

1 **Synthetic repertoires derived from convalescent COVID-19 patients enable discovery of**  
2 **SARS-CoV-2 neutralizing antibodies and a novel quaternary binding modality**

3 Authors: Jule Goike<sup>1,2#</sup>, Ching-Lin Hsieh<sup>2#</sup>, Andrew Horton<sup>1,2#</sup>, Elizabeth C. Gardner<sup>1,2#</sup>, Foteini  
4 Bartzoka<sup>1,2</sup>, Nianshuang Wang<sup>2</sup>, Kamyab Javanmardi<sup>1,2</sup>, Andrew Herbert<sup>8</sup>, Shawn Abbassi<sup>8</sup>,  
5 Rebecca Renberg<sup>4</sup>, Michael J. Johanson<sup>5</sup>, Jose A. Cardona<sup>1</sup>, Thomas Segall-Shapiro<sup>3</sup>, Ling  
6 Zhou<sup>2</sup>, Ruth H. Nissly<sup>11</sup>, Abhinay Gontu<sup>11</sup>, Michelle Byrom<sup>1</sup>, Andre C. Maranhao<sup>1</sup>, Anna M.  
7 Battenhouse<sup>1</sup>, Varun Gejji<sup>5</sup>, Laura Soto-Sierra<sup>5,13</sup>, Emma R. Foster<sup>5,13</sup>, Susan L. Woodard<sup>5</sup>,  
8 Zivko L. Nikolov<sup>5,13</sup>, Jason Lavinder<sup>2,6</sup>, Will N. Voss<sup>2</sup>, Ankur Annapareddy<sup>1,2,3</sup>, Gregory C.  
9 Ippolito<sup>1,2,10</sup>, Andrew D. Ellington<sup>1,2,7</sup>, Edward M. Marcotte<sup>1,2</sup>, Ilya J. Finkelstein<sup>1,2</sup>, Randall A.  
10 Hughes<sup>1,2,3</sup>, James M. Musser<sup>9</sup>, Suresh V. Kuchipudi<sup>11</sup>, Vivek Kapur<sup>12</sup>, George Georgiou<sup>1,2,6,7,10</sup>,  
11 John M. Dye<sup>8</sup>, Daniel R. Boutz<sup>1,2,3\*</sup>, Jason S. McLellan<sup>2\*</sup>, and Jimmy D. Gollihar<sup>1,2,3,9\*</sup>

12 Affiliations:

- 13 1. Center for Systems and Synthetic Biology, Department of Molecular Biosciences, The  
14 University of Texas at Austin, Austin, TX, USA
- 15 2. Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX,  
16 USA
- 17 3. CCDC Army Research Laboratory-South, Austin, TX, USA
- 18 4. CCDC Army Research Laboratory, Biotechnology Branch, Adelphi, MD, USA
- 19 5. National Center for Therapeutics Manufacturing, Texas A&M University, College Station,  
20 TX, USA
- 21 6. Department of Chemical Engineering, The University of Texas at Austin, Austin, TX,  
22 USA
- 23 7. Department of Chemistry, The University of Texas at Austin, Austin, TX, USA
- 24 8. U.S. Army Medical Research Institute of Infectious Diseases, Frederick, MD, USA
- 25 9. Center for Molecular and Translational Human Infectious Diseases Research,  
26 Department of Pathology and Genomic Medicine, Houston Methodist Research Institute,  
27 Houston Methodist Hospital, Houston, TX, USA
- 28 10. Department of Oncology, Dell Medical School, The University of Texas at Austin, Austin,  
29 TX, USA
- 30 11. Department of Veterinary and Biomedical Science and Animal Diagnostic  
31 Laboratory, The Pennsylvania State University, University Park, Pennsylvania,  
32 USA
- 33 12. Department of Animal Science and the Huck Institutes of the Life Sciences, The  
34 Pennsylvania State University, University Park, Pennsylvania, USA
- 35 13. Department of Agricultural and Biological Engineering, Texas A&M University,  
36 College Station, TX, USA

37  
38 # Authors contributed equally  
39 \* Corresponding authors

40 Contact: [jimmy.d.gollihar2.civ@mail.mil](mailto:jimmy.d.gollihar2.civ@mail.mil); [jmclellan@austin.utexas.edu](mailto:jmclellan@austin.utexas.edu); [dboutz@gmail.com](mailto:dboutz@gmail.com)

41 **Abstract**

42

43 The ongoing evolution of SARS-CoV-2 into more easily transmissible and infectious

44 variants has sparked concern over the continued effectiveness of existing therapeutic antibodies

45 and vaccines. Hence, together with increased genomic surveillance, methods to rapidly develop

46 and assess effective interventions are critically needed. Here we report the discovery of SARS-

47 CoV-2 neutralizing antibodies isolated from COVID-19 patients using a high-throughput platform.

48 Antibodies were identified from unpaired donor B-cell and serum repertoires using yeast surface

49 display, proteomics, and public light chain screening. Cryo-EM and functional characterization of

50 the antibodies identified N3-1, an antibody that binds avidly ( $K_{d,app} = 68$  pM) to the receptor binding

51 domain (RBD) of the spike protein and robustly neutralizes the virus *in vitro*. This antibody likely

52 binds all three RBDs of the trimeric spike protein with a single IgG. Importantly, N3-1 equivalently

53 binds spike proteins from emerging SARS-CoV-2 variants of concern, neutralizes UK variant

54 B.1.1.7, and binds SARS-CoV spike with nanomolar affinity. Taken together, the strategies

55 described herein will prove broadly applicable in interrogating adaptive immunity and developing

56 rapid response biological countermeasures to emerging pathogens.

## 57 **Introduction**

58           The rapid global dissemination of the severe acute respiratory syndrome coronavirus 2  
59 (SARS-CoV-2)<sup>1</sup>, the cause of coronavirus disease 19 (COVID-19)<sup>2</sup>, has highlighted our extreme  
60 vulnerability to novel microbial threats. The speed of SARS-CoV-2 transmission and absence of  
61 widespread adaptive immunity created a pandemic that overwhelmed the international medical  
62 community. This situation was exacerbated by the scarcity of treatment options, especially early  
63 in the pandemic. Functional immune repertoire analysis has the potential to efficiently address  
64 this scarcity. By analyzing primary immune responses directed towards emerging pathogens,  
65 newly elicited antibodies can be identified and rapidly deployed to treat patients.

66           The COVID-19 pandemic has stimulated global research efforts to identify SARS-CoV-2-  
67 neutralizing antibodies for therapeutic and prophylactic applications. Initial attempts to identify  
68 neutralizing antibodies focused on screening extant antibodies elicited against previous SARS-  
69 CoV strains. These efforts were largely unsuccessful owing to limited cross-reactivity<sup>3</sup>, primarily  
70 due to SARS-CoV-2's sequence divergence in the receptor-binding domain (RBD) of its trimeric  
71 spike protein<sup>4</sup>.

72           The ectodomain (ECD) of the SARS-CoV-2 spike (S) protein is essential for initial binding  
73 and subsequent entry of the virus into human cells and has been the primary target for  
74 therapeutics and vaccine formulations. The ECD consists of the S1 subunit, containing an N-  
75 terminal domain (NTD) and RBD, and the S2 subunit, containing the fusion machinery that  
76 mediates entry into host cells. The RBD initiates attachment through interaction with angiotensin-  
77 converting enzyme 2 (ACE2)<sup>5-7</sup>. The functional significance of the ACE2-RBD interaction makes  
78 the RBD a prime target for neutralizing antibodies<sup>8-10</sup>. Targeting this domain increases the  
79 selective pressure on the RBD, which may promote the emergence of escape mutants that

80 maintain virulence. Indeed, neutralizing antibodies targeting a single epitope can induce virus  
81 escape in cell culture, quickly rendering antibodies ineffective<sup>11</sup>.

82 The FDA approved antibody cocktail of REGN10933 and REGN10987<sup>12</sup> targets distinct  
83 epitopes of the RBD and sustains some neutralization activity against new SARS-CoV-2 variants  
84 B.1.1.7 and B.1.351, originally identified in the UK and South Africa, respectively<sup>13–15</sup>. Recently,  
85 a deep mutational scan of the RBD found that the single amino acid mutation E406W increased  
86 the IC<sub>50</sub> of REGN10933 by more than 300-fold and also detrimentally affected REGN10987  
87 binding<sup>16</sup>. The scan also identified amino acid changes that compromised the epitope of Eli Lilly's  
88 antibody LY-CoV016<sup>16</sup>. Moreover, the mass sequencing of 5,085 SARS-CoV-2 genomes from the  
89 Houston metropolitan area published in September 2020 already identified numerous spike  
90 protein mutations that affect the existing neutralizing antibodies' abilities to bind to their epitopes<sup>17</sup>.

91 The B.1.1.7 (UK), B.1.1.248 (Brazil), and B.1.351 (South Africa) variants are of special  
92 concern because they have each accumulated multiple spike protein mutations. These mutations  
93 are located in the NTD, RBD, and furin cleavage site. NTD-directed antibodies in particular show  
94 reduced or abolished binding to these strains<sup>14</sup>. More concerning than the loss of binding ability,  
95 however, is the reduced neutralizing activity of convalescent plasma from patients infected in the  
96 spring of 2020<sup>14,18</sup>. Furthermore, currently available vaccines are largely based on an earlier  
97 prefusion-stabilized spike variant (Wuhan). When assayed, sera from patients vaccinated with the  
98 Moderna and Pfizer/BioNTech vaccines showed a significant decrease in neutralization activity  
99 towards strain B.1.351, while B.1.1.7 was only mildly affected<sup>14</sup>. It is therefore critical to identify  
100 neutralizing antibodies to a broad spectrum of non-overlapping epitopes on the spike protein to  
101 achieve highly potent and persistent neutralization<sup>11</sup>.

