 3a-e, LCDR interaction panels**). In S2A4, P5S-

3B11, GH12 and 3D11, the LCDR3 tyrosine is interacting with the R in the HCDR3 "WLRG" motif by cation-π interactions (**Fig. 3f-i, LCDR interaction panels**). These specific interactions suggest LCDR3 supports HCDR3 in antigen binding.

The above analysis suggests that the "WLRG" motif within the 12-AA long HCDR3 and the "QSYDSS" motif within the LCDR3 are the convergent signatures of the SARS-CoV-2 RBD-specific VL6-57 mAbs, which may be widely induced in the COVID-19 population. To verify this, we next searched our previously published bulk antibody repertoires from 24 healthy donors and 33 COVID-19 convalescents (Niu et al., 2020; Yan et al., 2021; Zhang et al., 2022) (**Fig. 4a, see method**). The COVID-19 donors PtZ and PtK, from whom R1-26, R1-30, R2-3, R2-6 and R2-7 have been isolated, are also included in this analysis. In addition to IgH and IgL sequences identical to the 5 isolated mAbs, similar "WLRG" motif containing IgH sequences and "QSYDSS" motif containing VL6-57 IgL sequences can be readily detected in the repertoires of PtZ and PtK (**Fig. S9**). The search shows that "WLRG" motif containing IgH sequences can be detected in 29/33 COVID-19 convalescents and 10/24 healthy donors (**Fig. 4c**). Following SARS-CoV-2 exposure, there was a remarkable enrichment of "WLRG" motif containing IgH sequences (**Fig. 4d**). We observed clonal expansion of the "WLRG" motif containing IgH sequences in COVID-19 patients by tracking the longitudinal samples collected from 4-28 days post symptom onset (**Fig. 4e and f**). It is worth mentioning that the "WLRG" motif containing IgH sequences present in the SARS-CoV-2-exposed antibody repertoires are mainly IgG isotype (encoded by IGHG, 79.3%). However, 65.2% of the "WLRG" motif containing IgH sequences detected in the SARS-CoV-2-naïve antibody repertoires are IgM isotype (encoded by IGHM) that is usually expressed by naïve B cells (**Fig. 4g and h**). Unlike heavy chains, the "QSYDSS" motif containing VL6-57 light chains can be readily detected in both the healthy and SARS-CoV-2-exposed IgL repertoires, highlighting abundance of VL6-57 transcripts in human B cell repertoires even under resting state (**Fig. 4i**).

To further determine the abundance of B cells expressing mAbs of VL6-57 light chains paired with the HCDR3 "WLRG" motif, we analyzed published single-B V(D)J sequences from SARS-CoV-2-exposed and -naïve individuals (Barmada et al., 2023; Dugan et al., 2021; Ferreira-Gomes et al., 2021; Ren et al., 2021; Zhang et al., 2020) and our previously reported FACS-sorted RBD-reactive single-B memory cell sequences from COVID-19 convalescents (Yu et al., 2023) (**Fig. 4a, see method**). Interestingly, only two VL6-57 B cells with paired "WLRG" motif (2 IGHM) were detected among 87808 B cells (0.02‰) from SARS-CoV-2-naïve donors (**Fig. 4j**). Following SARS-CoV-2 vaccination or infection, there is a remarkable expansion of B cells expressing VL6-57 antibodies with paired "WLRG" motif. We identified 20 VL6-57 B cells with paired "WLRG" motif (15 IGHG, 4 IGHM, and 1 IGHA) among 396211 B cells (0.05‰) from COVID-19 patients, while 14 (12 IGHG, 1 IGHM, and 1 IGHA) were identified among 29838 B cells (0.47‰) from COVID-19 vaccinees. Furthermore, 14 (13 IGHG and 1 IGHM) were found among the 4642 FACS-sorted RBD-reactive B cells (3.0‰) from COVID-19 convalescents (**Fig. 4j**). After FACS sorting, there is an approximately 60-fold (3‰/0.05‰) enrichment of VL6-57 B cells with paired "WLRG" motif among the RBD-reactive B cells compared with the unsorted single B cell sequencing data (**Fig. 4j**). Consistently, the VL6-57 B cells with paired "WLRG" motif identified from SARS-CoV-2-naïve individuals are mostly expressing IgM, while those from COVID-19 patients and vaccinees are predominantly expressing IgG (**Fig. 4j**). In summary, the above observations indicate that naïve B-cells expressing VL6-57 antibodies with paired "WLRG" motif were activated after SARS-CoV-2 exposure and underwent class switching and clonal expansion.

Antibodies utilizing VL6-57 light chains identified from SARS-CoV-2 unexposed individuals bind to SARS-CoV-2 RBD

To confirm whether the VL6-57 light chain (IgL) sequences identified from bulk antibody repertoires of SARS-CoV-2-naïve individuals (**Fig. 4c**) are able to react with SARS-CoV-2 RBD, five different "WLRG" motif-containing heavy chains encoded by IGHV3-7, IGHV3-11, IGHV4-31, IGHV4-39 and IGHV4-61 respectively from 5 different healthy donors are selected to pair with

a common VL6-57 light chain shared by the 5 healthy donors to generate 5 recombinational mAbs (namely H4, H5, H14, H16 and H18) (**Fig. S10a-c**). The selected heavy chains are predicted to bind SARS-CoV-2 RBD based on their HCDR3 sequences. All these heavy chains have none or only a few somatic mutations (**Fig. S10**). BLI assays confirmed that all of the 5 recombinational mAbs are able to bind to wildtype SARS-CoV-2 RBD with affinities (K_D) in the range of 30-1504 nM (**Fig. 5a, Table S3**). Among of them, H18 exhibits detectable cross reactivities with RBDs of SARS-CoV-1, Bat CoV RaTG13, and Pangolin CoV GD1, with affinities ranging between 134-332 nM (**Figs. 5b, S11, Table S3**). H18 has mild neutralization activity towards wildtype, Alpha, Beta and Delta SARS-CoV-2 authentic viruses (**Fig. 5c**), it also possesses activity to trigger conformational change of SARS-CoV-2 S and exhibits inhibition activity in the S-protein-ACE2 interaction mediated cell-cell membrane fusion assay (**Fig. 2d-e**).

To confirm whether H18 recognizes the same epitope with other VL6-57 mAbs, we obtained cryo-EM structures of wildtype SARS-CoV-2 S-trimer bound to H18 Fab (**Figs. 5d, S1b, S2a-c**). Similar to cryo-EM structures of SARS-CoV-2 S-trimer:R1-26 complexes (**Fig. 2a**), we observed 3:2, 3:3 (S-protomer:Fab) SARS-CoV-2 S-trimer:H18 complexes and head-to-head aggregate of two S-trimers each bound by 3 H18 Fabs. Unfortunately, we were not able to uncover high resolution information of the H18 epitope from these structures. To characterize fine characteristics of the H18 epitope, we generated a ternary complex sample consisting of wildtype SARS-CoV-2 S-trimer, H18 Fab and R1-32 Fab (He et al., 2022) (**Figs. 5e, S1c, S2a-c**). In the ternary complex sample, most S-trimer particles are bound by three H18 Fabs and three R1-32 Fabs in a 3-RBD "up" conformation (**Figs. 5e and S1c**). Although the cryo-EM sample was generated using a H18 concentration the same as that of R1-26 in the S-GSAS/6P:R1-26 sample, we observed a greater degree of S-trimer disintegration likely due to ternary complex formation by the addition of extra R1-32 Fab, structures of S1 bound to both H18 and R1-32 Fabs and dimers of S1 bound to both H18 and R1-32 Fabs were observed (**Fig. 5e**). Similar S-trimer disintegration has been observed previously in the presence of mAb R1-32 (He et al., 2022; Yu et al., 2023).

A 3.5 Å resolution structure of the H18 binding interface is derived from a focused refinement of the S1:H18:R1-32 Fab complex (**Figs. 5f, S2d**). Although the heavy chain genes are different between H18 (VH4-61) and R1-26 (VH3-7), H18 binds the identified convergent epitope with HCDR3 residues W102 and R104 via hydrophobic contact and cation-π interaction respectively. Consistent with our expectation, the R104 residue within the "WLRG" motif is interacting with Y94 within the LCDR3 "QSYDSS" motif via a cation-π interaction. Antigen binding by H18 HCDR3 is further stabilized by hydrogen bonds from LCDR1 similar to R1-26 (**Figs. 5f, S2d**). These observations confirm that VL6-57 light chain can pair with multiple heavy chains to target the identified convergent epitope on SARS-CoV-2 spike RBD.

Molecular basis of evasion from the shared VL6-57 antibody response by SARS-CoV-2 Omicron variants

During the structural analysis of VL6-57 antibodies we noticed that residues S371, S373 and S375 are within the convergent epitope in the ancestral SARS-CoV-2 spike RBD (**Figs. 3, S5**), these residues have been considered as a feature of Omicron variants and shown to alter virological behaviour of the virus (Kimura et al., 2022a; Pastorio et al., 2022). We also noted that among the 26 RBD substitutions observed in SARS-CoV-2 so far (**Fig. 6a**), 18 substitution positions are shared among SARS-related-CoVs (**Fig. 6b**) and among the other 8 substituted positions, S371, S375, T376, and R408 fall within the epitope of the VL6-57 class 4 public antibodies (**Fig. 6a-b**). Mapping of substituted RBD positions and VL6-57 antibody epitope areas indicates that the epitope is relatively conserved among SARS-related-CoVs. Although R1-26 and H18 could bind and neutralize SARS-CoV-2 WT and early VOCs, both of them fully lose binding and neutralization abilities towards Omicron BA.1 (**Figs. 1d and e, 5c**). BLI assays show that simultaneous rescue mutations at 371, 373 and 375 on Omicron BA.1 RBD (Omicron-(L371S+P373S+F375S)) are able to completely recover binding by VL6-57 mAbs R1-26 and H18 (**Fig. 6c, Table S4**). This result indicates that the evasion of VL6-57 mAb neutralization is specifically mediated by substitutions at S371, S373, and S375. BLI assays using a series of RBD mutants revealed that single S371L, S373P or S375F

mutation greatly reduces binding by R1-26 and H18. Combinations of any two simultaneous mutations (S373P/S375F, S371L/S375F, S371L/S373P) are able to almost completely abolish R1-26 and H18 binding (**Fig. 6c, Table S4**). Structural comparison of WT, Omicron BA.1 and BA.2 RBDs reveals that backbone conformations at residues 373 to 375 are affected by the S373P introduced proline and the phenylalanine sidechain introduced by S375F (**Fig. 6d**), these epitope structural changes likely abolish binding by the VL6-57 mAbs.

Discussion

We identify a class of VL6-57 light chain utilizing antibodies with the ability to pair with diverse heavy chains to target a convergent epitope defined by featured Omicron mutations - S371L/F, S373P, S375F. The identified epitope is cryptic in "down" RBD and strongly hydrophobic, consistent with that HCDR3s of VL6-57 mAbs share hydrophobic residues to interact with the epitope. Notably, within the epitope, residues S371-S373-S375 are located at the entrance to the identified fatty acid binding pocket of sarbecovirus spike RBD. Binding of linoleic acid within the pocket has been associated with "locked" spike conformation with potential functions in virus assembly (Qu et al., 2022; Toelzer et al., 2020; Toelzer et al., 2022; Xiong et al., 2020; Zhang et al., 2023). Being a highly dynamic protein, SARS-CoV-2 S-protein has been observed to adopt RBD "up" and "down" conformations. D614G substitution became fixed in the spike protein shortly after the SARS-CoV-2 pandemic (Grubaugh et al., 2020; Qu et al., 2022; Yurkovetskiy et al., 2020; Zhang et al., 2021a), it has been shown that this substitution increased spike stability and a shift towards more RBD "up" open S-trimer has been observed by multiple studies (Qu et al., 2022; Yang et al., 2021; Zhang et al., 2021a). We speculate that the change in spike dynamics may have increased immune pressure posed by mAbs targeting cryptic epitopes. Indeed, we find that binding of R1-26, H18 and possibly other VL6-57 mAbs to the convergent cryptic epitope not only blocks ACE2 binding (Piccoli et al., 2020), but also prematurely triggers spike conformational change resulting in spike destruction. When paired with other RBD-specific antibodies, binding of antibodies to cryptic epitopes may also promote spike disintegration (He et al., 2022; Yu et al., 2023). The increased immune pressure may promote

the introduction of S371L/F-S373P-S375F in Omicron variants. The identified cryptic epitope of VL6-57 mAbs is relatively conserved among S-proteins of sarbecoviruses. Before the emergence of Omicron variants, the VL6-57 antibody epitope remains unchanged in early SARS-CoV-2 VOCs or VOIs. We and others also found that some of the VL6-57 mAbs including H18 and many others (Cao et al., 2023) are able to cross-react with SARS-CoV-1 RBD, suggesting that the VL6-57 epitope is relatively conserved among sarbecoviruses, therefore, the introduction of S371L/F-S373P-S375F demonstrates an extraordinary adaptability of SARS-CoV-2, providing a piece of evidence to suggest that unusual changes of a pandemic virus can occur in response to strong population immune pressures including those posed by population antibodies. Studies have shown that single mutation of S371F/L, S373P or S375F strongly impairs spike processing and spike mediated virus entry (Pastorio et al., 2022; Saito et al., 2022). Among them, S375F has the most detrimental effect (Kimura et al., 2022b), and notably, mutation at this position has not been observed among animal sarbecovirus S-proteins (**Fig. 6b**). When the 3 serine substitutions are combined the ability of the S-protein to mediate virus-cell entry is altered and it has been proposed that such change is associated with change of tissue tropism by Omicron variants and their attenuation (Pastorio et al., 2022; Willett et al., 2022). The above observations highlight the likely necessity of the S371-S373-S375 mutations for SARS-CoV-2 immune evasion even at the expense of drastic S-protein function change with associated impact on virus behaviour. It has been shown that antigenic changes can alter SARS-CoV-2 receptor binding (Niu et al., 2021), for other viruses, antigenic evolution of the pandemic H3N2 virus was found to significantly weaken hemagglutinin (HA) sialic acid binding affecting virological phenotypes (Lin et al., 2012). Therefore, antigenic change of viral surface protein can have complex consequences and a delicate balance between antibody evasion and transmissibility is pursued by circulating viruses (Carabelli et al., 2023).

It has been reported that SARS-CoV-2 infections induce several classes of germline antibodies of specific VH genes. Due to similarity of equivalent germline antibody genes among humans, induced classes of germline antibodies are widely present in the population. Certain germline antibodies of the same

VH gene can bind a convergent spike epitope with similar molecular interactions mediated by antibody residues encoded by germline sequences (Chen et al., 2021; Nielsen et al., 2020; Robbiani et al., 2020; Zhang et al., 2022). Similar phenomena have also been observed for mAbs induced by infections of other viruses (Ehrhardt et al., 2019; Jackson et al., 2014; Parameswaran et al., 2013). A common feature of these germline antibodies is that essential paratope interactions are primarily mediated by antibody heavy chains, such feature has been observed for VH3-53 (Yuan et al., 2020), VH1-58 (Dong et al., 2021), VH1-2 (Rapp et al., 2021; Yuan et al., 2021), VH1-69 (He et al., 2022), and VH1-24 (Cerutti et al., 2021) SARS-CoV-2 S-specific antibodies. These heavy chain mediated shared antibody responses have been thought to be a key driving force of SARS-CoV-2 antigenic drift (Clark et al., 2021; Patel et al., 2023; Yan et al., 2022). In this study, by contrast, the identified population antibody response is mediated by antibodies with shared usage of VL6-57 light chain gene. Generally, heavy chains dominate interactions in antibody-antigen complexes, while light chains help modulating heavy chain conformation (Xiao et al., 2019). It has been shown that antibodies of VH1-69 heavy chains can pair with irrelevant light chains to bind influenza HAs (Pappas et al., 2014; Sui et al., 2009). Our structural study reveals that interactions between the VL6-57-encoded mAbs and RBD are mainly mediated by germline gene encoded residues in HCDR3 and LCDR1, while germline gene encoded LCDR3 residues support HCDR3 in antigen binding. Although VL6-57 is most frequently observed to pair with VH3-7, VH4-39, and VH4-59 we provide evidence that many other heavy chains can also pair with VL6-57 to bind SARS-CoV-2 RBD. Therefore, VL6-57 germline sequence provides an efficient framework to allow pairing with diverse heavy chains to generate SARS-CoV-2 RBD-targeting antibodies, such unique property of VL6-57 gene increases the probability of VL6-57 class of public antibodies being induced across a broad population. This study provides a rare example that a light chain paired with various heavy chains can largely dictate the binding mode of a class of antiviral antibodies.

Given the epitope of the VL6-57 mAbs is largely conserved among animal sarbecoviruses, we speculate that the versatility of VL6-57 in generating SARS-

CoV-2 RBD antibodies may be unique within the human population. Wide induction of this class of antibodies contribute to selection pressure on SARS-CoV-2 within the human population, but similar response may be absent in other species. Therefore, S371L/F-S373P-S375F represents a specific constellation of immune escape mutations for SARS-CoV-2 adaptation and transmission within the human population. However, it is worth noting that there is still VL6-57 mAb that is resistant to Omicron variants (Huang et al., 2023), implying adaptation of antibody immunity to viral mutations. Therefore, exploring the co-evolution of virus and antibody immunity in future works should be beneficial to understanding co-adaptability between virus and host (Wu et al., 2015). Overall, this study provides a fresh example of light-chain mediated populational antibody immune pressure against SARS-CoV-2 at molecular level. Our findings further support the hypothesis that convergent antibody responses within the population drive viral antigenic drift leading to emergence of new SARS-CoV-2 variants.