102 To this end, we developed a multi-pronged antibody discovery and informatics strategy  
103 involving both proteomic analysis of donor sera and selection of combinatorially paired heavy and  
104 light chain (VH-VL) libraries from donor B-cell receptor repertoires. Specifically, we used Ig-Seq  
105 proteomics to identify candidate variable heavy chains (VHs) of serum antibodies binding SARS-  
106 CoV-2 RBD or ECD. We identified productive light chain pairs for these heavy chains either from  
107 a bioinformatically derived set of “public light chains” (PLCs) or through combinatorial pairing with  
108 donor light chain repertoires via yeast surface display (YSD) selection. In addition to Ig-Seq  
109 interrogation of circulating IgGs, we employed YSD for selection of high-affinity antibodies by  
110 combinatorial display of all donor VH-VL pairs as Fab fragments through multiple rounds of  
111 selection.

112

113 Together, this holistic strategy allowed us to probe the secreted circulating antibody  
114 repertoire and the nascent cellular repertoire of a primary immune response. By interrogating the  
115 combined repertoire, we gained valuable insight into the antibody response elicited by SARS-  
116 CoV-2 and discovered neutralizing antibodies to multiple distinct domains of the spike protein.  
117 We report the structural characterization of two such neutralizers, including a highly potent  
118 neutralizing antibody (N3-1) that binds a quaternary epitope of the trimeric spike protein via a  
119 novel binding modality. Finally, we tested binding to emerging spike variants of SARS-CoV-2 and  
120 demonstrated N3-1 maintains robust binding to these variants and neutralizes B.1.1.7.

121

122

## 123 **Results**

### 124 **Ig-Seq analysis of the serological repertoire reveals candidate heavy chains**

125 Ig-Seq provides a proteomic snapshot of the serum antibody repertoire by using mass  
126 spectrometry to identify heavy chain complementarity-determining region 3 (CDR3) peptides of  
127 antigen-enriched antibodies (**Fig. 1**)<sup>19</sup>. To identify relevant SARS-CoV-2 antibodies present in the

128 serological repertoire, we first isolated antibodies from the serum of two donors by antigen  
129 enrichment chromatography using immobilized SARS-CoV-2 RBD and ECD fragments. We then  
130 employed Ig-Seq mass spectrometry to identify abundant IgG clonotypes. A total of 15 and 21  
131 unique clonotypes were identified with high confidence from donors 1 and 2, respectively  
132 **(Extended Data Figs. 1-2)**. This is notably lower than what has been observed in previous Ig-  
133 Seq studies<sup>19-22</sup>, perhaps because the previous studies involved donors after boost vaccination  
134 which is likely to elicit a more robust and diverse response than the primary response to natural  
135 SARS-CoV-2 infection observed here. Regardless, the clonotypes identified provided a sufficient  
136 set of VH candidates to express and characterize as recombinant anti-SARS-CoV-2 monoclonal  
137 antibodies (mAbs) in combination with suitable light chains.

138

### 139 **Public light chains recover productive pairs for most VHs**

140 While Ig-Seq provides valuable information for antibody discovery, it does not identify light  
141 chain partners, which usually requires additional laborious techniques. To expedite light chain  
142 discovery, we analysed a published dataset of natively paired memory B-cell sequences<sup>19</sup> from  
143 three healthy donors and bioinformatically derived a panel of nine light chains that show high  
144 frequencies of productive pairings with a diverse set of VH sequences (**Fig. 1c**).

145 At the time of blood draw, the donors in the dataset were healthy and asymptomatic. This  
146 lack of polarization and relative homeostasis potentially affords broader insight into common VH-  
147 VL pairings, independent of antigen specificity. Indeed, many of the observed light chains paired  
148 promiscuously with a diverse set of VH genes. The V gene usage of undiscerning light chains  
149 was remarkably consistent among the three donors and therefore were termed “public light  
150 chains” (PLCs). From our secondary analysis of this dataset, we identified six kappa V genes that  
151 accounted for 61% of the kappa VH-VL pairs and three lambda V genes representing 22% of

152 lambda pairs (**Fig. 2a**). Together, these nine public light chain (PLC) V genes accounted for 43%  
153 of observed VH-VL pairs in the analyzed repertoires.

154 We hypothesized that this abundance could indicate an enhanced ability to form  
155 productive VH-VL pairs independent of VH sequence and antigen specificity. To test this  
156 hypothesis, we constructed germline versions of these light chains with their most frequently  
157 observed J gene as published on the iRepertoire website<sup>20-23</sup> (**Extended Data Table 1**) and  
158 obtained full-length light chains for each PLC. SARS-CoV-2 specific Ig-Seq heavy chains were  
159 then expressed with each PLC as a full-length IgG1. EC50 determination via ELISA using  
160 recombinant spike protein demonstrated that of the 22 VHs tested, all paired productively with  
161 one or more of the PLCs, resulting in several highly potent and specific antibodies (**Fig. 2b**).  
162 Through this simple and rapid screening method we identified a panel of functional antibody  
163 candidates for SARS-CoV-2 neutralization testing.

164

### 165 **Yeast Surface Display-IgSeq identifies additional candidate antibodies**

166 Our PLC screening strategy provides a method to rapidly generate functionally paired  
167 antibodies, but the decreased sequence space of the nine PLCs may limit the ability to efficiently  
168 identify potent binders. Thus, to search for additional high-affinity antibodies, we randomly paired  
169 24 Ig-Seq heavy chains with donor-derived light chain libraries for yeast surface display (YSD)  
170 selections (**Fig. 1b**), which allowed us to efficiently screen Fabs for each heavy chain/light chain  
171 library combination. In YSD, heavy/light chain pairs were displayed as Fabs on the surface of a  
172 humanized yeast strain that expresses human protein chaperones to improve antibody  
173 expression. We then selected antigen-binding Fabs using two-color fluorescence cell sorting to  
174 maximize Fab expression and binding to either SARS-CoV-2 spike (S-2P) or RBD (**Extended**  
175 **Data Figs. 3-4**). As Fab enrichment over successive rounds of selection indicates antigen

176 specificity, we performed three rounds and identified and quantified productive VH-VL pairs based  
177 on MinION nanopore sequencing of Fab amplicons with sequence error correction and  
178 complementary Illumina reads (**Fig. 2c, Extended Data Figs. 5-6**). When light chain preference  
179 was compared, we observed extensive correspondence between YSD selection and PLC  
180 screening.

181

### 182 **Parallel selections *via* yeast surface display yield functionally diverse antibodies**

183 Much of the early stage primary antibody response is inaccessible to Ig-Seq discovery, as  
184 detection is limited to secreted antibody proteins with sufficient abundance in serum<sup>24</sup>. To broaden  
185 our search for neutralizing antibodies, we employed our YSD platform to directly mine the cellular  
186 repertoire for high affinity B-cell receptor sequences (**Fig. 1d**). We expressed and displayed large  
187 VH-VL libraries ( $>10^7$ ) created by combinatorial pairing of donor IgG heavy and light chain  
188 amplicons. To maximize the functional diversity of selected antibodies, we performed three  
189 rounds of selection on the initial libraries, using various gating strategies to ensure phenotypic  
190 diversity (i.e. a variety of antibody expression and antigen binding strengths). Ultimately, the  
191 parallelized selection strategy yielded populations with distinct binding characteristics as  
192 confirmed by cytometric phenotyping of single clones (**Extended Data Figs. 4, 7**). We observed  
193 progressive enrichment of specific antibodies with the dominant clone from each course of  
194 selections generally approaching 20% of the total population after round three (**Extended Data**  
195 **Fig. 8**). Individual clones from enriched populations were selected as candidates for further  
196 characterization either at random or based on their abundance from the bioinformatic analysis.  
197 Importantly, YSD detected enrichment of VH sequences that were not observed by Ig-Seq  
198 proteomics, demonstrating the advantages and increased sensitivity of this multi-pronged  
199 approach to antibody discovery.



200

## 201 **Non-cognate VH-VL pairs are potent neutralizing antibodies to SARS-CoV-2**

202 We initially screened the isolated IgG against the ECD, NTD, and RBD of the spike protein  
203 using ELISA to identify the most promising candidates for authentic live virus neutralization of  
204 SARS-CoV-2. We prioritized non-RBD antibodies in our screens in order to gain insights into  
205 alternative neutralization mechanisms. From binding data of the full-length IgGs, we selected 94  
206 candidate VH-VL pairs for live virus neutralization, and 37 pairs successfully neutralized authentic  
207 SARS-CoV-2/WA1 (**Fig. 3a, Extended Data Table 2**). Curve fitting calculated IC<sub>50</sub> values as low  
208 as 18 pM, with seven antibodies showing sub-nanomolar values (8-32, 18 pM; N3-1, 251.9 pM;  
209 8-131, 300 pM; 8-114, 350 pM; 12C8, 836 pM; A7V3, 955 pM; and 7-6, 977 pM). Twenty-three  
210 neutralizing antibodies had IC<sub>50</sub> values less than 100 nM (**Fig. 3**). Our high affinity NTD VH-VL  
211 pairs (8-32, 8-131, 8-114, 12C8, A7V3, 7-6, 4C7, 4C8, 7A8) showed strong preference for IGHV1-  
212 24 gene usage. PLC screening successfully identified VL partners for both YSD and Ig-Seq  
213 derived VHs (**Fig. 3b-e**). Seven of the mAbs were shown to be directed to the S2 domain. N3-1  
214 possesses a highly potent RBD-directed IGHV4-31 heavy chain and displayed strong  
215 neutralization activity with several PLCs as well as YSD VLs (**Extended Data Table 2**).