Methods

Cells and viruses

Expi293F cells (Thermo Fisher Scientific, A14527) were maintained in Expi293F Expression medium (Thermo Fisher Scientific, A1435101) at 37 ℃ by shaking at 120 rpm under a humidified atmosphere with 8% CO_2. The human embryonic kidney (HEK) 293T (ATCC, CRL-3216) and Vero E6 (ATCC, CRL-1586) were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FBS. The authentic SARS-CoV-2 viruses, including WT (Wuhan-Hu-1), Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2), and Omicron (BA.1) were isolated from COVID-19 patients and preserved in Guangzhou Customs District Technology Center BSL-3 Laboratory. The SARS-CoV-2 Delta strain was a gift from Guangdong Provincial Center for Disease Control and Prevention, China. Experiments related to authentic SARS-CoV-2 viruses were conducted in Guangzhou Customs District Technology Center BSL-3 Laboratory.

Expression of monoclonal antibody

The antibody heavy- and light-V genes (VH/VL) were cloned into human IgG1 expression vectors using Clone Express II One Step Cloning Kit (Vazyme, China). When density of HEK293F cells reached 1×10^6 cells/mL, equal amounts of heavy- and light-chain plasmids were transfected into HEK293F cells using EZ cell transfection reagent (Life-iLab Biotech, China). Following transfection, HEK293F cells were cultured in CD 293 TGE medium (ACRO, China) containing 10% CD Feed X supplement (ACRO, China) at 37 °C in a humidified atmosphere with 5% CO_2 and shaking at 120 rpm. 6 days post transfection, supernatants were harvested and clarified by centrifugation. Supernatants were filtered through 0.22-μm filters (Merck Millipore, Germany) before incubation with Protein A Resin (Genscript, China) at room temperature for 2 h for antibody affinity purification. After washing, antibodies were eluted from the Protein A Resin using 0.1 M Na-Citrate (pH 3.25) and eluents were neutralized immediately with 1 M Tris-HCl (pH 8.8). Antibodies were

607 concentrated in PBS using Amicon Ultrafilter devices (Merck Millipore, USA)
608 and stored at −80 °C.
609
610 **Inhibition of SARS-CoV-2 S-mediated cell-cell fusion**
611 As previously described (Ma et al., 2023), we utilized a real-time, quantitative
612 live cell split-GFP fluorescence complementation phenotypic assay to analyze the
613 fusion inhibition activity of antibodies. Briefly, effector cells (HEK293T cells
614 expressing SARS-CoV-2 S and GFP1-10 proteins) were seeded into 96-well-
615 plate (5×10^4 cells per well in DMEM). Each well was cultured in the presence
616 or absence of a test mAb at the indicated concentrations for 2 h at 37 °C.
617 Subsequently, target cells (hACE2-293T cells expressing GFP11 protein, 5×10^4
618 cells per well in DMEM) were added uniformly to each well and fluorescence
619 images were taken 2 h after the addition of target cells using a Nikon fluorescence
620 microscope or a Keyence BZ-X800 all-in-one fluorescence microscope. The GFP
621 area in each well was quantified on Image J, and percentage inhibition of cell-
622 cell fusion was calculated using the following formula: $(1 - (E - N) / (P - N)) \times$
623 100%. “E” represents the GFP area in the experimental group. “P” represents the
624 GFP area in the positive control group, where no mAb was added. “N” is the GFP
625 area in the negative control group, where effector HEK293T cells only expressing
626 GFP1-10. Samples were tested in triplicate, and experiments were repeated at
627 least twice.
628
629 **Protein expression and purification**
630 The extracellular domain of SARS-CoV-2 spike (S) protein (residues 14-1211)
631 with an N-terminal mu-phosphatase signal peptide, a “R” substitution at the
632 multibasic furin cleavage site (R682-R685), and a C-terminal TEV protease
633 cleavage site followed by a T4 fibritin trimerization motif and a His_6 tag was
634 cloned into the mammalian expression vector pCDNA3.1(+), named “S-R”, as
635 previously described (Walls et al., 2019). To generate the stabilized S-protein

for cryo-EM sample preparation, the sequence of S protein was modified with six prolines at residues 817, 892, 899, 942, 986, and 987 and the furin cleavage site was changed to "GSAS" (Hsieh et al., 2020), named "S-GSAS/6P". To express the SARS-CoV-2 RBD, residues 319-541 of S-protein were fused with an N-terminal mu-phosphatase signal peptide and a C-terminal 6×His tag before the sequence was inserted into the pCDNA3.1(+) vector. S-protein or RBD expression vector was transiently transfected into Expi293F using polyethylenimine. Proteins were purified using IMAC (immobilized metal affinity chromatography) following previously described protocols (He et al., 2022; Xiong et al., 2020; Yu et al., 2023). All proteins were aliquoted, flash-frozen in liquid nitrogen and stored at -80 ℃.

Biolayer interferometry

Binding assays were carried out on an Octet RED96 instrument (Sartorius) using a previously established protocol (He et al., 2022). Briefly, each Protein A biosensors (Sartorius) was pre-equilibrated in PBST buffer (PBS, pH 7.4, 0.02% Tween-20, 1 mg/ml BSA) for 10 min. Subsequently, IgG at 11 μg/ml was loaded onto each biosensor to a response level between 1.6-1.8 nm. The IgG immobilized biosensors were submerged into twofold serially diluted (200-3.125 nM) RBD or S-protein solutions for 300 s to record association. The biosensors were subsequently submerged into PBST buffer for 600 s to record dissociation. For IgG binding to the generated RBD mutants, biosensors immobilized with IgGs were monitored for association (300 s) in RBD solutions at a fixed concentration of 200 nM before the sensors were submerged into PBST to monitor dissociation (600 s). IgG immobilized sensors were also submerged into PBST buffer to record references. Data were reference-subtracted and analyzed using Data Analysis HT v12.0.2.59 software (Sartorius) with a 1:1 fitting model for binding to RBDs and 2:1 fitting model for binding to S-trimers. Raw data and fits were plotted in GraphPad Prism v8.0.

Competition assays were performed on a Gator label-free bioanalysis system (GatorBio). 2 μg/ml of SARS-CoV-2 RBD (Sino Biological, Cat# 40592-V08B) was immobilized onto the pre-equilibrated Anti-His biosensors (GatorBio). Biosensors were statured with the first antibody for 300 s before submerging into the second antibody or ACE2 solutions for 200 s. Data were analyzed by the Gator data analysis software (GatorBio) and plotted in GraphPad Prism v8.0.

Ligand-induced conformational change of spike protein

S-R diluted to 1 mg/ml (7.09 μM) was incubated with ACE2-Fc or antibodies in IgG or Fab form at a 1:1.1 molar ratio at room temperature for 1 h. The samples were subsequently treated with 50 μg/ml proteinase K at 4 ℃ for 30 min. Each sample was boiled in 5 × non-reducing SDS loading buffer at 98 ℃ for 5 min to stop the reaction. Samples were separated by SDS-PAGE on 4-12% gradient gels, before protein bands were transferred onto a polyvinylidene difluoride membrane using a semi-dry blotting system. The membrane was blocked with 5% milk in PBST before the membrane was incubated with a primary antibody (rabbit anti-SARS-CoV-2 S2 polyclonal antibody, Sino Biological, Cat# 40590-T62, 1:2500 dilution) in PBST. After extensive washing, the membrane was incubated with the secondary antibody (horseradish peroxidase-conjugated goat anti-rabbit IgG, Beyotime, Cat# A0208, 1:2500 dilution) in PBST. Finally, protein blots were visualized by chemiluminescence using a Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific, Cat# 32106).

Cryo-EM sample preparation and data collection

To generate S-GSAS/6P:R1-26 Fab, S-GSAS/6P:H18 Fab complexes, S-GSAS/6P at 4.43 mg/ml was incubated with R1-26 Fab or H18 Fab at a 1:1 molar ratio. To generate S-GSAS/6P:H18 Fab:R1-32 Fab complex, S-GSAS/6P at 4.43 mg/ml was incubated with H18 and R1-32 Fabs at a 1:1:1 molar ratio.

After 1 min incubation at room temperature, each 3 µl sample was supplemented with 0.1% octyl-glucoside (Sigma-Aldrich, Cat# V900365) before it was applied onto a 300-mesh holey carbon-coated copper grid (Quantifoil, Cu R1.2/R1.3) pre-treated by glow-discharging at 15 mA for 30 s. Each grid was blotted for 2.5 s with a blot force of 4 at 22 °C and 100% humidity before plunge-freezing in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). The S-GSAS/6P:R1-26 complex cryo-grid was imaged in a Titan Krios electron microscope (Thermo Fisher Scientific) operating at 300 kV and equipped with Gatan BioQuantum energy filter (slit width 20 eV) and Post-GIF Gatan K3 Summit direct detection camera. Movie stacks were automatically recorded using EPU at a nominal magnification of ×81,000 in super-resolution mode with a calibrated pixel size of 0.5475 Å and nominal defocus values ranged between -0.8 to -2.0 µm. Each stack was fractionated into 38 frames and exposed at a dose rate of 25 e^-/pixel/s for 2.4 s resulting in a total dose of ~ 50 $e^-/Å^2$. The S-GSAS/6P:H18 complex cryo-grid was imaged in a Talos Arctica electron microscope (Thermo Fisher Scientific) operating at 200 kV. Using the SerialEM v3.8.7 software, movie stacks were recorded at a nominal magnification of ×45,000 on a K3 direct detection camera (Gatan) in super-resolution mode with a calibrated pixel size of 0.44 Å with nominal defocus values ranged between -0.8 to -2.5 µm. Each stack was fractionated into 27 frames and exposed at a dose rate of 24.4 e^-/pixel/s for 1.89 s, resulting in a total dose of ~ 60 $e^-/Å^2$. The S-GSAS/6P:H18:R1-32 complex cryo-grid was imaged in a 300 kV Titan Krios electron microscope (Thermo Fisher Scientific) equipped with a SelectrisX energy filter (slit width 10 eV) and a Falcon 4 direct electron detector. Movie stacks were automated collected using EPU software with the electron event representation (EER) mode at a nominal magnification of ×130,000 with a pixel size of 0.93 Å and nominal defocus values ranged between -0.6 to -2.0 µm. Each stack was recorded and exposed at a dose rate of 7.51 e^-/pixel/s for 5.79 s resulting a total dose of ~ 50

723 $e^{-}/Å^2$. These settings yielded EER stacks consisting of 199 frames. All movie
724 stacks were imported into cryoSPARC live (v3.3.2) (Kucukelbir et al., 2014) for
725 pre-processing, which includes patched motion correction, contrast transfer
726 function (CTF) estimation and bad images rejection. Movie stacks of S-
727 GSAS/6P:R1-26 complex and S-GSAS/6P:H18 complex datasets were binned
728 2× resulting in pixel sizes of 1.095 Å and 0.88 Å, respectively.
729
730 **Cryo-EM data processing**
731 Data processing was carried out using cryoSPARC v3.3.2/v4.2.0. After bad
732 image removal, particles were picked by blob-picking on 4343 S-GSAS/6P:R1-
733 26, 4822 S-GSAS/6P:H18 and 9861 S-GSAS/6P:H18:R1-32 images. After 2D
734 Classification, several good 2D classes were selected as the templates for
735 template-picking, resulting in initial datasets of 2,359,986 particles for S-
736 GSAS/6P:R1-26; 2,357,740 particles for S-GSAS/6P:H18; and 3,335,676
737 particles for S-GSAS/6P:H18:R1-32. Template-picked particles were extracted
738 for two or more rounds of 2D Classification to remove contaminants and low-
739 quality particles. For S-GSAS/6P:H18:R1-32, particles with intact S-trimers
740 were re-extracted with a larger box size before the last round of 2D
741 Classification. After several rounds of 2D Classification, well-aligned particles
742 with intact S-protein features were selected and subjected to Ab-initio
743 Reconstruction to generate 3D models. Particles of initial models with complete
744 S2 but incomplete S1 were further classified using Ab-initio Reconstruction
745 with class similarity value 0.3 for the S-GSAS/6P:R1-26 dataset, 0.7 for the S-
746 GSAS/6P:H18 and S-GSAS/6P:H18:R1-32 datasets to separate different
747 conformations.
748
749 3:2 and 3:3 S-protomer:Fab structures were observed in both S-GSAS/6P:R1-26
750 and S-GSAS/6P:H18 datasets. Head-to-head aggregates of 2 S-trimers bound by
751 3 Fabs, giving S-protomer:Fab stoichiometry of 6:6, were observed in 2D

Classifications and Ab-Initio Reconstructions of both S-GSAS/6P:R1-26 and S-GSAS/6P:H18 datasets. Particles of aggregate classes were selected and combined before particles were re-extracted with a larger box size. A S-trimer aggregate structure from the S-GSAS/6P:R1-26 dataset was obtained by a round of 2D Classification followed by Ab-Initio Reconstruction. Similarly, 2 S-trimer aggregate structures were obtained from the S-GSAS/6P:H18 dataset. To obtain a higher resolution structure of the interface between RBD and R1-26 Fab, particles of the 3:3 S-GSAS/6P:R1-26 structure were refined again with applied C3 symmetry by Non-uniform Refinement. Refined particles were subjected to Symmetry Expansion and Density Subtraction before a Local Refinement was carried out focusing on a region containing RBD and Fab-V.

For the S-GSAS/6P:H18:R1-32 dataset, one 3:2:2 and two 3:3:3 structures were observed. In the S-GSAS/6P:H18:R1-32 dataset, many well-defined 2D classes showing incomplete S-protein features were selected for additional Ab-Initio Reconstruction. One 1:1:1 S1:H18:R1-32 structure and one 2:2:2 S1:H18:R1-32 structure consisting of two head-to-head 1:1:1 S1:H18:R1-32 complexes were observed. Particles of these two structures were re-extracted with a smaller box size. Final maps of differently conformations were reconstructed using Non-uniform Refinement. All maps with resolutions higher than 4 Å were processed by a second round of Non-uniform Refinement with defocus refinement and global CTF refinement to improve map quality. All resolutions were estimated at the 0.143 criterion of the phase-randomization-corrected Fourier shell correlation (FSC) curve calculated between two independently refined half-maps multiplied by a soft-edged solvent mask in RELION v4.0 **(**Kucukelbir et al., 2014**)**. Additional data processing details are summarized in **Figs. S1 and S2**, and **Table S2**.

Model building and analysis

A previously determined structure of SARS-CoV-2 S-trimer in complex with 3 R1-32 Fabs and 3 ACE2s (PDB: 7YEG) (He et al., 2022) or a structure of SARS-CoV-2 S1 in complex with a YB9-258 Fab and an R1-32 Fab (PDB: 8HC5) (Yu et al., 2023) was used as starting model. An antibody light chain (PDB: 7D6I) was used as the starting model for R1-26 and H18 Fab light chains. Starting models for R1-26 and H18 Fab heavy chains were generated from heavy chain structures in PDBs - 5X8M and 7VSU respectively. Starting models were fitted into final refined maps in UCSF Chimera v1.14 (Pettersen et al., 2004). Iterative model building and real space refinement were carried out in Coot v0.9.6 (Emsley et al., 2010) and PHENIX v1.20.1-4487 (Afonine et al., 2018). Model refinement statistics are summarized in **Tables S2**. Interface analyses were performed in QtPISA v2.1.0 (Krissinel, 2015). Structure figures were generated in UCSF Chimera v1.14.

SARS-CoV-2 authentic virus neutralization assay

Antibodies were serial diluted with DMEM and mixed with 200 focus forming unit (FFU) Wuhan-Hu-1 (wildtype), Alpha, Beta, Delta, or Omicron BA.1 authentic SARS-CoV-2 viruses. After incubation at 37 °C for 1 h, antibody-virus mixtures were added to a 96-well plate cultured with Vero E6 cells and incubated at 37 °C in 5% CO_2 for 1 h. After removing the inoculum, each well was overlaid with 100 μL 1.6% carboxymethylcellulose warmed to 37 °C. After culturing for 24 h, overlays were removed and the cells were fixed with 4% paraformaldehyde (Biosharp, China) and permeabilized with 0.2% Triton X-100 (Sigma, USA). Cells were incubated with a human anti-SARS-CoV-2 nucleocapsid protein monoclonal antibody (obtained by laboratory screening) at 37 °C for 1 h. After washing with 0.15% PBST three times, cells were incubated with an HRP-labeled goat anti-human secondary antibody (Cat. No.: 609-035-213, Jackson ImmunoResearch Laboratories) at 37 °C for 1 h. Plates were washed with 0.15% PBST three times before the foci were visualized by

TrueBlue Peroxidase Substrate (KPL, Gaithersburg, MD), and counted with an ELISPOT reader (Cellular Technology Ltd. Cleveland, OH). The foci reduction neutralization test titre (FRNT50) was calculated by the Spearman-Karber method.

Analysis of the structurally characterized SARS-CoV-2 S-specific mAbs

A total of 376 antibody:RBD/S-protein complex structures were curated from the PDB (https://www.rcsb.org/) before May 18, 2023. Epitope residues, paratope residues, and buried surface area (BSA) were determined or calculated using the PDBe PISA server (https://www.ebi.ac.uk/msd-srv/prot_int/). BSA for each epitope residue is assigned as a feature of a certain antibody and used to construct a feature matrix M_{AxB} for downstream analysis, where A is the number of antibodies and B is the number of features (amino acid length of spike: 1273). Therefore, a 376×1273 BSA matrix was obtained, which was subsequently used as input for epitope classification with the R package UMAP (v0.2.9.0). Uniform Manifold Approximation and Projection (UMAP) is an algorithm for dimensional reduction. After the reduction analysis, clustering was performed using the K-means algorithm. UMAP and K-means clusterings were conducted independently. Antibodies clustered into the same cluster as the well-documented mAb CR3022 are defined as class 4 antibodies. After this workflow, 35 class 4 mAbs were identified among the 376 structurally characterized S-specific mAbs (**Fig. S5**). In addition, we found that class 4 epitope is largely overlapping with the F2 epitope which was previously defined by deep mutational scanning (Cao et al., 2022). Thus, a total of 69 F2 epitope mAbs were also included for germline gene usage analysis. Germline gene usages of these structurally characterized S-specific mAbs were inferred using IMGT/V-QUEST (https://www.imgt.org/). Visualization of the germline heavy and light gene usage and pairing among the 35 class 4 and the 69 F2 antibodies

were performed in R platform v4.2 using the R package circlize v0.4.10 (Gu et al., 2014).

Bioinformatic analysis

To investigate features of VL6-57 utilizing mAbs, a total of 290 VL6-57 utilizing mAbs were extracted from the CoV-AbDab database (Raybould et al., 2021) for downstream analysis. Analysis of germline gene usage and CDR3 length were perform using the built-in function in R platform v4.2. CDR3 amino acid composition analysis was performed using the R package ggseqlogo (Wagih, 2017). VL6-57 utilizing mAbs with 12-AA length HCDR3 were used for lineage structure reconstruction. R package igraph was employed for visualization of the lineage structure of VL6-57 utilizing mAbs with 12-AA length HCDR3s. After the above workflow, the results showed that WLRG motif within 12-AA long HCDR3 and QSYDSS motif within LCDR3 are the convergent signatures of the VL6-57 mAbs. To determine the occurrence of VL6-57 mAbs with paired HCDR3 “WLRG” motif, we analysed nearly 2 billion IgH and IgL sequences from 3 previously described datasets (Niu et al., 2020; Yan et al., 2021; Zhang et al., 2022). These datasets have been deposited by us in the National Genomics Data Center (https://bigd.big.ac.cn/), China National Center for Bioinformation (CNCB) under accession numbers PRJCA003775, PRJCA007067, and PRJCA017560. HCDR3 “WLRG” motif-containing IgH sequences were defined as those sequences which have 12-AA long HCDR3s with a “WLRG” motif. VL6-57 IgL sequences were defined as those sequences that use the germline gene VL6-57 with 9 to10-AA long LCDR3s, and encoding a “QSYDSS” motif within LCDR3 region. The abundance of “WLRG” motif-containing IgH sequences or VL6-57 IgL sequences were normalized by counts per million. The clonal expansion of B cells expressing “WLRG” motif-containing IgH or VL6-57 IgL sequences in COVID-19 patients after SARS-CoV-2 infection were tracked. The divergence

from germline genes of all IgH or IgL sequences was equal to somatic hypermutations, and sequence identity to queried sequence was calculated using the R package Biostrings v2.60.2 (http://bioconductor.org/packages/release/bioc/html/Biostrings.html). All IgH or IgL sequences were plotted as a function of sequence somatic hypermutations (x axis) and sequence identity (y axis) to heavy or light chains of R1-26, R1-30, R2-3, R2-6 or R2-7 with a colour gradient indicating sequence density. To determine the frequency of B cells expressing VL6-57 mAbs with paired HCDR3 "WLRG" motif, we analysis the single-B V(D)J sequences from SARS-CoV-2-exposed and -unexposed individuals (Barmada et al., 2023; Dugan et al., 2021; Ferreira-Gomes et al., 2021; Ren et al., 2021; Zhang et al., 2020) and FACS-sorted RBD+ B cells from COVID-19 convalescents (Yu et al., 2023). These datasets are available from the Gene Expression Omnibus (GEO) database under accession numbers: GSE230227, GSE171703, GSE158038, GSE158055; http://www.microbiome-bigdata.com/project/SARS-CoV-2/; and National Genomics Data Center (https://bigd.big.ac.cn/) under the accession number: PRJCA012020. B cells expressing VL6-57 mAbs with paired HCDR3 "WLRG" motif were defined as those B cells expressing heavy chains with a 12-AA long HCDR3 containing a "WLRG" motif and VL6-57 light chains with a "QSYDSS" motif in LCDR3.

Data availability

Cryo-EM density maps for the structures of R1-26 or H18 Fab in complex with S-trimer or S1 fragment have been deposited in the Electron Microscopy Data Bank (EMDB) with accession codes XXXX. Related atomic models have been deposited in the Protein Data Bank (PDB) under accession codes XXXX, respectively.

Acknowledgements

We thank the staffs at cryo-EM Facilities of GIBH-CAS, Guangzhou

Laboratory Bio-imaging Technology Platform and SUSTech for help on cryo-EM sample preparation and data collection. This study was supported by the National Natural Science Foundation of China (82201932 to Q.Y., 92269201 to L.C.), the Emergency Key Program of Guangzhou Laboratory (EKPG21-06 to X.X.), the R&D Program of Guangzhou Laboratory (SRPG22-002 to X.X.), the Natural Science Fund of Guangdong Province (2021A1515011289 to X.X.), China Postdoctoral Science Foundation (2022M710891 to Q.Y.), State Key Laboratory of Respiratory Disease (SKLRD-Z-202324 to Q.Y.), X.X. acknowledges start-up grants from the Chinese Academy of Sciences.

Contributions

Q.Y. and X.X. conceived the study. Q.Y., R.H. and P.H. conceived and initiated the antibody isolation; X.X. conceived and initiated the structural studies; R.H. and P.H. isolated antibodies with assistance from X.H., H.L., X.C. and X.N.; P.H. and R.H. expressed and purified antibodies; P.H., B.L., R.H. and Q.C. performed BLI assays; Y.M. performed the SARS-CoV-2 S-mediated cell-cell fusion assays; YZ performed authentic virus neutralization assays; B.L. and Q.C. purified spikes for cryo-EM and other experiments using constructs and protocols developed by X.X.; B.L. performed western blots to assay spike structural change; Q.Y. performed the bioinformatics analysis with assistance from X.X. and Y Z; B.L. and X.G. collected cryo-EM data under the supervision of J.H. and X.X.; X.G., Z.L. and J.W. processed cryo-EM data under the supervision of J.H. and X.X.; X.X., X.G., Q.Y. and J.H. analysed cryo-EM structures with assistance from B.L. and J.W.; Q.Y., B.L., X.G., R.H. and Y.M. prepared the figures under the supervision of X.X.; X.X. and Q.Y. wrote the paper with input from all co-authors; X.X., Z.J., J.H. and L.C. acquired funding and supervised the research.

Competing interests

The authors declare no competing interests.

Reference

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

Barmada, A., Klein, J., Ramaswamy, A., Brodsky, N.N., Jaycox, J.R., Sheikha, H., Jones, K.M., Habet, V., Campbell, M., Sumida, T.S.*, et al.* (2023). Cytokinopathy with aberrant cytotoxic lymphocytes and profibrotic myeloid response in SARS-CoV-2 mRNA vaccine-associated myocarditis. Science Immunology *8*.

Boni, M.F. (2008). Vaccination and antigenic drift in influenza. Vaccine *26 Suppl 3*, C8-14.

Cao, Y., Jian, F., Wang, J., Yu, Y., Song, W., Yisimayi, A., Wang, J., An, R., Chen, X., Zhang, N.*, et al.* (2023). Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. Nature *614*, 521-529.

Cao, Y., Yisimayi, A., Jian, F., Song, W., Xiao, T., Wang, L., Du, S., Wang, J., Li, Q., Chen, X.*, et al.* (2022). BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. Nature.

Carabelli, A.M., Peacock, T.P., Thorne, L.G., Harvey, W.T., Hughes, J., Consortium, C.-G.U., Peacock, S.J., Barclay, W.S., de Silva, T.I., Towers, G.J.*, et al.* (2023). SARS-CoV-2 variant biology: immune escape, transmission and fitness. Nat Rev Microbiol *21*, 162-177.

Cerutti, G., Guo, Y., Zhou, T., Gorman, J., Lee, M., Rapp, M., Reddem, E.R., Yu, J., Bahna, F., Bimela, J.*, et al.* (2021). Potent SARS-CoV-2 neutralizing antibodies directed against spike N-terminal domain target a single supersite. Cell Host Microbe *29*, 819-833 e817.

Chen, E.C., Gilchuk, P., Zost, S.J., Suryadevara, N., Winkler, E.S., Cabel, C.R., Binshtein, E., Chen, R.E., Sutton, R.E., Rodriguez, J.*, et al.* (2021). Convergent antibody responses to the SARS-CoV-2 spike protein in convalescent and vaccinated individuals. Cell Rep *36*, 109604.

Clark, S.A., Clark, L.E., Pan, J., Coscia, A., McKay, L.G.A., Shankar, S., Johnson, R.I., Brusic, V., Choudhary, M.C., Regan, J.*, et al.* (2021). SARS-CoV-2 evolution in an immunocompromised host reveals shared neutralization escape mechanisms. Cell.

Dejnirattisai, W., Huo, J., Zhou, D., Zahradnik, J., Supasa, P., Liu, C., Duyvesteyn, H.M.E., Ginn, H.M., Mentzer, A.J., Tuekprakhon, A.*, et al.* (2022). SARS-CoV-2 Omicron-B.1.1.529 leads to widespread escape from neutralizing antibody responses. Cell *185*, 467-484 e415.

Deng, X., Garcia-Knight, M.A., Khalid, M.M., Servellita, V., Wang, C., Morris, M.K., Sotomayor-González, A., Glasner, D.R., Reyes, K.R., Gliwa, A.S.*, et al.* (2021). Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. Cell *184*, 3426-3437.e3428.

Dong, J., Zost, S., Greaney, A., Starr, T.N., Dingens, A.S., Chen, E.C., Chen, R., Case, B., Sutton, R., Gilchuk, P.*, et al.* (2021). Genetic and structural basis for recognition of SARS-CoV-2 spike protein by a two-antibody cocktail. bioRxiv.

963 Dugan, H.L., Stamper, C.T., Li, L., Changrob, S., Asby, N.W., Halfmann, P.J., Zheng, N.Y.,
964 Huang, M., Shaw, D.G., Cobb, M.S*., et al.* (2021). Profiling B cell immunodominance after
965 SARS-CoV-2 infection reveals antibody evolution to non-neutralizing viral targets. Immunity
966 *54*, 1290-1303 e1297.

967 Ehrhardt, S.A., Zehner, M., Krahling, V., Cohen-Dvashi, H., Kreer, C., Elad, N., Gruell, H.,
968 Ercanoglu, M.S., Schommers, P., Gieselmann, L*., et al.* (2019). Polyclonal and convergent
969 antibody response to Ebola virus vaccine rVSV-ZEBOV. Nat Med *25*, 1589-1600.

970 Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of
971 Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

972 Ferreira-Gomes, M., Kruglov, A., Durek, P., Heinrich, F., Tizian, C., Heinz, G.A., Pascual-
973 Reguant, A., Du, W., Mothes, R., Fan, C*., et al.* (2021). SARS-CoV-2 in severe COVID-19
974 induces a TGF-beta-dominated chronic immune response that does not target itself. Nat
975 Commun *12*, 1961.

976 Grubaugh, N.D., Hanage, W.P., and Rasmussen, A.L. (2020). Making Sense of Mutation:
977 What D614G Means for the COVID-19 Pandemic Remains Unclear. Cell *182*, 794-795.

978 Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize Implements and enhances
979 circular visualization in R. Bioinformatics *30*, 2811-2812.

980 Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M.,
981 Ludden, C., Reeve, R., Rambaut, A., Peacock, S.J*., et al.* (2021). SARS-CoV-2 variants, spike
982 mutations and immune escape. Nature Reviews Microbiology *19*, 409-424.

983 He, P., Liu, B., Gao, X., Yan, Q., Pei, R., Sun, J., Chen, Q., Hou, R., Li, Z., Zhang, Y*., et al.* (2022).
984 SARS-CoV-2 Delta and Omicron variants evade population antibody response by mutations
985 in a single spike epitope. Nat Microbiol *7*, 1635-1649.

986 Hsieh, C.L., Goldsmith, J.A., Schaub, J.M., DiVenere, A.M., Kuo, H.C., Javanmardi, K., Le, K.C.,
987 Wrapp, D., Lee, A.G., Liu, Y*., et al.* (2020). Structure-based design of prefusion-stabilized
988 SARS-CoV-2 spikes. Science *369*, 1501-1505.

989 Huang, K.A., Chen, X., Mohapatra, A., Nguyen, H.T.V., Schimanski, L., Tan, T.K., Rijal, P.,
990 Vester, S.K., Hills, R.A., Howarth, M*., et al.* (2023). Structural basis for a conserved
991 neutralization epitope on the receptor-binding domain of SARS-CoV-2. Nat Commun *14*,
992 311.

993 Huo, J., Dijokaite-Guraliuc, A., Nutalai, R., Das, R., Zhou, D., Mentzer, A.J., Fry, E.E.,
994 Mongkolsapaya, J., Ren, J., Stuart, D.I*., et al.* (2022). Humoral responses against SARS-CoV-2
995 Omicron BA.2.11, BA.2.12.1 and BA.2.13 from vaccine and BA.1 serum. Cell Discovery *8*, 1-4.

996 Jackson, K.J., Liu, Y., Roskin, K.M., Glanville, J., Hoh, R.A., Seo, K., Marshall, E.L., Gurley, T.C.,
997 Moody, M.A., Haynes, B.F*., et al.* (2014). Human responses to influenza vaccination show
998 seroconversion signatures and convergent antibody rearrangements. Cell Host Microbe *16*,
999 105-114.

Kimura, I., Yamasoba, D., Nasser, H., Zahradnik, J., Kosugi, Y., Wu, J., Nagata, K., Uriu, K., Tanaka, Y.L., Ito, J., *et al.* (2022a). The SARS-CoV-2 spike S375F mutation characterizes the Omicron BA.1 variant. iScience *25*, 105720.

Kimura, I., Yamasoba, D., Tamura, T., Nao, N., Suzuki, T., Oda, Y., Mitoma, S., Ito, J., Nasser, H., Zahradnik, J., *et al.* (2022b). Virological characteristics of the SARS-CoV-2 Omicron BA.2 subvariants, including BA.4 and BA.5. Cell *185*, 3992-4007 e3916.

Krissinel, E. (2015). Stock-based detection of protein oligomeric states in jsPISA. Nucleic Acids Res *43*, W314-319.

Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nature Methods *11*, 63-65.

Lin, Y.P., Xiong, X., Wharton, S.A., Martin, S.R., Coombs, P.J., Vachieri, S.G., Christodoulou, E., Walker, P.A., Liu, J., Skehel, J.J., *et al.* (2012). Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. Proc Natl Acad Sci U S A *109*, 21474-21479.

Liu, L., Iketani, S., Guo, Y., Chan, J.F., Wang, M., Liu, L., Luo, Y., Chu, H., Huang, Y., Nair, M.S., *et al.* (2022). Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. Nature *602*, 676-681.

Ma, Y., Li, P., Hu, Y., Qiu, T., Wang, L., Lu, H., Lv, K., Xu, M., Zhuang, J., Liu, X., *et al.* (2023). Spike substitution T813S increases Sarbecovirus fusogenicity by enhancing the usage of TMPRSS2. PLoS Pathog *19*, e1011123.

Mallapaty, S. (2022). Where did Omicron come from? Three key theories. Nature *602*, 26-28.

Markov, P.V., Ghafari, M., Beer, M., Lythgoe, K., Simmonds, P., Stilianakis, N.I., and Katzourakis, A. (2023). The evolution of SARS-CoV-2. Nature Reviews Microbiology *21*, 361-379.

Nielsen, S.C.A., Yang, F., Jackson, K.J.L., Hoh, R.A., Roltgen, K., Jean, G.H., Stevens, B.A., Lee, J.Y., Rustagi, A., Rogers, A.J., *et al.* (2020). Human B Cell Clonal Expansion and Convergent Antibody Responses to SARS-CoV-2. Cell Host Microbe *28*, 516-525 e515.

Niu, X., Li, S., Li, P., Pan, W., Wang, Q., Feng, Y., Mo, X., Yan, Q., Ye, X., Luo, J., *et al.* (2020). Longitudinal Analysis of T and B Cell Receptor Repertoire Transcripts Reveal Dynamic Immune Response in COVID-19 Patients. Front Immunol *11*, 582010.

Niu, Z., Zhang, Z., Gao, X., Du, P., Lu, J., Yan, B., Wang, C., Zheng, Y., Huang, H., and Sun, Q. (2021). N501Y mutation imparts cross-species transmission of SARS-CoV-2 to mice by enhancing receptor binding. Signal Transduction and Targeted Therapy *6*.

Pappas, L., Foglierini, M., Piccoli, L., Kallewaard, N.L., Turrini, F., Silacci, C., Fernandez-Rodriguez, B., Agatic, G., Giacchetto-Sasselli, I., Pellicciotta, G., *et al.* (2014). Rapid development of broadly influenza neutralizing antibodies through redundant mutations. Nature *516*, 418-422.

Parameswaran, P., Liu, Y., Roskin, K.M., Jackson, K.K., Dixit, V.P., Lee, J.Y., Artiles, K.L., Zompi, S., Vargas, M.J., Simen, B.B., *et al.* (2013). Convergent antibody signatures in human dengue. Cell Host Microbe *13*, 691-700.

Pastorio, C., Zech, F., Noettger, S., Jung, C., Jacob, T., Sanderson, T., Sparrer, K.M.J., and Kirchhoff, F. (2022). Determinants of Spike infectivity, processing, and neutralization in SARS-CoV-2 Omicron subvariants BA.1 and BA.2. Cell Host Microbe *30*, 1255-1268 e1255.

Patel, A., Kumar, S., Lai, L., Chakravarthy, C., Valanparambil, R., Reddy, E.S., Gottimukkala, K., Bajpai, P., Raju, D.R., Edara, V.V., *et al.* (2023). Molecular basis of SARS-CoV-2 Omicron variant evasion from shared neutralizing antibody response. Structure.