216

## 217 **Common binding mode of an NTD-directed antibody**

218

219 A7V3 was of interest because of the different donor origins of the VH and VL. Its IGHV1-  
220 24 was discovered in both IgSeq and YSD from donor 1. However, the IGLV1-51 was mapped to  
221 donor 2 and is also related to PLC8. The VH of A7V3 showed a strong preference to PLC8 in  
222 ELISA screening, and related light chains appear in six other neutralizing VH-VL pairs, commonly  
223 pairing with IGHV1-24 VHs. To ascertain its potential binding mode and contributions from both  
224 variable regions, we determined a cryo-EM structure of A7V3 Fab complexed with SARS-CoV-2  
225 S-ECD. The initial 3D reconstruction from 715,398 particles had a global resolution of 3.0 Å (**Fig.**

226 **4a)**. Although the map revealed an A7V3 Fab molecule bound to each NTD of the trimeric spike,  
227 only one Fab defined clear density. Further 3D classification was performed to resolve potential  
228 heterogeneity in the particle set. One class was obtained consisting of one fourth of the total  
229 particles and exhibited defined density for an NTD and its bound Fab. A local refinement of this  
230 map focused on the NTD and Fab yielded a reconstruction with a well-resolved binding interface  
231 (**Fig. 4b**). Similar to the first structurally defined NTD-targeted antibody 4A8<sup>25</sup>, the heavy chain of  
232 A7V3 plays a dominant role in binding the N3 loop (residues 141-156) and N5 loop (residues 246-  
233 260) of the NTD. The light chain of A7V3 makes limited contact with the NTD, and is therefore  
234 likely to play a structural role in positioning the heavy chain. H-CDR1, H-CDR3, and L-CDR2  
235 collectively form a concave surface packed tightly against the N5 loop. Specifically, Trp100d in  
236 H-CDR3, and both Tyr49 and Pro55 in L-CDR2, form a hydrophobic cage to enclose Pro251 at  
237 the tip of the N5 loop (**Fig. 4c**). Glu31 from H-CDR1 is expected to form a salt bridge with Arg246,  
238 and Asp101 from H-CDR3 is expected to form a hydrogen bond with the hydroxyl group of Tyr248.  
239 Pro97 and Phe98 from H-CDR3 insert into a groove flanked by the N3 and N5 loops (**Fig. 4d**). As  
240 opposed to the major binding interface on the N5 loop, A7V3 only makes a minor contact with the  
241 N3 loop, mainly mediated by H-CDR1 and H-CDR2. Akin to a conserved Phe from H-CDR2 of  
242 4A8, Phe51 contacts Lys147 via a  $\pi$ -cation interaction (**Fig. 4d-e**). In stark contrast, H-CDR3 of  
243 A7V3 is much shorter than its counterpart in 4A8, which contains a Phe that forms a  $\pi$ - $\pi$  stacking  
244 interaction with Trp152 on the N3 loop. Collectively, the structural characterization of A7V3  
245 demonstrates that our semi-synthetic approach is capable of yielding a novel neutralizing antibody  
246 that binds to the NTD using an established modality.

247

248 **An RBD-targeted antibody exhibits a novel quaternary binding mode**

249 Structurally defined RBD-targeted antibodies or nanobodies complexed with SARS-CoV-  
250 2 S may be generally grouped into two types based on binding sites. The first type, mainly  
251 discovered from the IGHV3-53 germline family, recognizes the ACE2-binding site and directly  
252 blocks host receptor engagement (e.g. antibodies CC12.1<sup>26</sup>, C105<sup>27</sup> and VHH E<sup>28</sup>). The second  
253 type, from a more diverse germline family, recognizes a relatively conserved region on the RBD  
254 that is mostly buried and contacts a neighboring RBD in the down conformation (e.g. C135<sup>29</sup>,  
255 S309<sup>30</sup> and VHH V<sup>28</sup>). From negative stain EM analysis of the IgG N3-1-spike complex, we  
256 observed density for two Fabs, likely derived from the same IgG molecule, bound to multiple  
257 RBDs on a single trimeric spike (**Extended Data Figs. 12**). We also found the affinity of N3-1 IgG  
258 to SARS-CoV-2 spike is nearly 1000-fold stronger than N3-1 Fab to spike (**Extended Data Fig.**  
259 **13**).

260  
261 To investigate this unique binding mode, we determined a cryo-EM structure of N3-1 Fab  
262 bound to SARS-CoV-2 S-ECD to a global resolution of 2.8 Å (**Fig. 5a**). We observed that two N3-  
263 1 Fabs bind to a single trimeric spike with one Fab binding to RBD in the up conformation and the  
264 other Fab simultaneously engaging two RBDs: one in the up conformation and one in the down  
265 conformation. We performed focused refinement on the Fab bound to the two RBDs (**Fig. 5b**),  
266 which substantially improved the interpretability of the map in this region. The well-resolved Fab-  
267 RBD binding interface revealed two completely distinct epitopes on RBD-up and RBD-down (**Fig.**  
268 **5c**). For RBD-up interaction, contacts are made by the Fab H-CDR1, H-CDR3 and L-CDR2, which  
269 together bury 862 Å<sup>2</sup> surface area. For the RBD-down interaction, contacts are made by the Fab  
270 H-CDR2, H-CDR3 and L-CDR3, which buries 696 Å<sup>2</sup> surface area. Notably, the relatively long H-  
271 CDR3 loop (18 a.a.) engages both RBDs via hydrophobic and polar interactions. H-CDR3  
272 residues Tyr98, Phe99 and Arg100a pack against a hydrophobic pocket formed by Tyr369,  
273 Phe377, Lys378 and Pro384 on RBD-up, which are highly conserved between SARS-CoV and

274 SARS-CoV-2. This pocket is barely exposed when RBD is in the down conformation and is part  
275 of the shared epitopes targeted by cross-reactive antibodies CR3022 and COVA1-16<sup>31</sup>. Lys378,  
276 from the upper ridge of the pocket, forms a cation- $\pi$  interaction with Tyr98 of H-CDR3, and its  
277 amine group is expected to form a salt bridge with Asp100h of H-CDR3. In addition, main chain  
278 atoms of Cys379 and Tyr369 form hydrogen bonds with Phe99 and Arg100a from H-CDR3,  
279 strengthening this primary Fab binding interface on RBD-up. Furthermore, the sidechain  
280 guanidinium of Arg408 on RBD-up likely has polar interactions with Tyr49 and Glu55 from L-  
281 CDR2, which along with H-CDR1, constitutes the secondary binding interface on RBD-up.

282  
283 In contrast, the N3-1 binding site on RBD-down overlaps with the ACE2-binding site,  
284 where 11 of 19 N3-1 epitope residues are also involved in ACE2 binding (**Fig. 5d**). Phe486 on  
285 RBD-down inserts into a hydrophobic pocket formed by Trp96 (L-CDR3), Tyr52, Tyr58 (H-CDR2),  
286 Arg100e and Val100g (H-CDR3). The sidechain guanidinium of Arg100e not only forms a  
287 hydrogen bond with Tyr489, but also contacts Phe486 through cation- $\pi$  interactions. In addition,  
288 Asn487 and Gln 493 on RBD-down are expected to form polar interactions with Ser93 of L-CDR3  
289 and Thr57 of H-CDR2. The angle of approach of N3-1 prevents ACE2 binding by both trapping  
290 one RBD in the down position, thereby preventing exposure of the ACE2 binding site, as well as  
291 by sterically inhibiting ACE2 access to the bound 'up' RBD. (**Fig. 5e**). Collectively, the N3-1  
292 antibody engages an extensive quaternary epitope on neighboring RBDs through a novel binding  
293 modality.

294

### 295 **mAb N3-1 binds robustly to spike proteins from emerging variants of concern**

296 Next, we tested whether binding of A7V3 (NTD-targeting) or N3-1 (RBD-targeting) was  
297 detrimentally affected by mutations emerging in SARS-CoV-2 variants of concern (VOC). We  
298 assessed binding to the 12 most abundant SARS-CoV-2 variants from Houston<sup>17</sup> and did not

299 observe detrimental losses in binding avidity (**Extended Data Table 5**). Next, we used  
300 mammalian surface display of spike variants to test binding of N3-1 to mutations that were  
301 previously shown to reduce or abrogate binding to one or both of the Regeneron antibodies  
302 (REGN10987 and REG10933) as well as a commonly found mutation in Australia (S477N).  
303 Strikingly, N3-1 was able to bind each of the mutant spike variants. N3-1 binding was slightly  
304 enhanced by 10 of the 12 mutant spike proteins tested (L455A, G476D, E406W, Q493F, K444A,  
305 K444Q, V445A, G446A, G446V, and S477N) using this assay. Variant Y453A was  
306 indistinguishable from binding to the prefusion stabilized SARS-CoV-2 S D614G, and F486K was  
307 the only mutation shown to marginally reduce binding (~6%) (**Extended Data Fig. 15**). We then  
308 tested binding to B.1.1.7, B.1.1.248, and B.1.351 spike protein variants using the same  
309 mammalian surface display assay. Binding of A7V3 was reduced by all three variants, a result  
310 predicted from our structural work. In contrast, N3-1 binding was largely unperturbed by the  
311 accumulated mutations (**Fig. 6a**) of B.1.1.7, B.1.1.248, and B.1.351 spike proteins in this assay  
312 (**Fig. 6b**). To validate the mammalian display screening results, we examined the binding of the  
313 SARS-CoV-2 S Wuhan-Hu-1, and its variants B.1.1.7, and B.1.351 to N3-1 Fab or N3-1 IgG using  
314 surface plasmon resonance (SPR). N3-1 Fab binding to each of the variants was only able to be  
315 fit to a heterogenous binding model (**Extended Data Fig. 13**), consistent with the structural  
316 studies. N3-1 IgG exhibited comparable binding to Wuhan-Hu-1 and its variants, with 67.9 pM  
317 apparent affinity to Wuhan-Hu-1, 296 pM apparent affinity to B.1.1.7, and 300 pM apparent affinity  
318 to B.1.351 (**Fig. 6c-e**). We also tested binding via SPR to SARS-CoV spike protein and discovered  
319 that N3-1 IgG cross-reacts with SARS-CoV spike with 28 nM affinity (**Extended Data Figure 13**).  
320 Finally, we corroborated these findings with live virus neutralization assays with circulating VOC  
321 B.1.1.7. We found that the wildtype WA/1 strain was fully neutralized by both N3-1 and A7V3, but  
322 B.1.1.7 was only neutralized by N3-1 at the same titer as WA/1 (**Fig. 6f**), consistent with our  
323 structure-based predictions.