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

Piccoli, L., Park, Y.J., Tortorici, M.A., Czudnochowski, N., Walls, A.C., Beltramello, M., Silacci-Fregni, C., Pinto, D., Rosen, L.E., Bowen, J.E., *et al.* (2020). 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.

Qu, K., Chen, Q., Ciazynska, K.A., Liu, B., Zhang, X., Wang, J., He, Y., Guan, J., He, J., Liu, T., *et al.* (2022). Engineered disulfide reveals structural dynamics of locked SARS-CoV-2 spike. PLoS Pathog *18*, e1010583.

Qu, P., Evans, J.P., Faraone, J.N., Zheng, Y.M., Carlin, C., Anghelina, M., Stevens, P., Fernandez, S., Jones, D., Lozanski, G., *et al.* (2023). 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 *31*, 9-17 e13.

Rapp, M., Guo, Y., Reddem, E.R., Yu, J., Liu, L., Wang, P., Cerutti, G., Katsamba, P., Bimela, J.S., Bahna, F.A., *et al.* (2021). Modular basis for potent SARS-CoV-2 neutralization by a prevalent VH1-2-derived antibody class. Cell Reports *35*, 108950.

Raybould, M.I.J., Kovaltsuk, A., Marks, C., and Deane, C.M. (2021). CoV-AbDab: the coronavirus antibody database. Bioinformatics *37*, 734-735.

Ren, X., Wen, W., Fan, X., Hou, W., Su, B., Cai, P., Li, J., Liu, Y., Tang, F., Zhang, F., *et al.* (2021). COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. Cell *184*, 1895-1913 e1819.

Robbiani, D.F., Gaebler, C., Muecksch, F., Lorenzi, J.C.C., Wang, Z., Cho, A., Agudelo, M., Barnes, C.O., Gazumyan, A., Finkin, S., *et al.* (2020). Convergent antibody responses to SARS-CoV-2 in convalescent individuals. Nature *584*, 437-442.

Saito, A., Tamura, T., Zahradnik, J., Deguchi, S., Tabata, K., Anraku, Y., Kimura, I., Ito, J., Yamasoba, D., Nasser, H., *et al.* (2022). Virological characteristics of the SARS-CoV-2 Omicron BA.2.75 variant. Cell Host Microbe *30*, 1540-1555 e1515.

Starr, T.N., Greaney, A.J., Hannon, W.W., Loes, A.N., Hauser, K., Dillen, J.R., Ferri, E., Farrell, A.G., Dadonaite, B., McCallum, M.*, et al.* (2022). Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution. Science *377*, 420-424.

Sui, J., Hwang, W.C., Perez, S., Wei, G., Aird, D., Chen, L.M., Santelli, E., Stec, B., Cadwell, G., Ali, M.*, et al.* (2009). Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol *16*, 265-273.

Tchesnokova, V., Kulasekara, H., Larson, L., Bowers, V., Rechkina, E., Kisiela, D., Sledneva, Y., Choudhury, D., Maslova, I., Deng, K.*, et al.* (2021). Acquisition of the L452R Mutation in the ACE2-Binding Interface of Spike Protein Triggers Recent Massive Expansion of SARS-CoV-2 Variants. Journal of clinical microbiology *59*, e0092121.

Tian, X., Li, C., Huang, A., Xia, S., Lu, S., Shi, Z., Lu, L., Jiang, S., Yang, Z., Wu, Y.*, et al.* (2020). Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. Emerging Microbes & Infections *9*, 382-385.

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

Toelzer, C., Gupta, K., Yadav, S.K.N., Hodgson, L., Williamson, M.K., Buzas, D., Borucu, U., Powers, K., Stenner, R., Vasileiou, K.*, et al.* (2022). The free fatty acid-binding pocket is a conserved hallmark in pathogenic beta-coronavirus spike proteins from SARS-CoV to Omicron. Sci Adv *8*, eadc9179.

Tuekprakhon, A., Nutalai, R., Dijokaite-Guraliuc, A., Zhou, D., Ginn, H.M., Selvaraj, M., Liu, C., Mentzer, A.J., Supasa, P., Duyvesteyn, H.M.E.*, et al.* (2022). Antibody escape of SARS-CoV-2 Omicron BA.4 and BA.5 from vaccine and BA.1 serum. Cell *185*, 2422-2433.e2413.

Volz, E. (2023). Fitness, growth and transmissibility of SARS-CoV-2 genetic variants. Nature Reviews Genetics.

Wagih, O. (2017). ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics *33*, 3645-3647.

Walls, A.C., Xiong, X., Park, Y.J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A.*, et al.* (2019). Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. Cell *176*, 1026-1039 e1015.

Wang, Q., Guo, Y., Iketani, S., Nair, M.S., Li, Z., Mohri, H., Wang, M., Yu, J., Bowen, A.D., Chang, J.Y.*, et al.* (2022a). Antibody evasion by SARS-CoV-2 Omicron subvariants BA.2.12.1, BA.4 and BA.5. Nature.

Wang, Q., Iketani, S., Li, Z., Liu, L., Guo, Y., Huang, Y., Bowen, A.D., Liu, M., Wang, M., Yu, J.*, et al.* (2023). Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. Cell *186*, 279-286 e278.

Wang, Y., Zhan, W., Liu, J., Wang, Y., Zhang, X., Zhang, M., Han, L., Ma, Y., Lu, L., Wen, Y., *et al.* (2022b). A broadly neutralizing antibody against SARS-CoV-2 Omicron variant infection exhibiting a novel trimer dimer conformation in spike protein binding. Cell Res *32*, 862-865.

Willett, B.J., Grove, J., MacLean, O.A., Wilkie, C., De Lorenzo, G., Furnon, W., Cantoni, D., Scott, S., Logan, N., Ashraf, S., *et al.* (2022). SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway. Nature Microbiology *7*, 1161-1179.

Wu, X., Zhang, Z., Schramm, C.A., Joyce, M.G., Kwon, Y.D., Zhou, T., Sheng, Z., Zhang, B., O'Dell, S., McKee, K., *et al.* (2015). Maturation and Diversity of the VRC01-Antibody Lineage over 15 Years of Chronic HIV-1 Infection. Cell *161*, 470-485.

Xiao, H., Guo, T., Yang, M., Qi, J., Huang, C., Hong, Y., Gu, J., Pang, X., Liu, W.J., Peng, R., *et al.* (2019). Light chain modulates heavy chain conformation to change protection profile of monoclonal antibodies against influenza A viruses. Cell Discovery *5*.

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

Yan, Q., He, P., Huang, X., Luo, K., Zhang, Y., Yi, H., Wang, Q., Li, F., Hou, R., Fan, X., *et al.* (2021). Germline IGHV3-53-encoded RBD-targeting neutralizing antibodies are commonly present in the antibody repertoires of COVID-19 patients. Emerg Microbes Infect *10*, 1097-1111.

Yan, Q., Hou, R., Huang, X., Zhang, Y., He, P., Zhang, Y., Liu, B., Wang, Q., Rao, H., Chen, X., *et al.* (2022). Shared IGHV1-69-encoded neutralizing antibodies contribute to the emergence of L452R substitution in SARS-CoV-2 variants. Emerg Microbes Infect *11*, 2749-2761.

Yang, T.-J., Yu, P.-Y., Chang, Y.-C., and Hsu, S.-T.D. (2021). D614G mutation in the SARS-CoV-2 spike protein enhances viral fitness by desensitizing it to temperature-dependent denaturation. Journal of Biological Chemistry *297*, 101238.

Yu, H., Liu, B., Zhang, Y., Gao, X., Wang, Q., Xiang, H., Peng, X., Xie, C., Wang, Y., Hu, P., *et al.* (2023). Somatically hypermutated antibodies isolated from SARS-CoV-2 Delta infected patients cross-neutralize heterologous variants. Nat Commun *14*, 1058.

Yuan, M., Huang, D., Lee, C.-C.D., Wu, N.C., Jackson, A.M., Zhu, X., Liu, H., Peng, L., Gils, M.J.v., Sanders, R.W., *et al.* (2021). Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. Science *373*, 818-823.

Yuan, M., Liu, H., Wu, N.C., Lee, C.C.D., Zhu, X., Zhao, F., Huang, D., Yu, W., Hua, Y., Tien, H., *et al.* (2020a). Structural basis of a shared antibody response to SARS-CoV-2. Science *369*, 1119-1123.

Yuan, M., Wu, N.C., Zhu, X., Lee, C.-C.D.C.D., So, R.T.Y.Y., Lv, H., Mok, C.K.P.P., and Wilson, I.A. (2020b). A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. Science *368*, 630-633.

1147 Yurkovetskiy, L., Wang, X., Pascal, K.E., Tomkins-Tinch, C., Nyalile, T.P., Wang, Y., Baum, A.,
1148 Diehl, W.E., Dauphin, A., Carbone, C.*, et al.* (2020). Structural and Functional Analysis of the
1149 D614G SARS-CoV-2 Spike Protein Variant. Cell *183*, 739-751.e738.

1150 Zhang, F., Gan, R., Zhen, Z., Hu, X., Li, X., Zhou, F., Liu, Y., Chen, C., Xie, S., Zhang, B.*, et al.*
1151 (2020). Adaptive immune responses to SARS-CoV-2 infection in severe versus mild
1152 individuals. Signal Transduct Target Ther *5*, 156.

1153 Zhang, J., Cai, Y., Xiao, T., Lu, J., Peng, H., Sterling, S.M., Walsh, R.M., Rits-Volloch, S., Zhu, H.,
1154 Woosley, A.N.*, et al.* (2021a). Structural impact on SARS-CoV-2 spike protein by D614G
1155 substitution. Science *372*, 525-530.

1156 Zhang, Q., Ju, B., Ge, J., Chan, J.F., Cheng, L., Wang, R., Huang, W., Fang, M., Chen, P., Zhou,
1157 B.*, et al.* (2021b). Potent and protective IGHV3-53/3-66 public antibodies and their shared
1158 escape mutant on the spike of SARS-CoV-2. Nat Commun *12*, 4210.

1159 Zhang, X., Li, Z., Zhang, Y., Liu, Y., Wang, J., Liu, B., Chen, Q., Wang, Q., Fu, L., Wang, P.*, et al.*
1160 (2023). Disulfide stabilization reveals conserved dynamic features between SARS-CoV-1 and
1161 SARS-CoV-2 spikes. Life Sci Alliance *6*.

1162 Zhang, Y., Yan, Q., Luo, K., He, P., Hou, R., Zhao, X., Wang, Q., Yi, H., Liang, H., Deng, Y.*, et al.*
1163 (2022). Analysis of B Cell Receptor Repertoires Reveals Key Signatures of the Systemic B Cell
1164 Response after SARS-CoV-2 Infection. Journal of Virology *96*, 1-23.
1165
1166

a **Features of isolated IGLV6-57 (VL6-57) light chain utilizing mAbs**

		Gene analysis								Binding of mAb to RBD	Pseudovirus neutralization
Donor	mAb	IGHV	IGHD	IGHJ	HCDR3	SHM%	IGLV	IGLJ	LCDR3	SHM%	K_D (nM)	IC_{50} (nM)
PtK	R1-26	IGHV3-7	IGHD6-6	IGHJ4	ARGQLGPWVGVDY	0	IGLV6-57	IGLJ3	QSYDSSNWV	0	3.84	2.71
PtK	R1-30	IGHV4-59	IGHD4-23	IGHJ4	ARQGWLRGNFDY	0	IGLV6-57	IGLJ2	QSYDSSIHVV	0	35.38	4.06
PtZ	R2-3	IGHV3-7	IGHD5-18	IGHJ3	ASQLWLRGAFDI	2.78	IGLV6-57	IGLJ3	QSYDSSNPWV	1.72	13.28	3.46
PtZ	R2-6	IGHV4-31	IGHD5-24	IGHJ3	ARKGWLRGAFDI	1.37	IGLV6-57	IGLJ2	QSYDSGVV	0.69	22.8	13.91
PtZ	R2-7	IGHV4-31	IGHD5-24	IGHJ3	ARKGWLRGAFDI	1.37	IGLV6-57	IGLJ2	QSYDSSNHLVV	0	62.2	>333

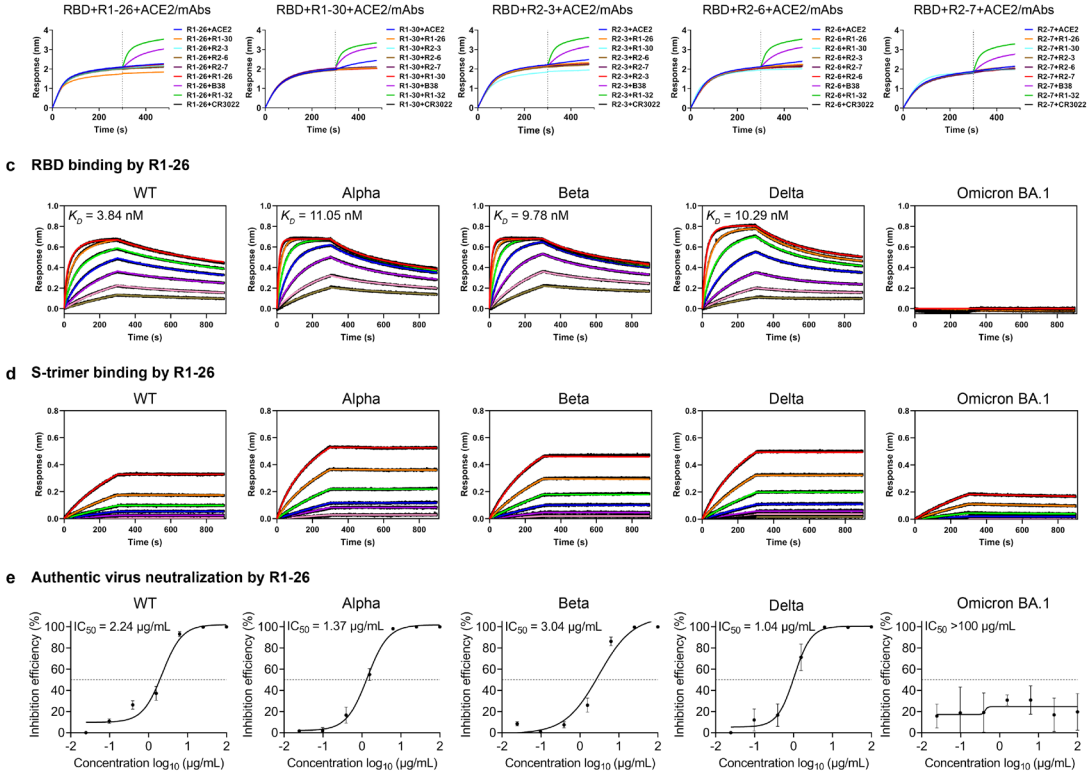

Fig. 1 | Characteristics of the 5 VL6-57 light chain utilizing antibodies. a, The genetic and functional properties of the 5 isolated VL6-57 light chain utilizing antibodies (He et al., 2022). Neutralization titres (IC_{50}) of antibodies were determined against a pseudovirus displaying WT SARS-CoV-2 S-protein. Binding affinities (K_D) of antibodies to SARS-CoV-2 WT RBD were determined by biolayer interferometry (BLI) assays. Germline usage and somatic hypermutation (SHM) analysis of antibodies were performed using IMGT/V-QUEST. **b**, Pairwise binding competition to SARS-CoV-2 RBD was assessed by BLI. A biosensor immobilized with WT SARS-CoV-2 RBD was first saturated with one of the VL6-57 utilizing mAbs until the dashed line before submerging into a solution of another mAb or ACE2 to assess competition. An ACE2 competing antibody B38 (purple) and two non-ACE2 competing antibodies recognizing different RBD epitopes, CR3022 (black) (Yuan et al., 2020b) and R1-32 (green) (He et al., 2022) were used as controls. **c-d**, BLI binding curves of R1-26 to 2-fold serially diluted RBD (**c**) or S-protein

(**d**) solutions of WT SARS-CoV-2 and VOCs. The black lines represent the experimentally recorded sensorgram traces, the colored lines represent corresponding fits. Detailed binding kinetics parameters are summarized in **Supplementary Table 1**. **e**, Neutralization activities of R1-26 towards SARS-CoV-2 WT and VOC authentic viruses in cell culture (data are presented as mean values ± SD).