324

## 325 **Discussion**

326 SARS-CoV-2 and its evolved progeny represent a moving target for therapeutic  
327 development. While the rapid response of the global research community has led to the discovery  
328 of many neutralizing mAbs<sup>32, 25,33</sup> the emergence and widespread dissemination of viral variants  
329 such as the U.K. variant B.1.1.7, Brazilian variant B.1.1.248, and South African variant B.1.351  
330 stresses the need for continued efforts to rapidly identify and develop antibodies that can bind to  
331 diverse epitopes. By combining proteomic analyses of antibody repertoires from patient samples  
332 and yeast-display with public light chain pairing, we developed a high-throughput antibody  
333 discovery strategy that identified 37 neutralizing antibodies that target multiple regions of the  
334 SARS-CoV-2 spike protein, including the RBD, NTD, and the S2 domain, and that broadly bind  
335 multiple viral variants.

336 To accelerate the discovery of neutralizing antibodies, we first performed Ig-Seq proteomic  
337 analyses on patient samples. However, because Ig-Seq does not provide information on relative  
338 affinities, epitope targeting, and light chain pairings, we adopted two strategies for light chain  
339 discovery: a rapid screening method utilizing public light chains (PLCs) and a YSD light chain  
340 selection. Ultimately, we found that both methods independently yielded multiple neutralizing  
341 antibodies, with several of the most potently neutralizing mAbs converging in two or more of our  
342 approaches.

343 The neutralizing antibodies identified relied on eight IGHV genes, including IGHV3-30 and  
344 IGHV1-2, consistent with studies that show these genes are overrepresented in SARS-CoV-2  
345 neutralizing antibodies<sup>32</sup>. In addition, IGHV1-24 was previously found to contribute to a number  
346 of NTD-directed neutralizing antibodies<sup>25,34</sup> and was observed in 10 of the neutralizing antibodies  
347 herein. We found representative neutralizers from each of our discovery strategies, although each

348 method yielded one or more unique IGVH or IGVL genes, validating the high coverage afforded  
349 by this multi-pronged approach.

350         The rapid PLC screening method appears particularly useful not only for expediting  
351 functional testing of Ig-Seq-derived heavy chains, but also to generate high-affinity antibodies.  
352 For instance, N3-1 exhibits a sub-nanomolar IC-50 (252 pM) and binding affinity (68 pM), while  
353 A7V3 has an IC-50 of 950 pM. A7V3 is a particularly interesting demonstration of the utility of  
354 the strategy, because its heavy chain was first identified from one donor, while the light chain was  
355 mapped to another donor. Initial PLC screening indicated a nearly identical IGLC VJ gene (PLC8)  
356 as a productive partner, which further confirmed the utility of PLC screening for productive  
357 antibody VH-VL pairings.

358         We resolved high-resolution cryo-EM structures of these two newly discovered  
359 neutralizing antibodies, and found that N3-1 targets the RBD, whereas A7V3 targets the NTD.  
360 Structural analysis of N3-1 revealed that the L-CDR2 and L-CDR3s were both engaged in binding,  
361 which suggests that our bioinformatically derived PLCs may directly contribute to neutralization.  
362 A7V3 bound to a common neutralization site on the NTD<sup>34</sup>, and ultimately, PLC8 was found to  
363 productively pair with three different neutralizing antibodies that were most commonly associated  
364 with NTD-binding heavy chains. Our work not only establishes the first empirical demonstration  
365 that public light chains can form functional pairings across heavy chains, but also debuts their  
366 utility as a prolific method for high affinity antibody discovery, potentially improving rapid-response  
367 antibody discovery for neo-antigens in future epidemics. Our results further suggest that naturally  
368 paired repertoires may be unnecessary for efficient antibody discovery, in contrast to a recently  
369 published work where cognate pairing was deemed necessary for effective neutralization<sup>35</sup>.

370           The cryo-EM structure of N3-1 is of particular importance because it reveals a novel  
371 binding mechanism at a quaternary epitope of the spike trimer, thus allowing this antibody to bind  
372 the RBD in both the ‘up’ and ‘down’ conformations. This greatly enhances the binding affinity of  
373 the full mAb as it enables binding to all three RBD subunits in the spike protein trimer. As a  
374 consequence of this unique mechanism of action, N3-1 is minimally perturbed by spike mutations  
375 in the newly emerged variants B.1.1.7, B.1.1.248, and B.1.351, where it was shown to still have  
376 low picomolar affinities and neutralized B.1.1.7 with a similar titer to WA/1. N3-1 is also capable  
377 of binding the distantly related SARS-CoV spike protein with low nanomolar affinity (28 nM). Given  
378 the mechanism of binding and the fact that the N3-1 binding epitope for the RBD-down  
379 conformation corresponds to the ACE2-binding site, we fully expect N3-1 to show neutralization  
380 of VOCs not tested here.

381           The strategy we described here will aid in the rapid identification and development of  
382 monoclonal antibodies to extant SARS-CoV-2 variants, and can be readily utilized for identifying  
383 antibody therapeutics for variants yet to be discovered and for other explosive pathogen  
384 zoonoses. Candidate mAbs may be produced in less than 10 days, and have desirable  
385 characteristics – presence in immune responses, public light chain robustness, validation of  
386 expression via yeast display – that may enable production at scale. Because of the modular  
387 nature of the synthetic biology-like discovery process, the overall strategy could also be readily  
388 extended to affinity maturation or directed evolution of tighter binding mAbs to distinct antigens or  
389 epitopes. In the long term, the ability to store, combine, and continually reassess repertoires  
390 should prove broadly useful in not only discovering novel antibodies with potential therapeutic  
391 applications, but also for studying the evolution of adaptive immunity.



392 **Reference**

- 393 1. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat  
394 origin. *Nature* **579**, 270–273 (2020).
- 395 2. Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature*  
396 **579**, 265–269 (2020).
- 397 3. Tian, X. *et al.* Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-  
398 specific human monoclonal antibody. *Emerg. Microbes Infect.* **9**, 382–385 (2020).
- 399 4. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation.  
400 *Science* **367**, 1260–1263 (2020).
- 401 5. Tai, W. *et al.* Characterization of the receptor-binding domain (RBD) of 2019 novel  
402 coronavirus: implication for development of RBD protein as a viral attachment inhibitor and  
403 vaccine. *Cell. Mol. Immunol.* **17**, 613–620 (2020).
- 404 6. Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS  
405 coronavirus. *Nature* **426**, 450–454 (2003).
- 406 7. Yan, R. *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2.  
407 *Science* **367**, 1444–1448 (2020).
- 408 8. Shi, R. *et al.* A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2.  
409 *Nature* **584**, 120–124 (2020).
- 410 9. Yuan, M. *et al.* A highly conserved cryptic epitope in the receptor binding domains of SARS-  
411 CoV-2 and SARS-CoV. *Science* **368**, 630–633 (2020).
- 412 10. Yi, C. *et al.* Key residues of the receptor binding motif in the spike protein of SARS-CoV-  
413 2 that interact with ACE2 and neutralizing antibodies. *Cell. Mol. Immunol.* **17**, 621–630  
414 (2020).
- 415 11. Baum, A. *et al.* Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational  
416 escape seen with individual antibodies. *Science* **369**, 1014–1018 (2020).

- 417 12. Hansen, J. *et al.* Studies in humanized mice and convalescent humans yield a SARS-  
418 CoV-2 antibody cocktail. *Science* **369**, 1010–1014 (2020).
- 419 13. Shen, X. *et al.* SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies  
420 elicited by ancestral Spike vaccines. *bioRxiv* 2021.01.27.428516 (2021)  
421 doi:10.1101/2021.01.27.428516.
- 422 14. Wang, P. *et al.* Increased Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7 to  
423 Antibody Neutralization. *bioRxiv* 2021.01.25.428137 (2021) doi:10.1101/2021.01.25.428137.
- 424 15. Ferrareze, P. A. G. *et al.* E484K as an innovative phylogenetic event for viral evolution:  
425 Genomic analysis of the E484K spike mutation in SARS-CoV-2 lineages from Brazil. *bioRxiv*  
426 2021.01.27.426895 (2021) doi:10.1101/2021.01.27.426895.
- 427 16. Starr, T. N. *et al.* Prospective mapping of viral mutations that escape antibodies used to  
428 treat COVID-19. *Science* (2021) doi:10.1126/science.abf9302.
- 429 17. Long, S. W. *et al.* Molecular Architecture of Early Dissemination and Massive Second  
430 Wave of the SARS-CoV-2 Virus in a Major Metropolitan Area. *mBio* **11**, (2020).
- 431 18. Hoffmann, M. *et al.* SARS-CoV-2 variants B.1.351 and B.1.1.248: Escape from  
432 therapeutic antibodies and antibodies induced by infection and vaccination.  
433 <http://biorxiv.org/lookup/doi/10.1101/2021.02.11.430787> (2021)  
434 doi:10.1101/2021.02.11.430787.
- 435 19. DeKosky, B. J. *et al.* In-depth determination and analysis of the human paired heavy-  
436 and light-chain antibody repertoire. *Nat. Med.* **21**, 86–91 (2015).
- 437 20. Immune Repertoire Sequencing.  
438 <https://irweb.irepertoire.com/nir/showsample?sampleid=21746#>.
- 439 21. Immune Repertoire Sequencing.  
440 <https://irweb.irepertoire.com/nir/showsample?sampleid=21745>.
- 441 22. Immune Repertoire Sequencing.

- 442 <https://irweb.irepertoire.com/nir/showsample?sampleid=21744>.
- 443 23. Immune Repertoire Sequencing.
- 444 <https://irweb.irepertoire.com/nir/showsample?sampleid=21743>.
- 445 24. Lavinder, J. J. *et al.* Identification and characterization of the constituent human serum  
446 antibodies elicited by vaccination. *Proc. Natl. Acad. Sci.* **111**, 2259–2264 (2014).
- 447 25. Chi, X. *et al.* A neutralizing human antibody binds to the N-terminal domain of the Spike  
448 protein of SARS-CoV-2. *Science* **369**, 650–655 (2020).
- 449 26. Yuan, M. *et al.* Structural basis of a shared antibody response to SARS-CoV-2. *Science*  
450 **369**, 1119–1123 (2020).
- 451 27. Barnes, C. O. *et al.* Structures of Human Antibodies Bound to SARS-CoV-2 Spike  
452 Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **182**, 828-842.e16  
453 (2020).
- 454 28. Koenig, P.-A. *et al.* Structure-guided multivalent nanobodies block SARS-CoV-2  
455 infection and suppress mutational escape. *Science* **371**, (2021).
- 456 29. Barnes, C. O. *et al.* SARS-CoV-2 neutralizing antibody structures inform therapeutic  
457 strategies. *Nature* **588**, 682–687 (2020).
- 458 30. Pinto, D. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV  
459 antibody. *Nature* **583**, 290–295 (2020).
- 460 31. Liu, H. *et al.* Cross-Neutralization of a SARS-CoV-2 Antibody to a Functionally  
461 Conserved Site Is Mediated by Avidity. *Immunity* **53**, 1272-1280.e5 (2020).
- 462 32. Xiaojie, S., Yu, L., lei, Y., Guang, Y. & Min, Q. Neutralizing antibodies targeting SARS-  
463 CoV-2 spike protein. *Stem Cell Res.* **50**, 102125 (2021).
- 464 33. Liu, L. *et al.* Potent Neutralizing Antibodies Directed to Multiple Epitopes on SARS-CoV-  
465 2 Spike. *bioRxiv* 2020.06.17.153486 (2020) doi:10.1101/2020.06.17.153486.
- 466 34. Voss, W. N. *et al.* Prevalent, protective, and convergent IgG recognition of SARS-CoV-2

- 467 non-RBD spike epitopes in COVID-19 convalescent plasma. *bioRxiv* 2020.12.20.423708  
468 (2020) doi:10.1101/2020.12.20.423708.
- 469 35. Banach, B. B. *et al.* Paired Heavy and Light Chain Signatures Contribute to Potent  
470 SARS-CoV-2 Neutralization in Public Antibody Responses.  
471 <https://papers.ssrn.com/abstract=3754549> (2020) doi:10.2139/ssrn.3754549.
- 472 36. Wick, R. R., Judd, L. M. & Holt, K. E. Performance of neural network basecalling tools  
473 for Oxford Nanopore sequencing. *Genome Biol.* **20**, 129 (2019).
- 474 37. Bolotin, D. A. *et al.* MiXCR: software for comprehensive adaptive immunity profiling. *Nat.*  
475 *Methods* **12**, 380–381 (2015).
- 476 38. Javanmardi, K. *et al.* Rapid characterization of spike variants via mammalian cell surface  
477 display. *bioRxiv* 2021.03.30.437622 (2021) doi:10.1101/2021.03.30.437622.
- 478 39. Brown, J. F. *et al.* Anti–Ebola Virus Antibody Levels in Convalescent Plasma and Viral  
479 Load After Plasma Infusion in Patients With Ebola Virus Disease. *J. Infect. Dis.* **218**, 555–562  
480 (2018).
- 481 40. Liu, X. *et al.* Neutralizing Aptamers Block S/RBD-ACE2 Interactions and Prevent Host  
482 Cell Infection. *Angew. Chem. Int. Ed Engl.* (2021) doi:10.1002/anie.202100345.
- 483 41. Libster, R. *et al.* Early High-Titer Plasma Therapy to Prevent Severe Covid-19 in Older  
484 Adults. *N. Engl. J. Med.* **384**, 610–618 (2021).
- 485 42. Gontu, A. *et al.* Limited window for donation of convalescent plasma with high live-virus  
486 neutralizing antibody titers for COVID-19 immunotherapy. *Commun. Biol.* **4**, 267 (2021).
- 487 43. Sui, J. *et al.* Potent neutralization of severe acute respiratory syndrome (SARS)  
488 coronavirus by a human mAb to S1 protein that blocks receptor association. *Proc. Natl. Acad.*  
489 *Sci. U. S. A.* **101**, 2536–2541 (2004).
- 490 44. Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single-particle  
491 image processing. *eLife* **7**,.

- 492 45. Xiong, X. *et al.* A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct.*  
493 *Mol. Biol.* **27**, 934–941 (2020).
- 494 46. Henderson, R. *et al.* Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat.*  
495 *Struct. Mol. Biol.* **27**, 925–933 (2020).
- 496 47. McCallum, M., Walls, A. C., Bowen, J. E., Corti, D. & Veesler, D. Structure-guided  
497 covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation.  
498 *Nat. Struct. Mol. Biol.* **27**, 942–949 (2020).
- 499 48. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with  
500 Warp. *Nat. Methods* **16**, 1146–1152 (2019).
- 501 49. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for  
502 rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
- 503 50. Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume  
504 post-processing. *bioRxiv* 2020.06.12.148296 (2020) doi:10.1101/2020.06.12.148296.
- 505 51. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and  
506 analysis. *Protein Sci.* **27**, 14–25 (2018).
- 507 52. Hsieh, C.-L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes.  
508 *Science* **369**, 1501–1505 (2020).
- 509 53. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta*  
510 *Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
- 511 54. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic  
512 structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
- 513 55. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-  
514 resolution electron-density maps. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 519–530 (2018).

515  
516  
517

518 **Online Methods**

519 **Strains and media.** Yeast strain EBY100 (MATa AGA1::GAL1-AGA1::URA3 ura3-52 trp1 leu2-  
520 delta200 his3-delta200 pep4::HIS3 prbd1.6R can1 GAL) was acquired from ATCC (cat. no. MYA-  
521 4941) and used for antibody expression and selection. To improve antibody expression the  
522 human chaperones BIP (binding immunoglobulin protein) and PDI (protein disulfide isomerase)  
523 were genomically integrated as an expression cassette in the HO locus. Yeast were grown in rich  
524 medium (YPD; Takara, cat. no. 630409) or in selective medium for leucine prototrophs after library  
525 transformation (Takara cat. no. 630310). YEP-galactose was used for expression of displayed  
526 antibody libraries (1% yeast extract, 1% bacto-peptone, 0.5% NaCl, 2% galactose, 0.2% glucose).

527 **Antigens and antibodies.** Spike antigen was biotinylated using the EZ-link kit (Thermo Scientific,  
528 cat. no. 21435) and labeled with streptavidin-AF647 (Invitrogen, cat. no. S32357). RBD was  
529 labeled with a mouse anti-human Fc-AF647 (Southern Biotech, cat. no. 9042-31). Fab library light  
530 chains were labeled with anti-FLAG M2-FITC (Sigma, cat. no. F4049).

531 **Donors.** Blood was collected from 3 PCR-confirmed, symptomatic patients. Donors 1 and 2 were  
532 collected at day 12 post-onset of symptoms. Donor 3 was collected on day 11. None of the donors  
533 were hospitalized or experienced severe disease. PBMCs and plasma were both collected by  
534 density gradient centrifugation using Histopaque-1077 (Sigma-Aldrich).

535 **Preparation of serum antibodies for Ig-Seq proteomic analysis.** Serum samples were  
536 prepared for Ig-Seq analysis as previously described<sup>34</sup>. Briefly, total IgG was isolated from plasma  
537 using Pierce Protein G Plus Agarose (Pierce Thermo Fisher Scientific) and cleaved into F(ab')<sub>2</sub>  
538 fragments with IdeS protease. Antigen-specific F(ab')<sub>2</sub> was enriched by affinity chromatography  
539 against recombinant SARS-CoV-2 S-2P or RBD protein cross-linked to NHS-activated agarose  
540 resin (Thermo Fisher Scientific). Eluted F(ab')<sub>2</sub> fractions were concentrated by vacuum

541 centrifugation and prepared for mass spectrometry-based proteomic analysis as previously  
542 described.<sup>24</sup>

543 **LC-MS/MS analysis of antigen-enriched antibodies.** Liquid chromatography-tandem mass  
544 spectrometry analysis was carried out on a Dionex Ultimate 3000 RSLCnano system coupled to  
545 an Orbitrap Fusion Lumos Mass Spectrometer (Thermo Scientific). Samples were loaded onto an  
546 Acclaim PepMap 100 trap column (75  $\mu\text{m}$   $\times$  2 cm; Thermo Scientific) and separated on an Acclaim  
547 PepMap RSLC C18 column (75  $\mu\text{m}$   $\times$  25 cm; Thermo Scientific) with a 3%-40% acetonitrile  
548 gradient over 60 min at a flow-rate of 300 nl/min. Peptides were eluted directly into the Lumos  
549 mass spectrometer using a nano-electrospray source. Mass spectra were acquired in data-  
550 dependent mode with a 3 sec. cycle time. Full (MS1) scans were collected by FTMS at 120,000  
551 resolution (375-1600 m/z, AGC target = 5E5). Parent ions with positive charge state of 2-6 and  
552 minimum intensity of 3.4E4 were isolated by quadrupole (1 m/z isolation window) and fragmented  
553 by HCD (stepped collision energy = 30+/-3%). Fragmentation (MS2) scans collected by ITMS  
554 (rapid scan rate, AGC target = 1E4). Selected ions and related isotopes were dynamically  
555 excluded for 20 sec (mass tolerance = +/-10ppm).