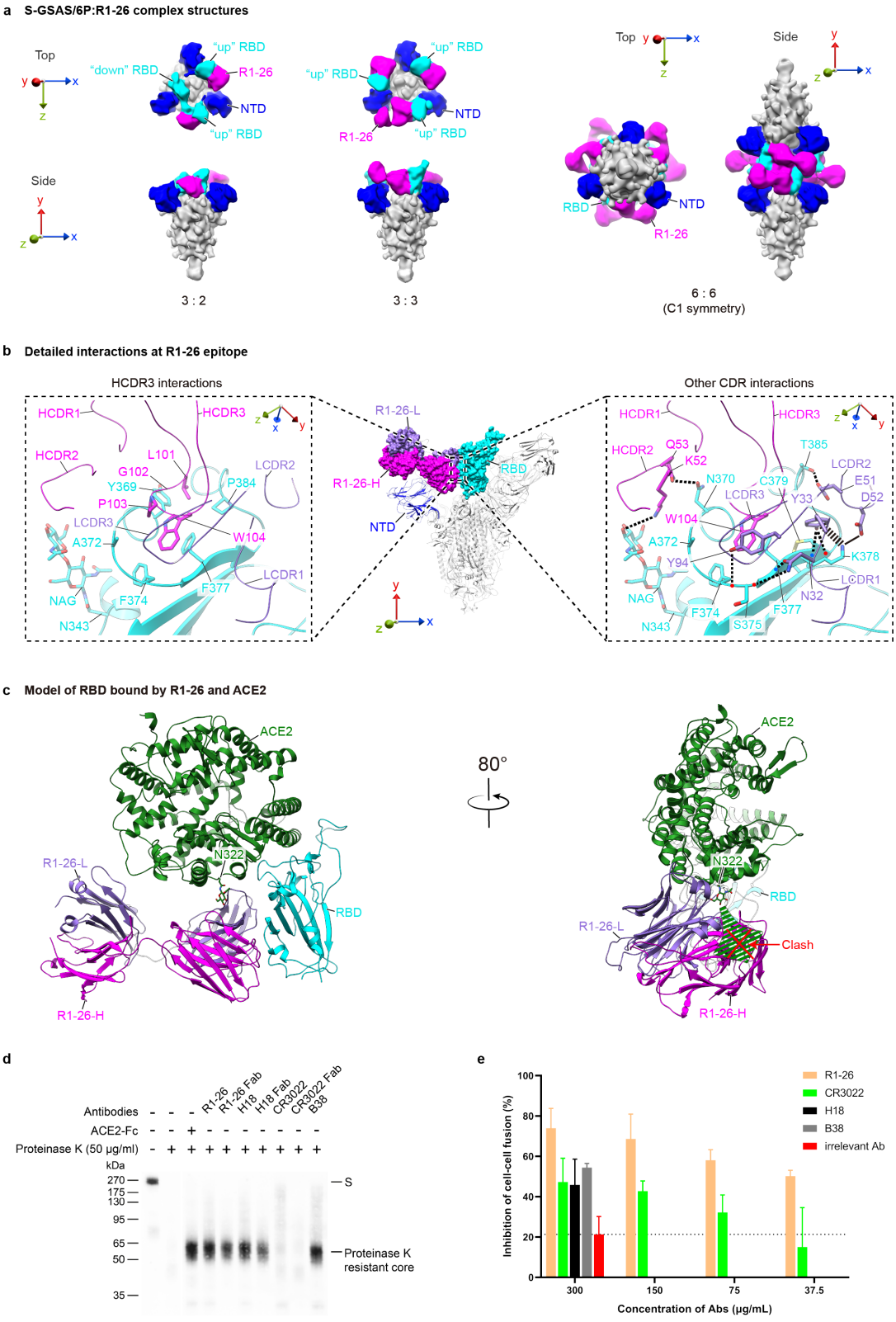

Fig. 2 | Structural analysis of R1-26 epitope binding and biochemical activities of R1-26. a, Structures (low-pass filtered to 12 Å) of S-GSAS/6P S-trimers in complex with R1-26 Fabs in different stoichiometries. R1-26 Fab, NTD and RBD are highlighted in magenta, blue and cyan, respectively; the rest

of the S-trimer is colored gray. **b**, Epitope of R1-26 on spike RBD. R1-26-H and R1-26-L chains are colored in magenta and purple. Detailed interactions between R1-26 and RBD are shown in the dashed boxes. CDR loops are indicated, selected interacting residues in antibody-RBD interface are shown. Thick and thin dashed lines indicate cation-π and hydrogen bond interactions; The solid line represents a salt bridge. **c**, A hypothetical model showing simultaneous RBD binding by R1-26 and ACE2 (also see **Fig. S4** for comparison with simultaneous RBD binding by CR3022 and ACE2). The model is based on RBD complex structures of R1-26 and ACE2 (PDB: 7YDI). The glycan attached to ACE2 residue N322 is labelled and the green dashed arrow indicates a possible direction of glycan chain extension. **d**, Ligand-induced conformational change assays to probe the induction of post-fusion structures, presence of proteinase K resistance core is indicative of post-fusion structure (He et al., 2022; Matsuyama and Taguchi, 2009; Walls et al., 2019). **e**, R1-26 inhibits spike-ACE2 interaction mediated cell-cell membrane fusion.

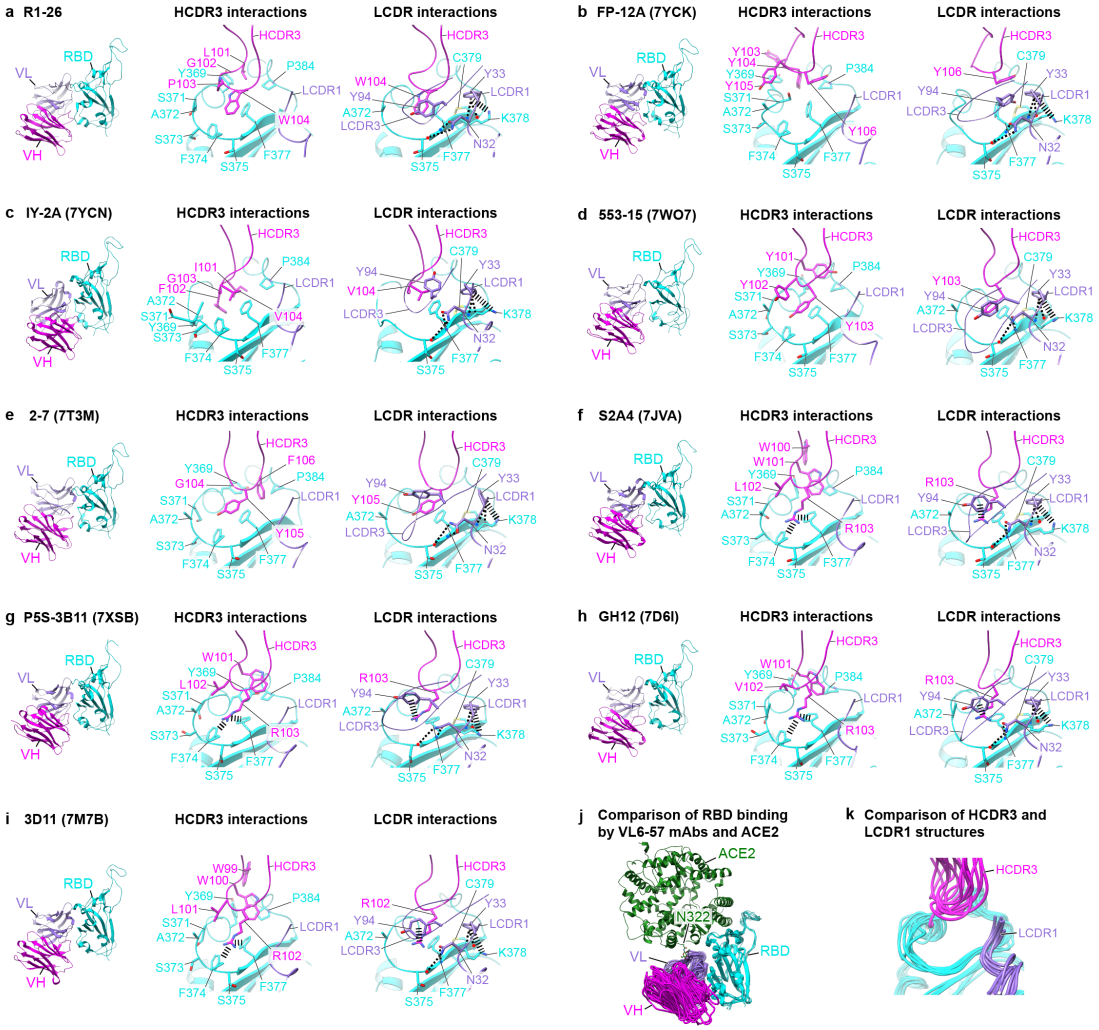

Fig. 3 | Epitopes and binding modes of VL6-57 antibodies. a-i, Structures of R1-26 (**a**), FP-12A (**b**), IY-2A (**c**), 553-15 (**d**), 2-7 (**e**), S2A4 (**f**), P5S-3B11 (**g**), GH12 (**h**), and 3D11 (**i**) bound to their epitopes. Fab-VH, Fab-VL and RBD are colored in magenta, purple and cyan, respectively. Selected interactions engaged by HCDR3, LCDR1 and LCDR3 are shown. All shown antibodies bind almost the same surface area on RBD, hydrophobic, particularly aromatic amino acid residues are found in all the HCDR3 loops of the shown VL6-57 antibodies for epitope recognition. All LCDR1 loops have the same interactions with RBD. **j**, Comparison of RBD binding by the shown VL6-57 mAbs and ACE2. **k**, Superposition of HCDR3 and LCDR1 loops of the shown VL6-57 mAbs. Variability is observed for the HCDR3 loops while the LCDR1 loops show high similarity.

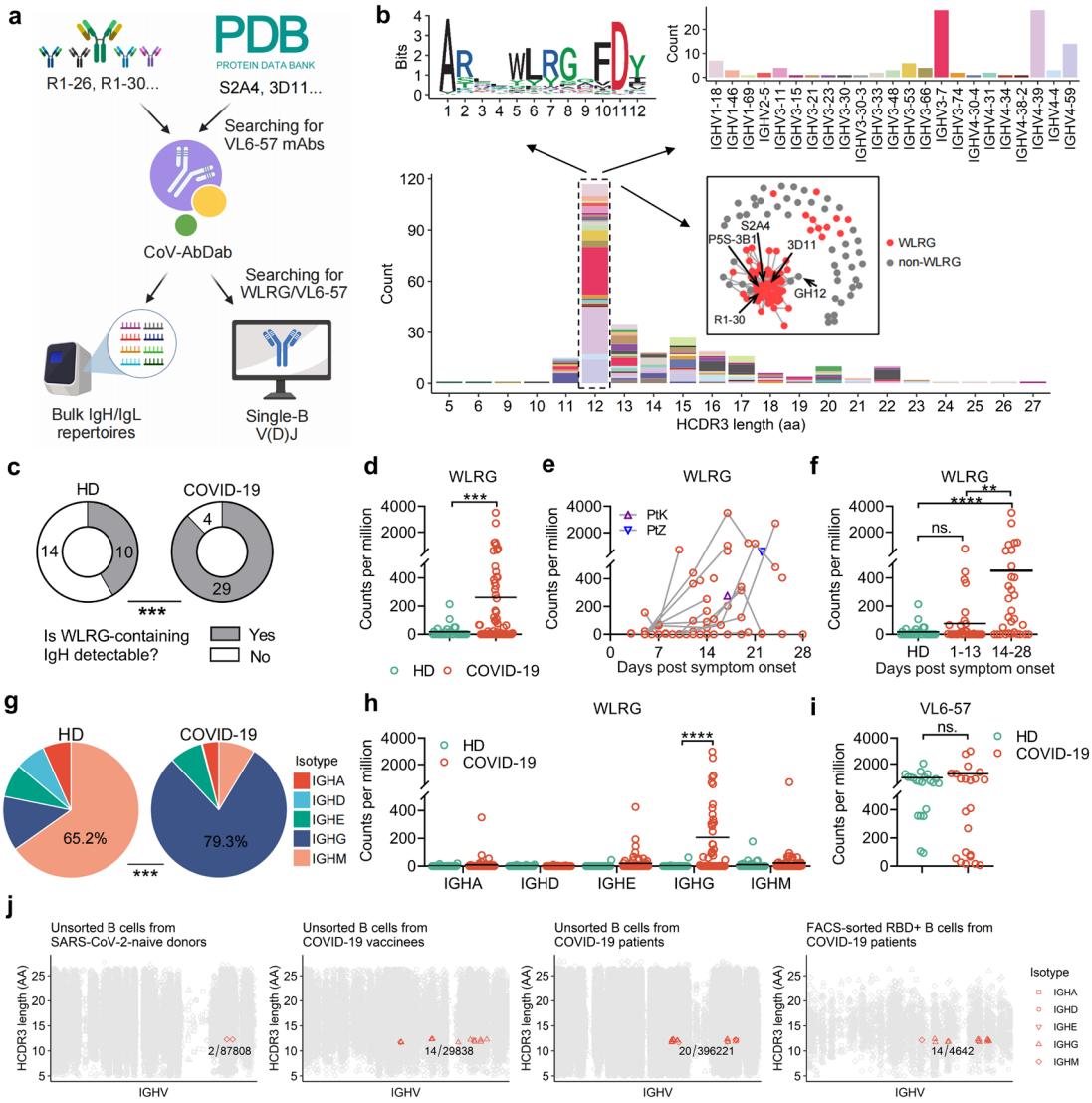

Fig. 4 | Analysis of VL6-57 light chain sequences and associated heavy chain sequences in antibody repertoires. a, A schematic diagram showing the search workflow for VL6-57 light chain and the associated "WLRG" motif containing heavy chain sequences from CoV-AbDab or published antibody repertoire datasets. Created with BioRender.com. **b**, Heavy chain sequence analysis of the 290 VL6-57 utilizing mAbs identified from the CoV-AbDab. A histogram showing the HCDR3 length distribution among the 290 mAbs (bottom). Sequence logo plot showing consensus HCDR3 amino acids of the VL6-57 utilizing mAbs with 12-AA long HCDR3s (top left). A histogram showing the heavy chain germline gene usage among the VL6-57 utilizing mAbs with 12-AA long HCDR3s (top right). Lineage structure of the VL6-57 utilizing mAbs with 12-AA long HCDR3s based on similarity of HCDR3 sequences (boxed). When similarity of any two mAbs' HCDR3s is ≥ 80%, they

are linked by a line. mAbs containing the HCDR3 "WLRG" motif are colored red. **c**, Pie chart showing the occurrence of "WLRG" motif containing IgH sequences in a cohort of 33 COVID-19 convalescents and 24 healthy donors. **d**, Read-count comparison of "WLRG" motif containing IgH sequences in the IgH repertoires of COVID-19 convalescents and healthy donors. **e**, Read-count dynamics of "WLRG" motif containing IgH sequences in the IgH repertoires of COVID-19 convalescents, each line connects data of different timepoints from the same patient. **f**, Read-count comparison of "WLRG" motif containing IgH sequences in the IgH repertoires of COVID-19 convalescents and healthy donors across different timepoints. **g**, Pie chart showing the isotype distribution of "WLRG" motif containing IgH sequences in COVID-19 patients and healthy donors. **h**, Read-count comparison of "WLRG" motif containing IgH sequences between COVID-19 convalescents and healthy donors across different isotypes. **i**, Read-count comparison of "QSYDSS" containing VL6-57 light chain sequences in the IgL repertoires of COVID-19 convalescents and healthy donors. Antibody sequence analyses in panels **c-i** used published bulk sequencing datasets (Niu et al., 2020; Yan et al., 2021; Zhang et al., 2022). **j**, The occurrence of VL6-57 mAbs with paired heavy chains containing the HCDR3 "WLRG" motif in the single-B V(D)J repertoires of COVID-19 patients, vaccinees and healthy donors (Barmada et al., 2023; Dugan et al., 2021; Ferreira-Gomes et al., 2021; Ren et al., 2021; Yu et al., 2023; Zhang et al., 2020). The two-sided chi-square test was performed in panels (**c**) and (**g**). Student's t test was performed in panels (**d**), (**f**), (**h**), and (**i**). (**** $p < 0.001$, *** $p < 0.001$, ** $p < 0.01$, ns. $p \geq 0.05$).

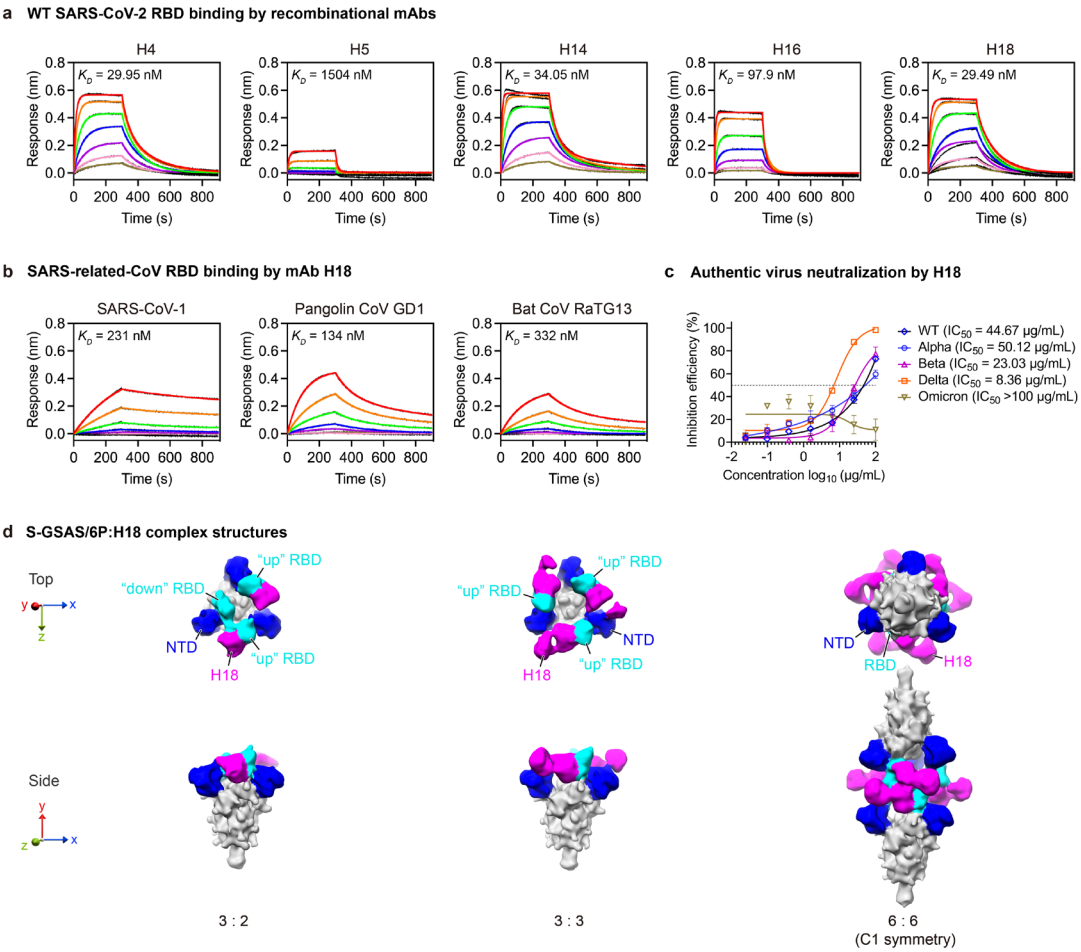

Fig. 5 | Characteristics of recombinational mAbs derived from healthy donors. a, WT SARS-CoV-2 RBD binding by five recombinational mAbs were assessed by BLI. Black and colored lines represent experimentally recorded sensorgram traces and corresponding fits. 2-fold serially diluted RBD solutions (200 nM to 3.125 nM) were used in the BLI assays. **b**, Binding of mAb H18 to SARS-CoV-1, Pangolin CoV GD1 and Bat CoV RaTG13 RBDs were measured by BLI. Kinetic parameters for panels (**a**) and (**b**) are summarized in **Supplementary Table 3**. **c**, Neutralization activities (mean ± SD) of mAb H18 towards SARS-CoV-2 authentic viruses in cell culture. **d**, Structures of S-GSAS/6P S-trimers in complex with H18 Fabs in different stoichiometries and conformations. **e**, Structures of S-GSAS/6P S-trimers or S1s in complex with H18 Fabs and R1-32 Fabs. Structures in **d-e** are low-pass filtered to 12 Å to reveal flexible regions (**also see Fig. S2**). H18 Fab, NTD, and RBD are highlighted in magenta, blue, cyan, respectively; other structures are colored gray. **f**, Detailed H18 epitope structure. H18-H and H18-L chains are colored in magenta and purple; CDR loops are indicated, selected interacting residues between RBD and H18 are shown and indicated; thick and thin dashed lines indicate cation-π interactions and hydrogen bonds.