556 **Antibody variable chain sequencing.** Peripheral blood mononuclear cells (PBMCs) from  
557 processed donor samples were provided in Trizol as a kind gift from Dr. Gregory C. Ippolito<sup>34</sup>. IgG  
558 and IgM VH and VL cDNA libraries were separately amplified from PBMC RNA of three donors  
559 and sequenced to create donor specific reference databases, from which the complete amino  
560 acid sequences of serum IgG proteins could subsequently be determined based on their mass  
561 spectral identifications. **Extended data Figure 1** shows the distribution and diversity of clonotypes  
562 and V-gene usage for the variable heavy chain repertoires of donors 1 and 2.

563 **Ig-Seq MS data analysis.** Mass spectra were analyzed using Proteome Discoverer 2.2 software  
564 (Thermo Scientific). Precursor masses were first recalibrated with the Spectrum File RC node  
565 using a consensus human reference proteome database (UniProt) with common contaminants  
566 (MaxQuant) and precursor mass tolerance of 20 ppm. Recalibrated mass spectra were searched  
567 against a custom database for each donor consisting of donor-derived VH sequences, VL  
568 sequences, and the human and contaminant sequences using the Sequest HT node. Mass  
569 tolerances of 5 ppm (precursor) and 0.6 Da (fragment) were used. Static carbamidomethylation  
570 of cysteine (+57.021 Da) and dynamic oxidation of methionine (+15.995 Da) were considered.  
571 False discovery rates for peptide-spectrum matches (PSMs) were estimated by decoy-based  
572 error modelling through the Percolator node. Label-free quantitation (LFQ) abundances were  
573 calculated from precursor areas using the Minora Feature Detector and Precursor Ions Quantifier  
574 nodes.

575 Resulting PSMs were filtered according to methods described (Boutz, 2014). Briefly, peptide  
576 sequences differing only by isoleucine/leucine substitution were considered equivalent and  
577 combined into a single PSM. PSMs were re-ranked by posterior error probability, q-value, and  
578 Xscore. Only top-ranked, high-confidence PSMs (FDR < 1%) were retained for each scan. If two  
579 or more PSMs had identical top-ranked scores, they were considered ambiguous and removed.  
580 PSMs for the same peptide sequence were summed and the average mass deviation (AMD) was  
581 calculated for each peptide. Peptides with AMD greater than 2 ppm were filtered out. Peptides  
582 mapping to VH sequences from a single clono-group were considered clono-specific. Clono-  
583 specific peptides overlapping the CDR3 sequence by four amino acids or more were considered  
584 CDR3-informative.



585 For each clono-group, PSMs and LFQ abundances of clono-specific CDR3-informative peptides  
586 were summed. Ratios of elution:flow-through PSMs and LFQ abundances were calculated; only  
587 clono-groups with both ratios > 5 were considered elution-specific.

588 **Library assembly and bacterial transformation.** Donor B-cell VH and VL amplicons were  
589 amplified via PCR to include adapters for cloning into yeast expression vectors. Assembly into  
590 the yeast kappa and lambda expression vectors was done via Golden Gate assembly. Library  
591 assemblies were prepared in 20  $\mu$ L reactions as follows: 2  $\mu$ L 10X AARI buffer (ThermoFisher  
592 Scientific, cat. no. B27), 0.4  $\mu$ L 50X oligo buffer (ThermoFisher Scientific, cat. no. ER1582), 0.2  
593  $\mu$ L 100 mM ATP (ThermoFisher Scientific, cat. no. R0441), 20 fmol backbone DNA, 40 fmol VH  
594 and VL amplicons, 0.5  $\mu$ L (2 U/ $\mu$ L) AARI endonuclease (ThermoFisher Scientific, cat. no.  
595 ER1582), and 0.5  $\mu$ L T7 ligase (3000 U/ $\mu$ L) (NEB, cat. no. M0318). Each assembly was scaled  
596 up to 16 total reactions in 8-well strips. Thermocycling consisted of the following protocol: 37°C,  
597 15 minutes; 37°C, 2 minutes, 16°C, 1 minute; go to step 2, x74; 37°C, 60 minutes; 80°C, 15  
598 minutes; hold at 4°C. Assemblies were consolidated and column purified using Promega Binding  
599 Solution (Promega, cat. no. A9303) to bind DNA to a Zymo-spin II column (Zymo Research, cat.  
600 no. C1008). The column was washed twice with DNA Wash Buffer (Zymo Research, cat. no.  
601 D4003) and eluted in 30  $\mu$ L nuclease-free water. For library transformations, DH10B cells were  
602 diluted 1:100 from confluent culture into 50 mL Superior broth (AthenaES, cat. no. 0105). When  
603 cells reached an OD<sub>600</sub> of 0.4-0.6, they were washed 3X with cold 10% glycerol and resuspended  
604 to a final volume of 600  $\mu$ L. The purified library was added to cells and electroporated at 1.8 kV  
605 in an E. coli Pulser electroporator (Bio-Rad) using Genepulser 0.2 cm cuvettes (Bio-Rad, cat. no.  
606 1652086) at 200  $\mu$ L per transformation.

607 **Library transformation into yeast and protein expression.** Purified libraries were linearized  
608 for integration into the yeast genome via homologous recombination at the Leu2 locus. For each

609 1 µg library plasmid, 0.5 µL NotI (10 units/µL) (NEB, cat. no. R0189) was used with the supplied  
610 Buffer 3.1 in 10 µL. Reactions were incubated at 37°C overnight and heat inactivated at 80°C for  
611 20 minutes. Digests were pooled and column purified as described in previous sections and  
612 eluted in 25 µL nuclease-free water. Our strain was electroporated as described elsewhere<sup>1</sup>. We  
613 found that 10 µg of linearized DNA was sufficient for library sizes of 10<sup>6</sup>, and that library sizes  
614 could reach >10<sup>7</sup> with 20 µg DNA. Transformed yeast were recovered in SD -Leu medium (see  
615 “strains and media” section). Libraries were passaged once at 1:100 before protein expression to  
616 reduce contamination from untransformed cells. To express Fab libraries, yeast were washed in  
617 YEP-galactose (see “strains and media”) and diluted 1:10 into 10 mL final volume. Cells were  
618 induced for 48 hours at 20°C with shaking.

619 **Fab library labeling and selection.** Expressed yeast libraries were harvested at 100 µL  
620 (representing approximately 10<sup>7</sup> cells) and washed with PBSA buffer (1X PBS, 2 mM EDTA, 0.1%  
621 Tween-20, 1% BSA, pH 7.4). Antigen was incubated with cells in 1 mL in PBSA at 200 nM at RT  
622 for one hour, washed with PBSA at 4°C, and labeled with secondary antibodies (mouse anti-human  
623 FITC, 1:100; streptavidin-AF647, 1:100; mouse anti-human Fc-AF647, 1:50). Cells were washed  
624 2X and resuspended in 2 mL in cold PBSA for sorting. Cell sorting was performed using a Sony  
625 SH800 fluorescent cell sorter. For first round libraries, 10<sup>7</sup> events were sorted into 2 mL SD-Leu  
626 medium supplemented with penicillin/streptomycin (Gibco, cat. no. 15140122). Cells were  
627 recovered by shaking incubation for 1-2 days for further rounds of selection or plated directly for  
628 phenotyping clones.

629 **Next generation sequencing.** Genome extraction was performed on yeast cultures of libraries  
630 and sorted rounds underwent genome extraction using a commercial kit (Promega, cat. no.  
631 A1120) with zymolyase (Zymo Research, cat. no. E1004). 100 ng genomic template was used to  
632 amplify the heavy and light chains separately or as one amplicon for short or long-read

633 sequencing, respectively. For amplification of heavy chain genes only, primers JG.VHVLK.F and  
634 JG.VH.R were used. For amplification of light chain genes only, primers JG.VL.F and JG.VHVK.R  
635 or JG.VHVL.R were used for kappa and lambda vectors, respectively. For amplification of paired  
636 genes, primers JG.VHVLK.F and JG.VHVK.R or JG.VHVL.R were used. Amplicons were column  
637 purified and deep sequenced with an iSeq. In parallel, we obtained ~1.8 kb sequences spanning  
638 the entire VH and VL using MinION nanopore sequencing (Oxford Nanopore Technologies Ltd.,  
639 MinION R10.3).

640 **Colony PCR and Sanger sequencing.** Sorted yeast populations were plated on SD -Leu and 8-  
641 32 colonies per plate were picked into 2 mL microplates either by hand or using a QPIX 420  
642 (Molecular Devices) automatic colony picker. Cultures were grown at 1000 rpm at 3 mm orbit at  
643 30°C overnight. Cells (20  $\mu$ L) were transferred to a fresh microplate and washed with 1 mL TE  
644 buffer (10 mM Tris, 1 mM EDTA). Cells were incubated with 20  $\mu$ L zymolyase solution (5 mg/mL  
645 zymolyase, 100T in TE) at 37°C for 1 hour. Cells (5  $\mu$ L) were then used in colony PCR to amplify  
646 the paired heavy and light chains. Amplicons were column purified with the Wizard SV 96 PCR  
647 Clean-Up System (Promega, cat. no. A9342) and yields were quantified with a Nanodrop  
648 spectrophotometer or the Quant-it Broad-Range dsDNA kit (Invitrogen, cat. no. Q33130).  
649 Approximately 10 ng (2.5-5  $\mu$ L) of purified PCR products were then subjected to Sanger  
650 sequencing.

651 **Long-read sequencing (donors 1 & 2)** Sequencing libraries were prepared from 18 amplicon  
652 samples using the Native Barcoding Kit (Oxford Nanopore Technologies; cat. no. EXP-NBD103)  
653 paired with the Ligation Sequencing Kit (Oxford Nanopore Technologies; cat. no. SQK-LSK109)  
654 according to the manufacturer's directions. Between four and eight sequencing libraries per flow  
655 cell were pooled for sequencing on three MinION flow cells (Oxford Nanopore Technologies;  
656 R9.4.1) for 72 hours on an Oxford Nanopore Technologies MinION Mk1B device (Oxford

657 Nanopore Technologies). Raw data was basecalled using the high accuracy model in Guppy  
658 (v.3.5.2).

659

660 **Short-read sequencing (phenotyping plates)** Sequencing libraries were prepared from 308  
661 amplicon samples using the Nextera DNA Flex Library Preparation kit (Illumina; cat. no.  
662 20018705) according to the manufacturer's directions. Sequencing libraries were pooled and  
663 sequenced (2x151bp) on an iSeq 100 (Illumina; California, USA) using iSeq 100 i1 Reagents v.1  
664 (Illumina; cat. no. 20021533).

665

666 **Long-read sequencing (donor 3)** Sequencing libraries were prepared from 32 amplicon  
667 samples using the Native Barcoding Kit (Oxford Nanopore Technologies; cat. no. EXP-NBD104)  
668 paired with the Ligation Sequencing Kit (Oxford Nanopore Technologies; cat. no. SQK-LSK109)  
669 according to the manufacturer's directions. Between five and eight sequencing libraries per flow  
670 cell were pooled for sequencing on five GridION flow cells (Oxford Nanopore Technologies;  
671 R9.4.1) for 72 hours on a GridION Mk1 device (Oxford Nanopore Technologies; Oxford, England,  
672 UK). Raw data was live basecalled using the high accuracy model in Guppy (v.3.2.10).

673

674 **Short-read sequencing (donor 3)** Sequencing libraries were prepared from 32 samples using  
675 the Nextera DNA Flex Library Preparation kit (Illumina; cat. no. 20018705) according to the  
676 manufacturer's directions. Sequencing libraries were pooled and sequenced (2x151bp) on an  
677 iSeq 100 (Illumina; California, USA) using the iSeq 100 i1 Reagents v2 (Illumina; cat. no.  
678 2009584).

679

680 **Long-read sequencing (YSD-IgSeq)** Sequencing libraries were prepared from 16 amplicon  
681 samples using the Native Barcoding Kit (Oxford Nanopore Technologies; Cat. No. EXP-NBD104)

682 paired with the Ligation Sequencing Kit (Oxford Nanopore Technologies; Cat. No. SQK-LSK109)  
683 according to the manufacturer's directions. Four sequencing libraries were pooled per flow cell  
684 and sequenced on four GridION flow cells (Oxford Nanopore Technologies; R9.4.1) for 72 hours  
685 on a GridION Mk1 device (Oxford Nanopore Technologies; Oxford, England, UK). Raw data was  
686 live basecalled using the high accuracy model in Guppy (v.4.0.11).

687

688 **Short-read sequencing (YSD-IgSeq)** Sequencing libraries were prepared from 16 samples  
689 using the Nextera DNA Flex Library Preparation kit (Illumina; cat. no. 20018705) according to the  
690 manufacturer's directions. Sequencing libraries were pooled and sequenced (2x150 bp) on an  
691 iSeq 100 (Illumina; California, USA) using the iSeq 100 i1 Reagents v2 (Illumina; Cat. No.  
692 2009584).

693 **Sequence processing and consolidation into VHVL clones** Individual reads after Guppy base  
694 calling typically average more than 10% error per base, and numerous tools exist to align and  
695 reduce reads into consensus sequences with substantially improved accuracy<sup>36</sup>. However, such  
696 tools were not designed for antibody library sequencing and its huge populations of subtly different  
697 sequences which, even assuming successful alignment, group into a myriad of very short,  
698 disconnected assemblies. We therefore implemented a bioinformatic pipeline to obtain accurate  
699 VHVL sequences from the MinION and iSeq data and to estimate, within each YSD round, the  
700 relative abundance of individual VHVL pairs.

701 Our methods proceed through antibody V(D)J annotation of raw MinION reads using MiXCR  
702 (v3.0.13)<sup>37</sup>; iteratively growing and shrinking sequence clusters based on annotated features from  
703 each read; sequence error correction and consolidation within each cluster, optionally including  
704 high quality Illumina reads; and finally, VHVL clone definition within each sample and quantitation  
705 by number of reads mapped to each clone. For enumeration, we only include counts for reads

706 with a length of 1700 to 2100 base pairs. This points to a secondary advantage of MinION  
707 sequencing, as shorter reads proliferate during PCR and inflate apparent abundance of particular  
708 species. Without length-filtering, relative VHVL abundance calculated from both MinION and iSeq  
709 reads are strongly correlated.

710 **Tissue culture and transient transfection of ECD and RBD.** Spike ECD protein and RBD  
711 proteins were expressed in Expi293F cells using the manufacturer provided guidelines with slight  
712 modifications. In short, a 1 ml frozen working cell bank of Expi293F cells at  $1 \times 10^7$  viable cells/mL  
713 were thawed in a bead bath at 37°C for 2-3 mins. The vial was sprayed with 70% isopropyl alcohol  
714 and transferred into a biosafety cabinet. The thawed cells were transferred to a 125 mL non-  
715 baffled vented shake flask containing 29 mL of fresh pre-warmed ExpiExpression medium at  
716 37°C. Cells were incubated in at 37°C with  $\geq 80\%$  humidity and 8% CO<sub>2</sub> on an orbital shaker at  
717 120 rpm and grown until they reached a cell density of  $3 \times 10^6$  viable cells/ml. Fresh pre-warmed  
718 ExpiExpression medium was added to 1 L non-baffled vented shake flask and the 30 mL cell  
719 suspension was carefully introduced, making the seeding density of  $0.4 \times 10^6$  viable cells/mL in a  
720 final culture volume of 225 mL. After the cell density reached  $3 \times 10^6$  viable cells/ml, the culture  
721 was expanded to a final volume of 2.25 L in two 2.8 L Thomson Optimum Growth flasks with a  
722 seeding density of  $0.3 \times 10^6$  viable cells/mL. After the cell density reached  $3 \times 10^6$  viable cells/mL,  
723 2 L of the culture was-split into four Thompson flasks with each flask containing 500 mL of culture  
724 medium and 500 mL of fresh pre-warmed ExpiExpression medium. The final culture volume in  
725 each of the four flasks was 1 L. The cells were incubated at 37°C with  $\geq 80\%$  humidity and 8%

726 CO<sub>2</sub> on an orbital shaker at 100 rpm. When the cells reached a density of  $3 \times 10^6$  viable cells/mL,  
727 the culture was transferred to two 500 mL sterile centrifuge bottles and the cells were spun down  
728 at 100 x g for 10 min. The supernatant was removed, and the cells were resuspended in 4 L of  
729 fresh, prewarmed medium. The cells were allowed to re-stabilize in the incubator for 24 h and  
730 were transfected at  $3 \times 10^6$  viable cells/mL in 4 L.

731 Transfection was performed by diluting 4 mg of plasmid DNA (pDNA) in 240 mL of OptiMEM  
732 medium in a sterile bottle and gently inverting 3 - 4 times before incubating at room temperature  
733 (RT) for 5 min. ExpiFectamine 293 reagent (13 mL) was then diluted in 225 mL OptiMEM in a  
734 sterile bottle and inverted 3 - 4 times before incubating for 5 min at RT. The diluted pDNA and  
735 ExpiFectamine reagent were carefully mixed and incubated at RT for 15 min. One-fourth of the  
736 combined complex was then slowly transferred to each of the flasks while gently swirling the cells  
737 during addition. The cultures were again placed in the incubator with shaking at 100 rpm for 18 h.  
738 Post-transfection, ExpiFectamine 293 Transfection Enhancer 1 (6 mL) and ExpiFectamine 293  
739 Transfection Enhancer 2 (60 mL) were added to each flask. The cell viability was monitored every  
740 24 h and the cells were harvested when the viability dropped below 70% or after approximately 3  
741 d. Harvesting was done by centrifugation at 15,900 x g for 45 min at 4°C and the supernatant was  
742 transferred into sterile bottles.

743 **Antigen purification.** The feed was prepared adding 0.2 M NaCl and 10 mM imidazole to the  
744 supernatant while mixing. The feed material was filtered using a 0.45 µm PES filter membrane  
745 pre-wetted with PBS before loading on a prepared Ni-IMAC column. A metal affinity column was  
746 prepared by packing IMAC FF beads (Cytiva) into an AxiChrom 70 column housing to a bed height  
747 of 9.5 cm and then charging with 50% column volume (CV) of a 0.2 M nickel sulfate solution,  
748 washing with water and then 50% CV of 100% "B" buffer (50 mM sodium phosphate buffer

749 containing 300 mM NaCl and 250 mM imidazole, pH 7.8) to remove weakly bound nickel ions.  
750 The column was then washed with 100% “A” buffer (50 mM sodium phosphate buffer containing  
751 300 mM NaCl and 20 mM imidazole, pH 7.2) prior to sample loading. The prepared (4.8 L) feed  
752 was then loaded onto the column at a linear flow rate of 90 cm/h. After loading, the column was  
753 washed with two CVs of 100% A buffer and two CVs of a 13% B buffer (containing 50 mM  
754 imidazole) before eluting the protein using 100% B (250 mM imidazole) for 3 CVs. All steps except  
755 for the loading step were done at 150 cm/h.