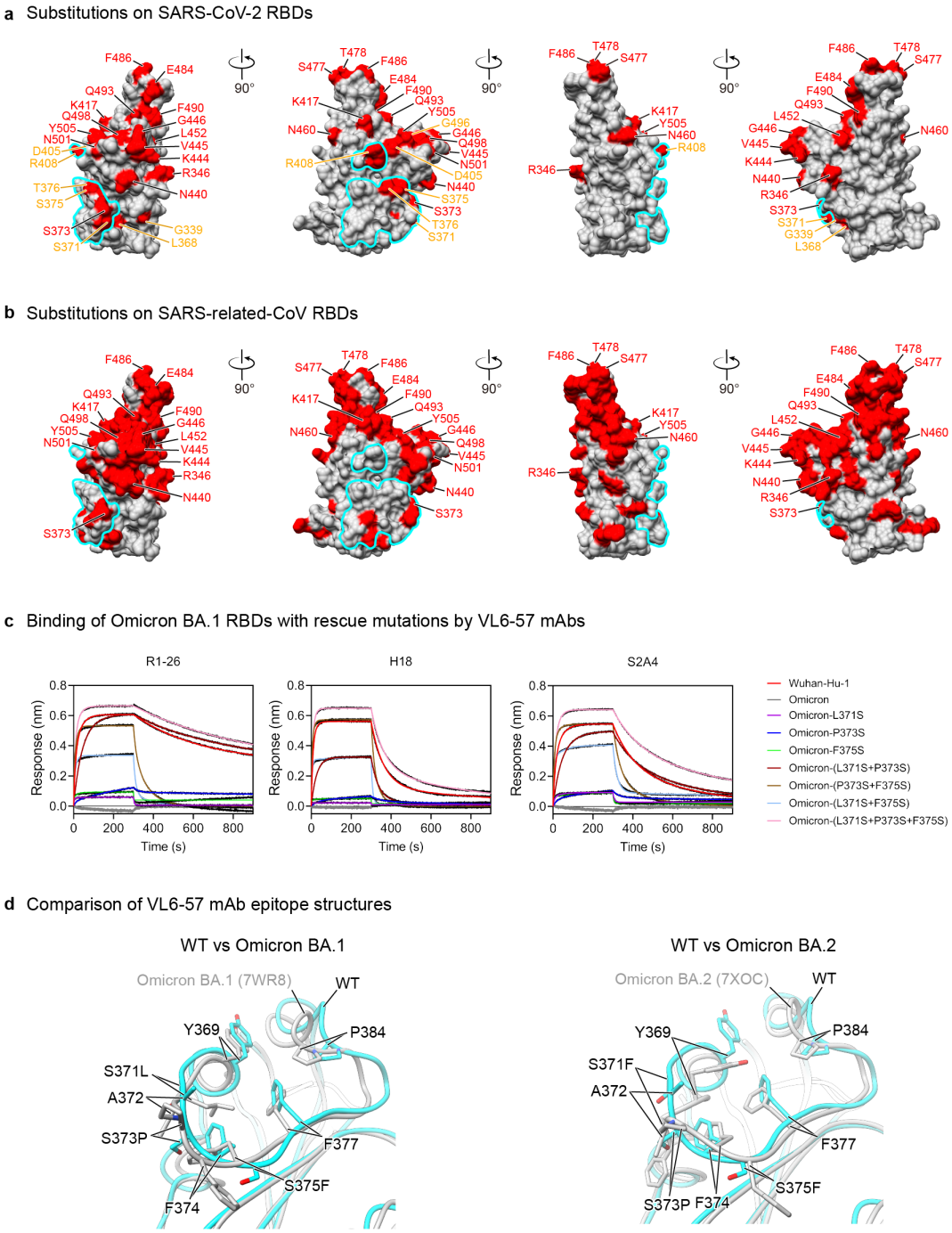

Fig. 6 | SARS-CoV-2 Omicron variants evade VL6-57 antibodies by mutations at residues S371-S373-S375. a, Substituted residues among SARS-CoV-2 variants are highlighted in red on the surface of SARS-CoV-2 RBD. Among the 26 substitutions, 18 are shared among SARS-related-CoVs and are labeled in red, the other 8 substitutions are labeled in orange. **b**, Substituted residues among RBDs of SARS-related-CoVs are highlighted in red. Among the 72 substitutions, 18 are shared with SARS-CoV-2 variants and are labeled in red. Epitope of R1-26 on RBD is encircled by cyan outlines in panels (**a**) and

(**b**). **c**, Binding of R1-26, H18 and S2A4 to a series of Omicron BA.1 RBD mutants carrying 1-3 rescue mutations at S371, S373, and S375. BLI binding kinetic parameters are summarized in **Supplementary Table 4**. **d**, Comparison of the VL6-57 mAb epitope structures on WT and Omicron BA.1 (left) or BA.2 (right) RBDs. Superposition shows differences in amino-acid sidechain orientations and main chain backbone structures.

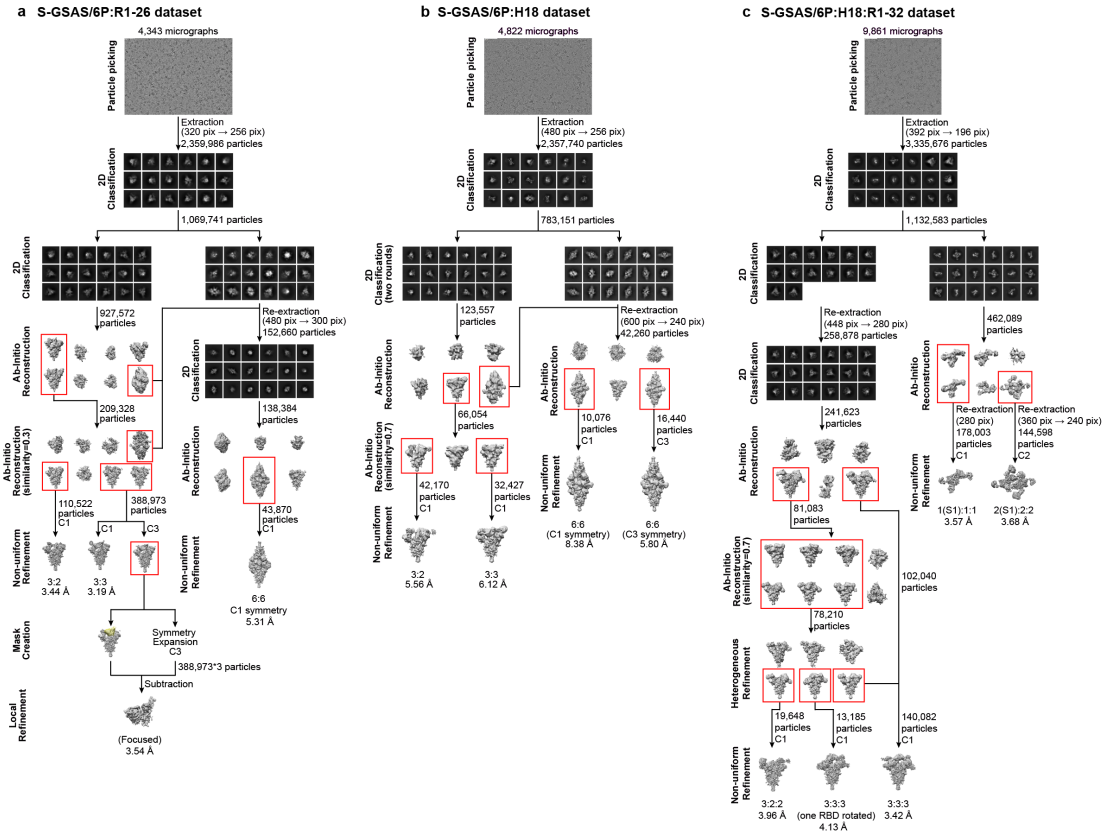

Fig. S1 | Cryo-EM data processing pipelines. a, Data processing pipeline for the S-GSAS/6P:R1-26 complex dataset, the mask for subtraction is colored in yellow. **b**, Data processing pipeline for the S-GSAS/6P:H18 complex dataset. **c**, Data processing pipeline for the S-GSAS/6P:H18:R1-32 complex dataset. Red boxes indicate selected classes.

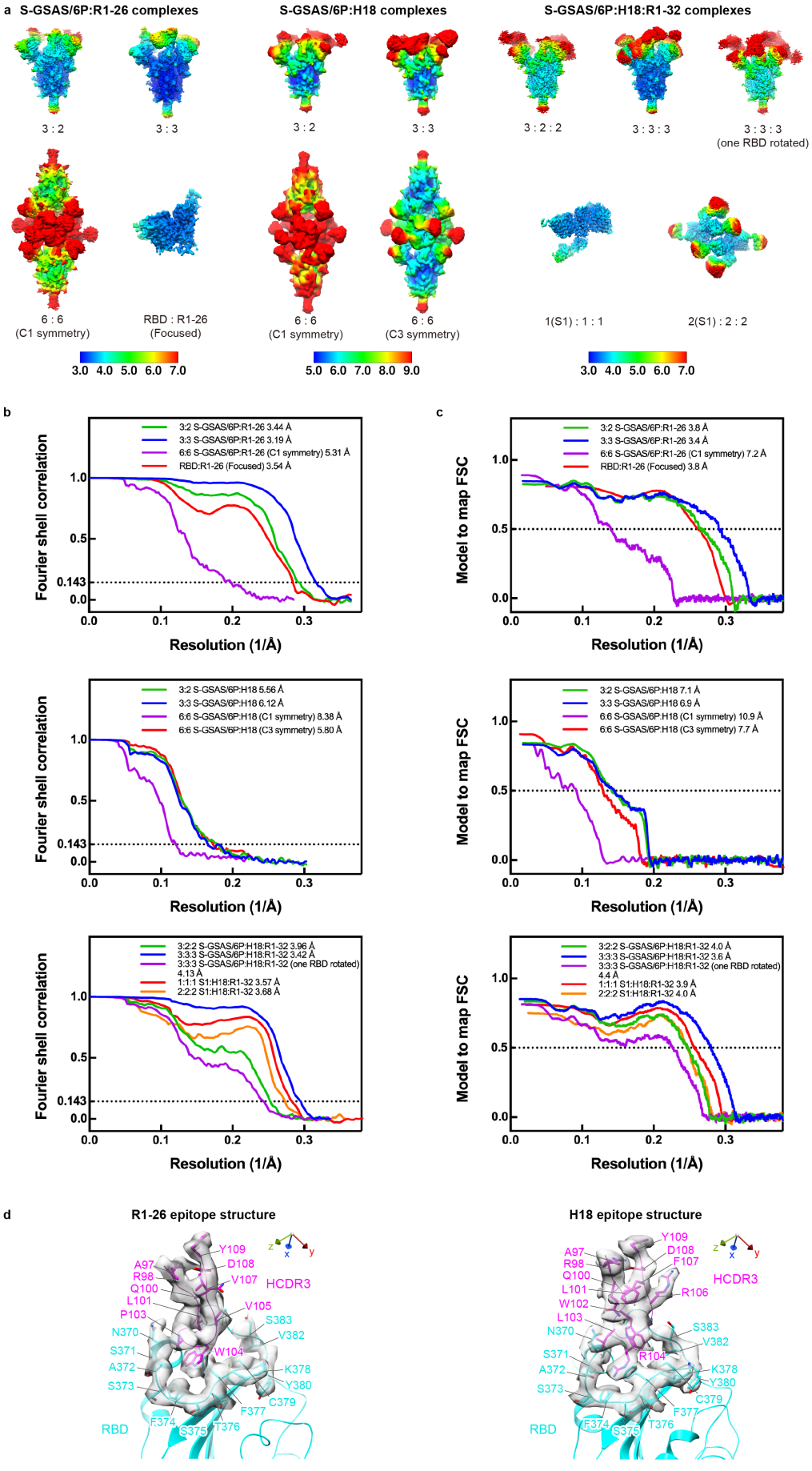

Fig. S2 | Resolution assessment of cryo-EM structures. a, Local resolution assessment for the complex structures determined from S-GSAS/6P:R1-26, S-GSAS/6P:H18 and S-GSAS/6P:H18:R1-32 datasets. **b**, Global resolution assessment by Fourier shell correlation (FSC) at the 0.143 criterion. **c**, Correlations of model vs map by FSC at the 0.5 criterion. **d**, Representative cryo-EM densities around antibody-RBD binding interfaces. HCDR3 loops of R1-26 or H18 are colored magenta and RBD is colored cyan.

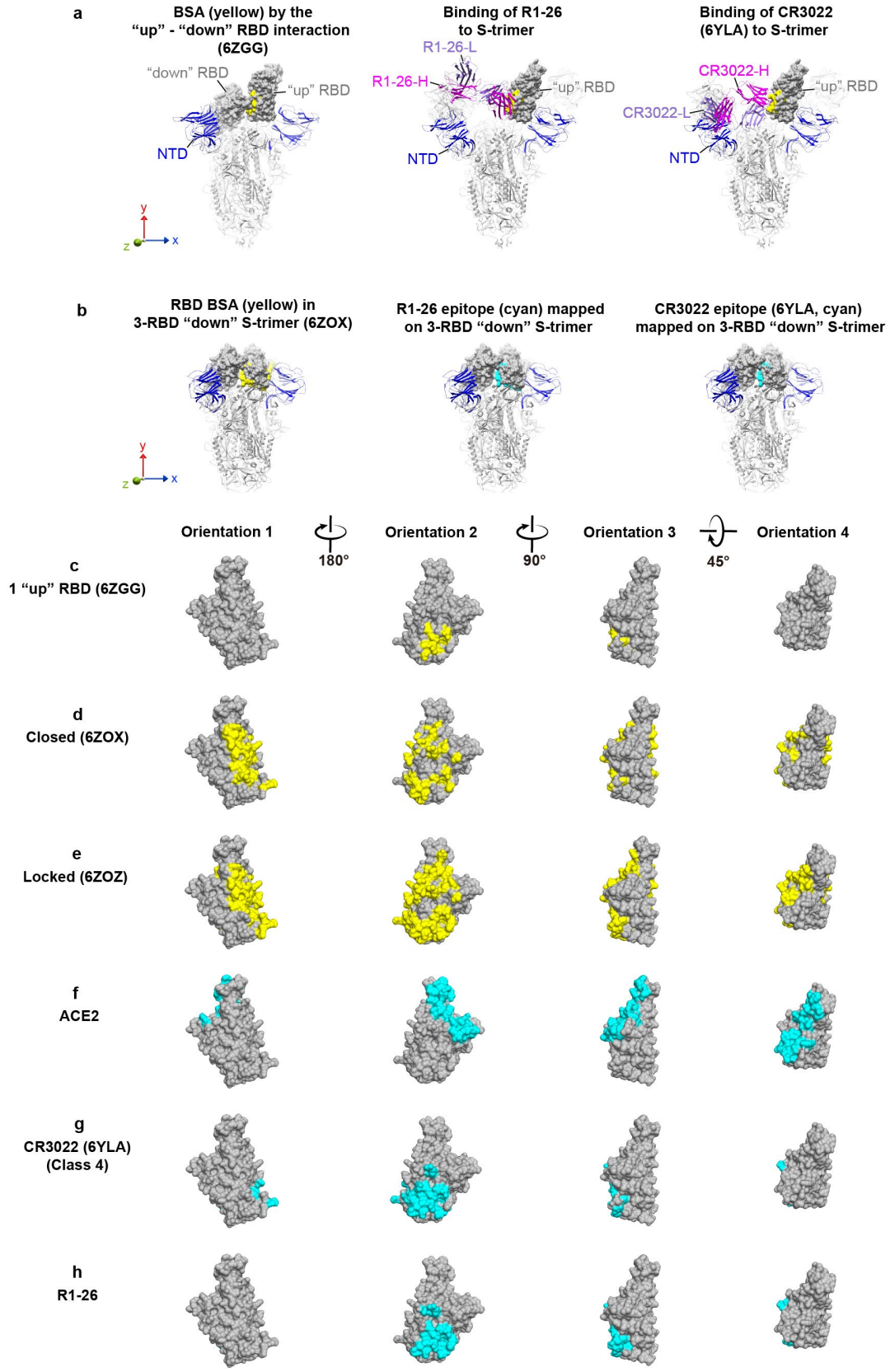

Fig. S3 | Epitope of R1-26 in the context of different S-trimer conformations and compared to other buried surface areas (BSAs) on

RBD. RBD molecular surfaces are colored in gray, BSAs in different S-trimer conformations are highlighted in yellow, BSAs by ACE2 or antibodies are highlighted in cyan. **a**, In each of the shown 1 RBD "up" S-trimers (Wrobel et al., 2020), buried surface area on the "up" RBD buried by the adjacent "down" RBD is colored yellow (also shown in **c**). Binding of R1-26 (middle) or CR3022 (right) (Yuan et al., 2020a) is shown and compared with the buried surface area by the "down" RBD. **b**, In each of the shown 3 RBD "down" closed S-trimer structures (Xiong et al., 2020), buried RBD surface area (left, colored yellow, also shown in **d**) or epitope of R1-26 (middle, cyan, also shown in **h**) or CR3022 (right, cyan, also shown in **g**) is shown to illustrate R1-26 and CR3022 epitopes being cryptic. **e**, RBD buried surface area in the 3 RBD "down" locked S-trimer structure (Xiong et al., 2020). **f-h**, Binding surface of ACE2 (**f**) (Lan et al., 2020) and epitope of CR3022 (**g**) or R1-26 (**h**).

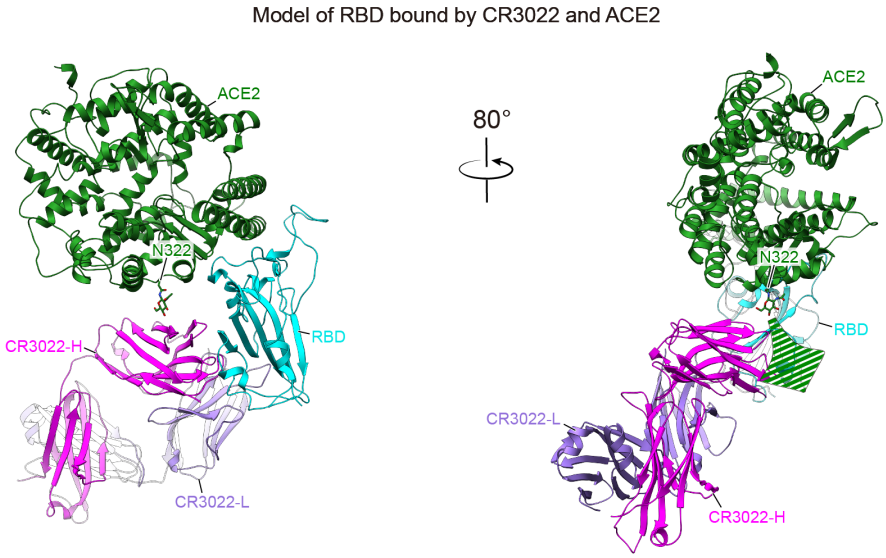

Fig. S4 | A hypothetical model showing simultaneous RBD binding by CR3022 and ACE2. The model is based on RBD complex structures of CR3022 (PDB: 6W41) and ACE2 (PDB: 7YDI). The glycan attached to ACE2 residue N322 is labelled and the green dashed arrow indicates a possible direction of glycan chain extension. Also see **Fig. 2c** for comparison with simultaneous RBD binding by R1-26 and ACE2.

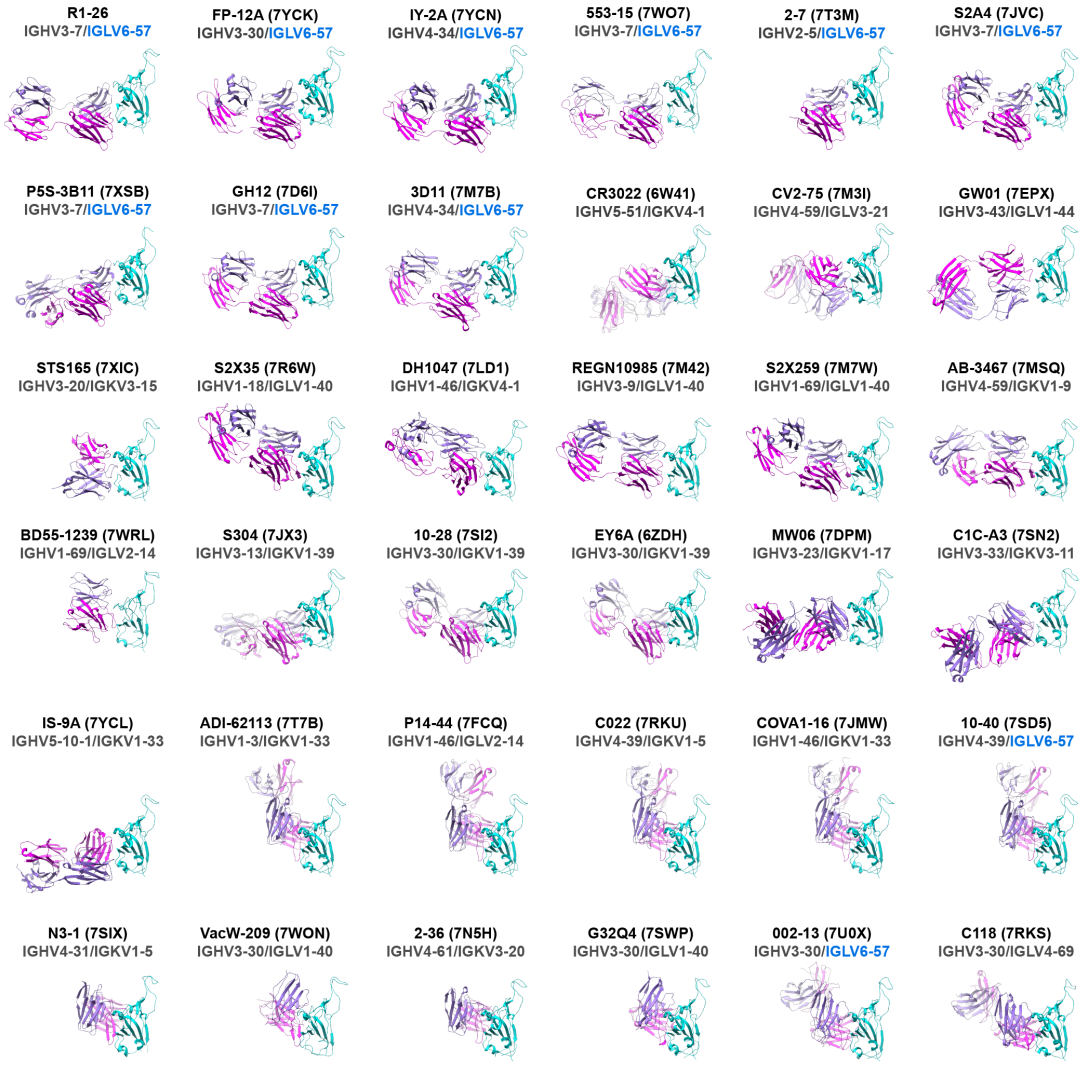

Fig. S5 | Structure gallery of class 4 antibodies in complex with RBD. Antibody H chains and L chains are colored in magenta and purple. RBDs are colored in cyan. Antibody heavy and light chain V genes encoding each antibody are shown under each antibody name with corresponding PDB code in backet. IGLV6-57 encoded antibody structures are highlighted by their light chain V genes shown in blue. Structures are arranged by similarity of their binding modes.

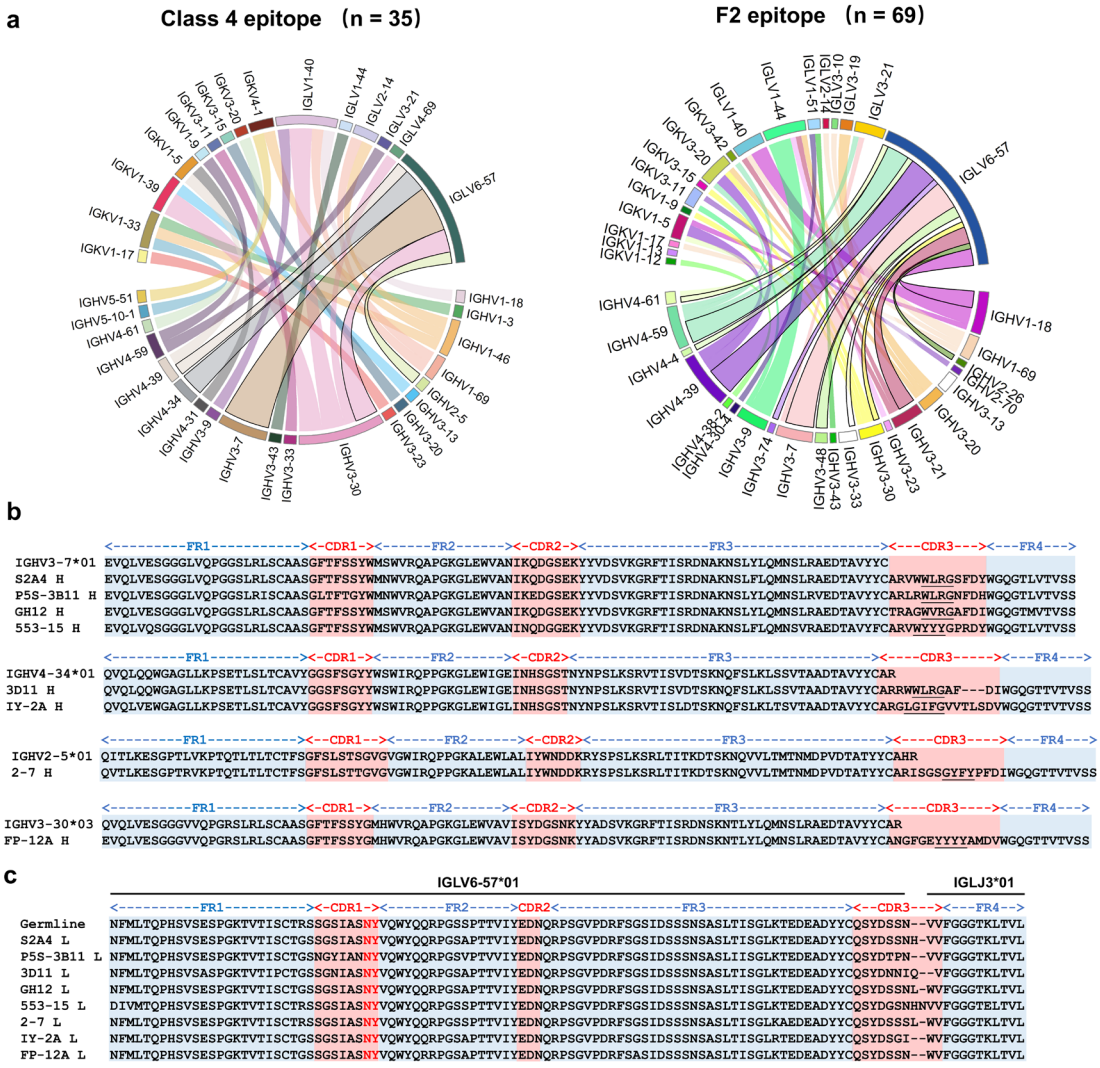

Fig. S6 | Sequence analysis of class 4 mAbs and VL6-57 utilizing mAbs. a, Germline heavy and light chain gene usage and pairing among the 35 class 4 (left) and the 69 class F2 (right) antibodies. In the chord diagrams, each of the germline heavy and light chain pair is linked by an arc, the size of arc is proportional to the number of antibodies identified with such heavy-light chain pairing. **b**, Heavy chain sequence alignment of the 8 structurally characterized VL6-57 utilizing mAbs with their germline gene sequences. Hydrophobic, particularly aromatic amino acid residues within HCDR3 regions are underlined. **c**, Light chain sequence alignment of the 8 structurally characterized VL6-57 utilizing mAbs with their germline gene sequences. The RBD-contact motif "NY" encoded by VL6-57 light chain LCDR1 germline sequence is shown in red letters.

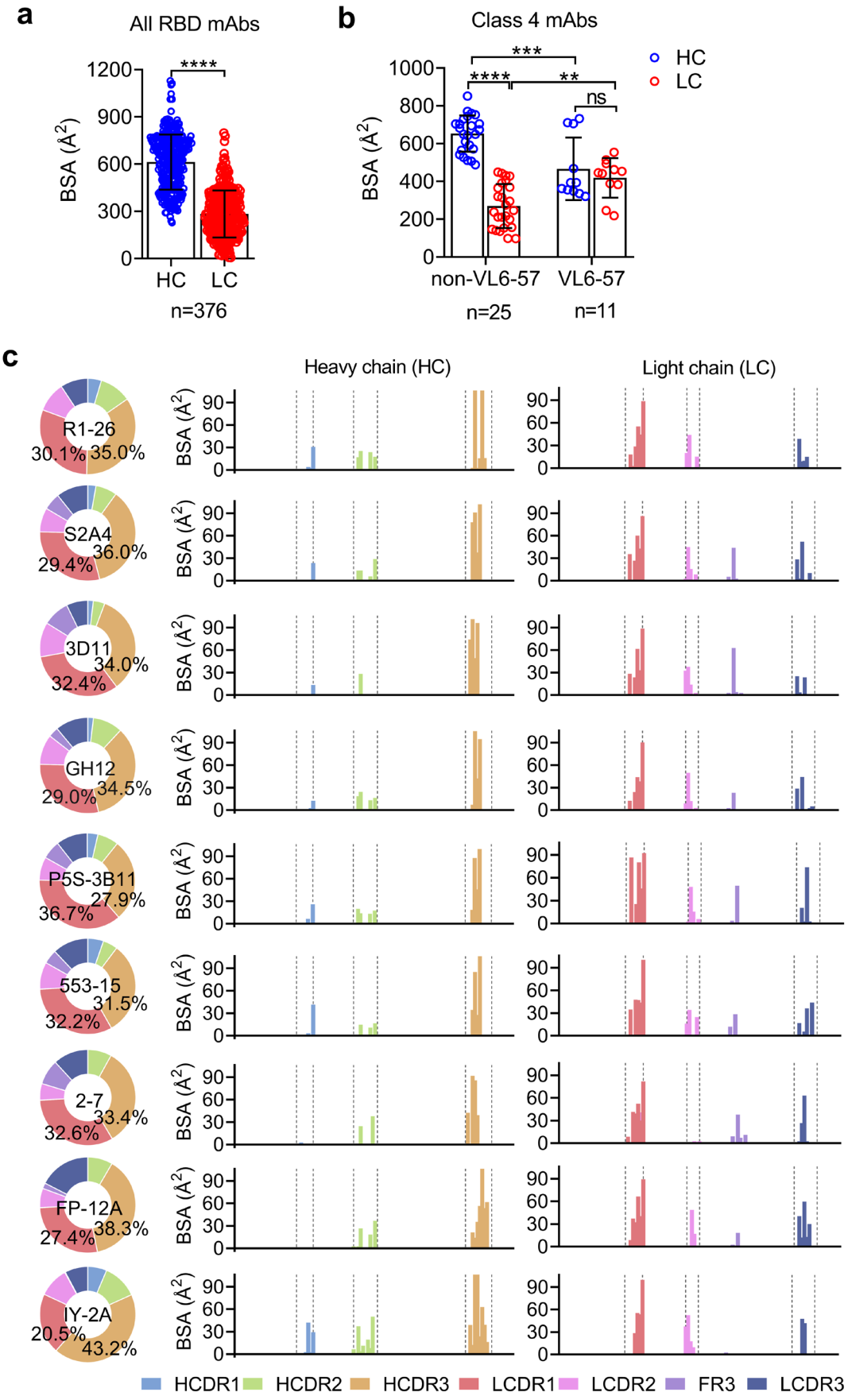

Fig. S7 | Buried surface area (BSA) analysis of the structurally characterized spike-specific mAbs. a, Comparison of BSAs by heavy (HC) and light chains (LC) of the 376 structurally characterized spike-specific mAbs available in the PDB. **b**, Differences in BSAs by heavy (HC) and light chain (LC) are compared between VL6-57 (R1-26 is included in this analysis) and non-VL6-57 class 4 mAbs. **c**, Per-site BSA analysis on heavy chains (HC) and light chains (LC) of R1-26 and the other 8 VL6-57 mAbs. Student's t test was performed in panels (**a**) and (**b**). (**** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, ns. $p \geq 0.05$).

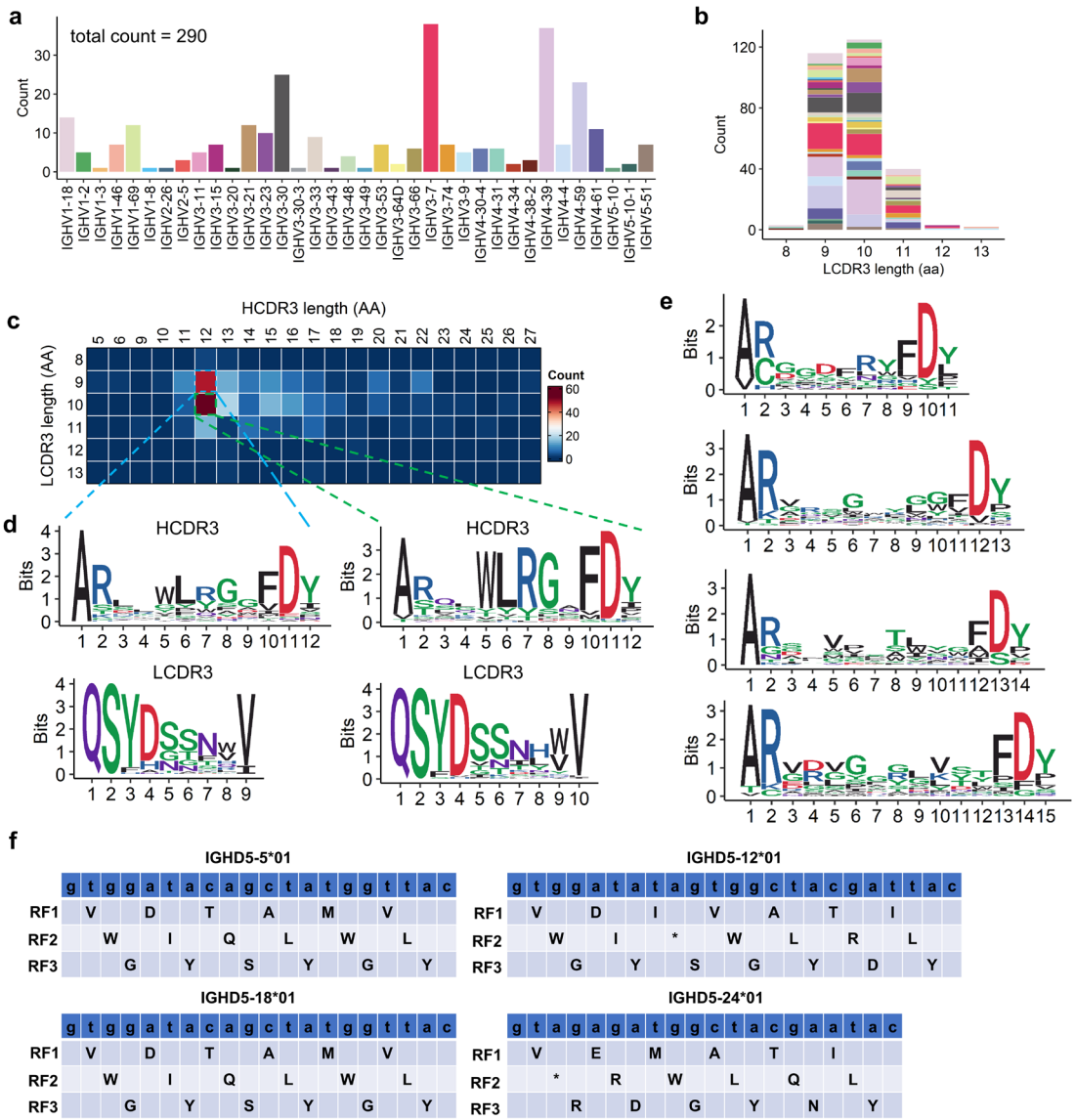

Fig. S8 | Sequence analysis of VL6-57 utilizing mAbs from the CoV-AbDab. **a**, A histogram showing the heavy chain germline gene usage of the 290 VL6-57 utilizing mAbs curated from the CoV-AbDab. **b**, A histogram showing the LCDR3 length distribution of the 290 VL6-57 utilizing mAbs. **c**, A heatmap showing the HCDR3 and LCDR3 length pairing and distribution among the 290 VL6-57 utilizing mAbs. **d**, Sequence logo plots showing consensus amino acids within 12-aa length HCDR3 loops and their paired 9-aa (left) or 10-aa (right) length LCDR3 loops of VL6-57 mAbs. **e**, Sequence logo plots showing consensus amino acids of HCDR3 loops of 11-, 13-, 14-, or 15-aa length among the 290 VL6-57 utilizing mAbs. **f**, Three reading frames (RF1-3) of IGHD5 germline genes are shown. The "WLR" sequence is encoded by RF2 of IGHD5-12 (top right).

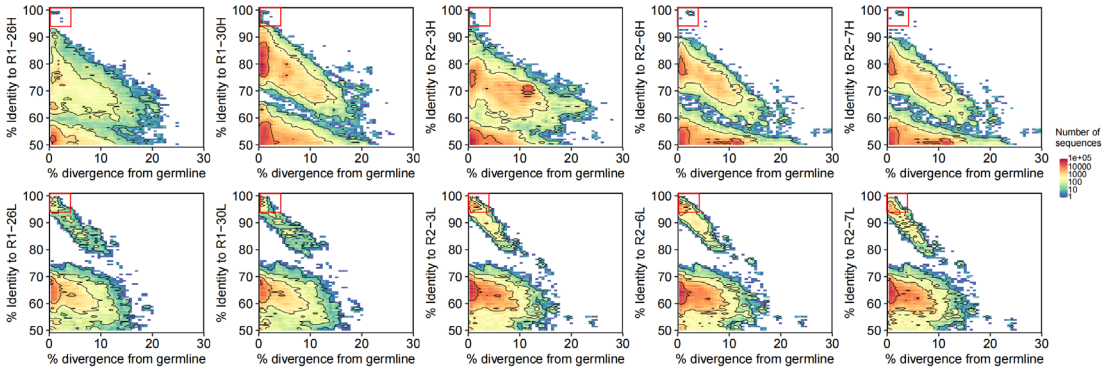

Fig. S9 | Identity-divergence plots showing the presence of R1-26, R1-30, R2-3, R2-6, or R2-7-like IgH or IgL lineages in the repertoires of their donors. All sequences in the repertoires are plotted as a function of sequence divergence from their germline genes (x axis) and sequence identity to heavy chain or light chain of R1-26, R1-30, R2-3, R2-6, or R2-7 (y axis). Color gradient indicates sequence density. Red boxes indicate peaks containing IgH or IgL sequences highly similar to those of R1-26, R1-30, R2-3, R2-6, or R2-7.

	Heavy chain						Light chain					
mAb Id	IGHV	IGHD	IGHJ	HCDR3	HCDR3 length	SHM%	IGLV	IGLJ	LCDR3	LCDR3 length	SHM%	Origin of donor
H4	IGHV3-7	IGHD5-18	IGHJ3	ASQLWLRGAFDI	12	0						Healthy donor H4
H5	IGHV3-11	IGHD5-18	IGHJ4	ARAKWLRGHFDY	12	0						Healthy donor H5
H14	IGHV4-31	IGHD5-18	IGHJ3	ARKGWLRGAFDI	12	0	IGLV6-57	IGLJ3	QSYDSSNVV	9	0	Healthy donor H14
H16	IGHV4-39	IGHD5-12	IGHJ4	ARREWLRGPFDY	12	0						Healthy donor H16
H18	IGHV4-61	IGHD5-18	IGHJ4	ARQLWLRGRFDY	12	1.3						Healthy donor H18

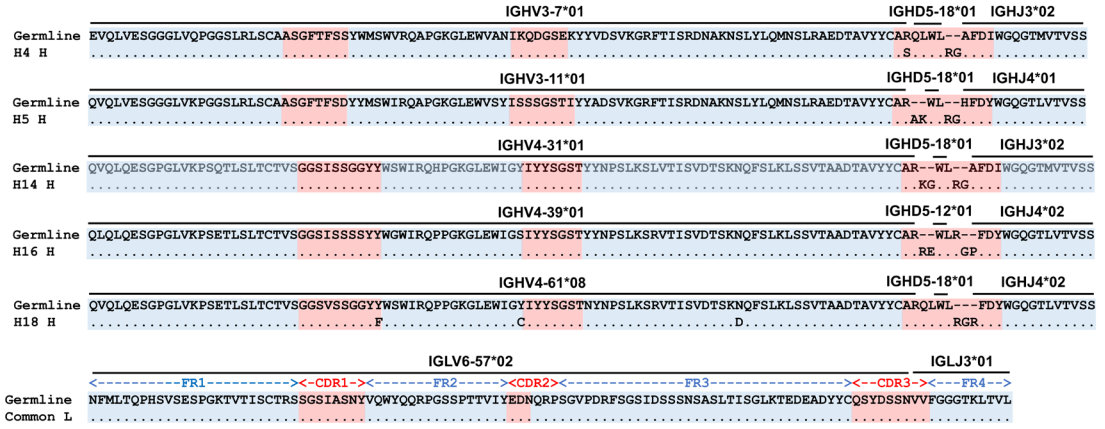

Fig. S10 | Characteristics of the recombinational mAbs derived from healthy donors. a, Summary of genetic features of the 5 recombinational mAbs derived from healthy donors. Sequence analysis of the mAbs was performed using IMGT/V-QUEST. **b**, Heavy chain sequence alignments of the 5 recombinational mAbs with their germline gene sequences. Amino acid residues identical to their germline sequence are dotted. **c**, Sequence alignment of the common light chain shared by the 5 recombinational mAbs with its germline gene sequence.

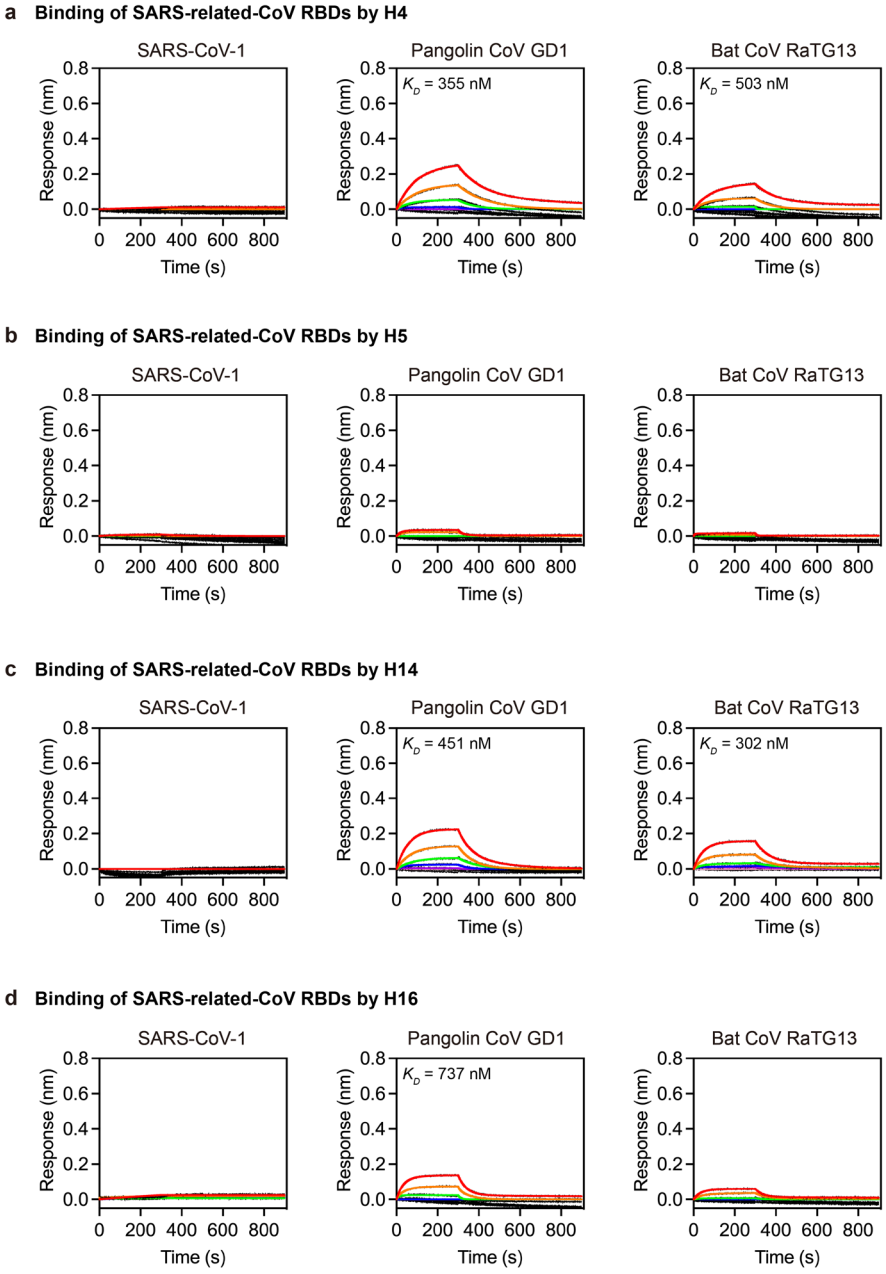

Fig. S11 | SARS-related-CoV RBD binding by the recombinational mAbs. **a-d**, Binding of H4 (**a**), H5 (**b**), H14 (**c**) and H16 (**d**) to SARS-CoV-1, Pangolin CoV GD1, and Bat CoV RaTG13 RBDs were measured by BLI. 2-fold serially diluted RBD solutions (200 nM to 3.125 nM) were used in the BLI assays. Kinetic parameters are summarized in **Supplementary Table 3**.

Barmada, A., Klein, J., Ramaswamy, A., Brodsky, N.N., Jaycox, J.R., Sheikha, H., Jones, K.M., Habet, V., Campbell, M., Sumida, T.S., et al. (2023). Cytokinopathy with aberrant cytotoxic lymphocytes and profibrotic myeloid response in SARS-CoV-2 mRNA vaccine-associated myocarditis. Science Immunology *8*.

Dugan, H.L., Stamper, C.T., Li, L., Changrob, S., Asby, N.W., Halfmann, P.J., Zheng, N.Y., Huang, M., Shaw, D.G., Cobb, M.S., et al. (2021). Profiling B cell immunodominance after SARS-CoV-2 infection reveals antibody evolution to non-neutralizing viral targets. Immunity *54*, 1290-1303 e1297. 10.1016/j.immuni.2021.05.001.

Ferreira-Gomes, M., Kruglov, A., Durek, P., Heinrich, F., Tizian, C., Heinz, G.A., Pascual-Reguant, A., Du, W., Mothes, R., Fan, C., et al. (2021). SARS-CoV-2 in severe COVID-19 induces a TGF-beta-dominated chronic immune response that does not target itself. Nat Commun *12*, 1961. 10.1038/s41467-021-22210-3.

He, P., Liu, B., Gao, X., Yan, Q., Pei, R., Sun, J., Chen, Q., Hou, R., Li, Z., Zhang, Y., et al. (2022). SARS-CoV-2 Delta and Omicron variants evade population antibody response by mutations in a single spike epitope. Nat Microbiol *7*, 1635-1649. 10.1038/s41564-022-01235-4.

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., and Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature *581*, 215-220. 10.1038/s41586-020-2180-5.

Matsuyama, S., and Taguchi, F. (2009). Two-Step Conformational Changes in a Coronavirus Envelope Glycoprotein Mediated by Receptor Binding and Proteolysis. Journal of Virology.

Niu, X., Li, S., Li, P., Pan, W., Wang, Q., Feng, Y., Mo, X., Yan, Q., Ye, X., Luo, J., et al. (2020). Longitudinal Analysis of T and B Cell Receptor Repertoire Transcripts Reveal Dynamic Immune Response in COVID-19 Patients. Front Immunol *11*, 582010. 10.3389/fimmu.2020.582010.

Ren, X., Wen, W., Fan, X., Hou, W., Su, B., Cai, P., Li, J., Liu, Y., Tang, F., Zhang, F., et al. (2021). COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. Cell *184*, 1895-1913 e1819. 10.1016/j.cell.2021.01.053.

Walls, A.C., Xiong, X., Park, Y.J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., et al. (2019). Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. Cell *176*, 1026-1039 e1015. 10.1016/j.cell.2018.12.028.

Wrobel, A.G., Benton, D.J., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and Gamblin, S.J. (2020). SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. Nat Struct Mol Biol *27*, 763-767. 10.1038/s41594-020-0468-7.

Xiong, X., Qu, K., Ciazynska, K.A., Hosmillo, M., Carter, A.P., Ebrahimi, S., Ke, Z., Scheres, S.H.W., Bergamaschi, L., Grice, G.L., et al. (2020). A thermostable, closed SARS-CoV-2 spike protein trimer. Nat Struct Mol Biol *27*, 934-941. 10.1038/s41594-020-0478-5.

Yan, Q., He, P., Huang, X., Luo, K., Zhang, Y., Yi, H., Wang, Q., Li, F., Hou, R., Fan, X., et al. (2021). Germline IGHV3-53-encoded RBD-targeting neutralizing antibodies are commonly present in the antibody repertoires of COVID-19 patients. Emerg Microbes Infect *10*, 1097-1111. 10.1080/22221751.2021.1925594.

Yu, H., Liu, B., Zhang, Y., Gao, X., Wang, Q., Xiang, H., Peng, X., Xie, C., Wang, Y., Hu, P., et al. (2023). Somatically hypermutated antibodies isolated from SARS-CoV-2 Delta infected patients cross-neutralize heterologous variants. Nat Commun *14*, 1058. 10.1038/s41467-023-36761-0.

1459 Yuan, M., Wu, N.C., Zhu, X., Lee, C.-C.D., So, R.T.Y., Lv, H., Mok, C.K.P., and Wilson, I.A.
1460 (2020a). A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2
1461 and SARS-CoV. Science, 630-633.
1462 Yuan, M., Wu, N.C., Zhu, X., Lee, C.-C.D.C.D., So, R.T.Y.Y., Lv, H., Mok, C.K.P.P., and Wilson,
1463 I.A. (2020b). A highly conserved cryptic epitope in the receptor binding domains of SARS-
1464 CoV-2 and SARS-CoV. Science.
1465 Zhang, F., Gan, R., Zhen, Z., Hu, X., Li, X., Zhou, F., Liu, Y., Chen, C., Xie, S., Zhang, B., et al.
1466 (2020). Adaptive immune responses to SARS-CoV-2 infection in severe versus mild
1467 individuals. Signal Transduct Target Ther *5*, 156. 10.1038/s41392-020-00263-y.
1468 Zhang, Y., Yan, Q., Luo, K., He, P., Hou, R., Zhao, X., Wang, Q., Yi, H., Liang, H., Deng, Y., et al.
1469 (2022). Analysis of B Cell Receptor Repertoires Reveals Key Signatures of the Systemic B Cell
1470 Response after SARS-CoV-2 Infection. Journal of Virology *96*, 1-23.
1471

1472