756 Fractions from the elution peak were pooled and then concentrated 4- to 8-fold by ultrafiltration  
757 (UF) using a 115 cm<sup>2</sup> hollow fiber cartridge with either a 50 kDa (S-2P) or a 10 kDa (RBD)  
758 molecular weight cutoff membrane (Repligen) and then diafiltered after concentration by  
759 exchanging with 5 volumes of PBS.

760 The concentration of diafiltered protein was determined by measuring the absorbance at 280 nm  
761 versus a PBS blank. The protein concentration in mg/mL was obtained using a divisor of 1.03  
762 mL/mg-cm for S-2P and 1.19 mL/mg-cm for RBD. A qualitative assessment of protein quality was  
763 made using SDS-PAGE with SYPRO Ruby staining (BioRad) for reduced and non-reduced  
764 samples. Only those preparations showing predominantly full-length S-2P (160 kDa subunits) or  
765 RBD (70 kDa) were used in ELISAs for assessing neutralizing antibodies.

766 **Monoclonal antibody expression and purification.** VHVL candidates were cloned into custom  
767 Golden Gate compatible pCDNA3.4 vectors for IgG1 expression. For transfections, VL was mixed  
768 3:1 with a corresponding VH. Plasmids were transfected into Expi293F (Invitrogen) cells using  
769 the recommended protocol. Monoclonal antibodies were harvested at 5-7 days post-transfection.  
770 Expi293F cells were centrifuged at 300 x g for 5 min, supernatants were collected and centrifuged  
771 at 3000 x g for 20 min at 4°C and diluted to 1X PBS final concentration. Each supernatant was



772 passed through a Protein G or A agarose affinity column (Thermo Scientific). Flow through was  
773 collected and passed through the column three times. Columns were washed with 10 CV of PBS  
774 and antibodies were eluted with 5 mL 100 mM glycine, pH 2.7 directly in neutralization buffer  
775 containing 500  $\mu$ L 1 M Tris-HCl, pH 8.0.

776 **Phenotyping assays.** Sorted clones from rounds two or three of selection were picked into  
777 microplates as described previously. After antibody expression, 10  $\mu$ L of cells were transferred to  
778 a fresh 2-mL microplate and washed 2X with 200  $\mu$ L cold PBSA buffer. Cells were labeled with  
779 200 nM spike or RBD antigen in 50-100  $\mu$ L PBSA at RT for 1 h with shaking at 1000 rpm, 3 mm  
780 orbit. Labeled cells were washed 2X, and secondary labels were applied as previously described.  
781 Cells were resuspended in 200  $\mu$ L ice-cold PBSA just before analysis. Samples were analyzed  
782 on a Sony SA3800 Spectral Cell Analyzer.

783 **Analysis of mammalian cell surface displayed spike proteins with flow cytometry.** The  
784 assay has been described in detail in the recent paper by Javanmardi et al<sup>38</sup>. Briefly, plasmids  
785 expressing full-length SARS-CoV-2 spike proteins (Spike-Linker-3XFLAG-TM), WT (HexaPro-  
786 D614G) and variants, were transfected into HEK293T cells using Lipofectamine 2000 according  
787 to manufacturer's instructions. After 48 h, cells were collected and resuspended in PBS-BSA and  
788 incubated with anti-FLAG (mouse) and anti-spike (human) mAbs for 1 h, shaking at RT. Cells  
789 were washed 3X with PBS-BSA, resuspended in \_\_\_ mL and incubated with Alexa Fluor 488 (anti-  
790 mouse) and Alexa Fluor 647 (anti-human) antibodies for 30 min, shaking at 4°C. Cells were  
791 washed again and resuspended in PBS-BSA prior to flow cytometry analysis (SA3900 Spectral  
792 Analyzer, Sony Biotechnology). All data was analyzed with FlowJo (BD Bioscience).

793 **ELISA.** Antigen ELISA plates were made using high-binding plates (Corning, cat. no. 3366) with  
794 antigen diluted in PBS to a final concentration of 2  $\mu$ g/mL. Antigen solution (50  $\mu$ L) was added to

795 microplates and incubated overnight at 4°C with shaking at 100 rpm, 3 mm orbit. Plates were  
796 blocked with PBSM (2% milk in PBS) at RT for 1 h. Plates were washed 3X with 300 µL PBS-T  
797 (0.1% Tween-20). Purified antibodies were prepared to 10 µg/mL in PBSM and serially diluted.  
798 Antibodies were incubated for 1 h at RT. Plates were washed 3X with PBS-T, and secondary goat  
799 anti-human Fab-HRP (Sigma-Aldrich, cat. no. A0293) was applied at 1:5000 in PBSM in 50 µL  
800 and incubated at RT for 45 min. HRP substrate (50 µL) was added to wells and the reaction  
801 proceeded for 5-15 min until quenched with 50 µL 4M H<sub>2</sub>SO<sub>4</sub> and analyzed for absorbance at 450  
802 nm in a plate reader.

803 **Live virus neutralization assays.** A SARS-CoV-2 microneutralization assay was adapted from  
804 an assay used to study Ebola virus<sup>39</sup>. This assay also used SARS-CoV-2 strain WA1. Antibodies  
805 were diluted in cell culture medium in triplicate. A SARS-CoV-2 monoclonal antibody was used  
806 as a positive control. An antibody that does not bind SARS-CoV-2A was used as negative control.  
807 Diluted antibodies were mixed with the SARS-CoV-2 WA1 strain, incubated at 37°C for 1 h, then  
808 added to Vero-E6 cells at target MOI of 0.4. Unbound virus was removed after 1 h incubation at  
809 37°C, and culture medium was added. Cells were fixed 24 h post-infection, and the number of  
810 infected cells was determined using SARS-CoV-S specific mAb (Sino Biological, cat. no. 401430-  
811 R001) and fluorescently-labeled secondary antibody. The percent of infected cells was  
812 determined with an Operetta high-content imaging system (PerkinElmer) and Harmonia software.  
813 Percent neutralization for each monoclonal antibody at each dilution was determined relative to  
814 untreated, virus-only control wells.

815 Live virus neutralization (VN) for variant of concern B.1.1.7 and WA/1 were performed in an  
816 orthogonal assay. The ability of the monoclonal antibodies to neutralize SARS-CoV-2 was  
817 determined with a traditional VN assay using SARS-CoV-2 strain USA-WA1/2020 (NR-52281-BEI  
818 resources), as previously described<sup>40,41,42,43</sup>. All experiments with SARS-CoV-2 were performed

819 in the Eva J Pell BSL-3 laboratory at Penn State and were approved by the Penn State Institutional  
820 Biosafety Committee (IBC # 48625). For each mAb a series of 12 two-fold serial dilutions were  
821 assessed from a stock concentration of 1 mg/ml. Triplicate wells were used for each antibody  
822 dilution. 100 tissue culture infective dose 50 (TCID<sub>50</sub>) units of SARS-CoV-2 were added to 2-fold  
823 dilutions of the diluted mAb. After incubating for 1 hour at 37°C, the virus and mAb mixture was  
824 then added to Vero E6 cells (ATCC CRL-1586) in a 96-well microtiter plate and incubated at 37°C.  
825 After 3 days, the cells were stained for 1 hour with crystal violet–formaldehyde stain (0.013%  
826 crystal violet, 2.5% ethanol, and 10% formaldehyde in 0.01 M PBS). The endpoint of the  
827 microneutralization assay was determined as the highest mAb dilution, at which all 3, or 2 of 3,  
828 wells are not protected from virus infection. Percent neutralization ability of each dilution of the  
829 mAb was calculated based on the number of wells protected, 3, 2, 1, 0 of 3 wells protected was  
830 expressed as 100%, 66.6%, 33.3%, or 0%.

831 **Surface plasmon resonance.** To investigate the binding kinetics of mAb N3-1 binding to the  
832 spikes, purified His-tagged spike variants (SARS-CoV Tor2 S-2P, SARS-CoV-2 Wuhan-Hu-1 S-  
833 HexaPro, SARS-CoV-2 B.1.1.7 S-Hexapro and SARS-CoV-2 B.1.351 S-HexaPro) were  
834 immobilized on a Ni-NTA sensor chip (GE Healthcare) using a Biacore X100 (GE Healthcare).  
835 For Fab binding experiments, we immobilized spike proteins to a level of ~450 response units  
836 (RUs). Serial dilutions of purified Fab N3-1 were injected at concentrations ranging from 400 to  
837 6.25 nM over spike-immobilized flow cell and the control flow cell in a running buffer composed  
838 of 10 mM HEPES pH 8.0, 150 mM NaCl and 0.05% Tween 20 (HBS-T). Between each cycle, the  
839 sensor chip was regenerated with 0.35 M EDTA, 50 mM NaOH and followed by 0.5 mM NiCl<sub>2</sub>.  
840 For IgG binding experiments, spike immobilization of 200 RUs was used instead to avoid mass  
841 transport effect. Serial dilutions of purified IgG N3-1 were injected at concentrations ranging from  
842 25 to 1.56 nM over a spike-immobilized flow cell and the control flow cell. For the SARS-CoV Tor2

843 S-2P binding experiments, IgG N3-1 concentrations ranging from 100 to 6.25 nM were used.  
844 Response curves were double-reference subtracted and fit to a 1:1 binding model or  
845 heterogeneous ligand binding model using Biacore X100 Evaluation Software (GE Healthcare).

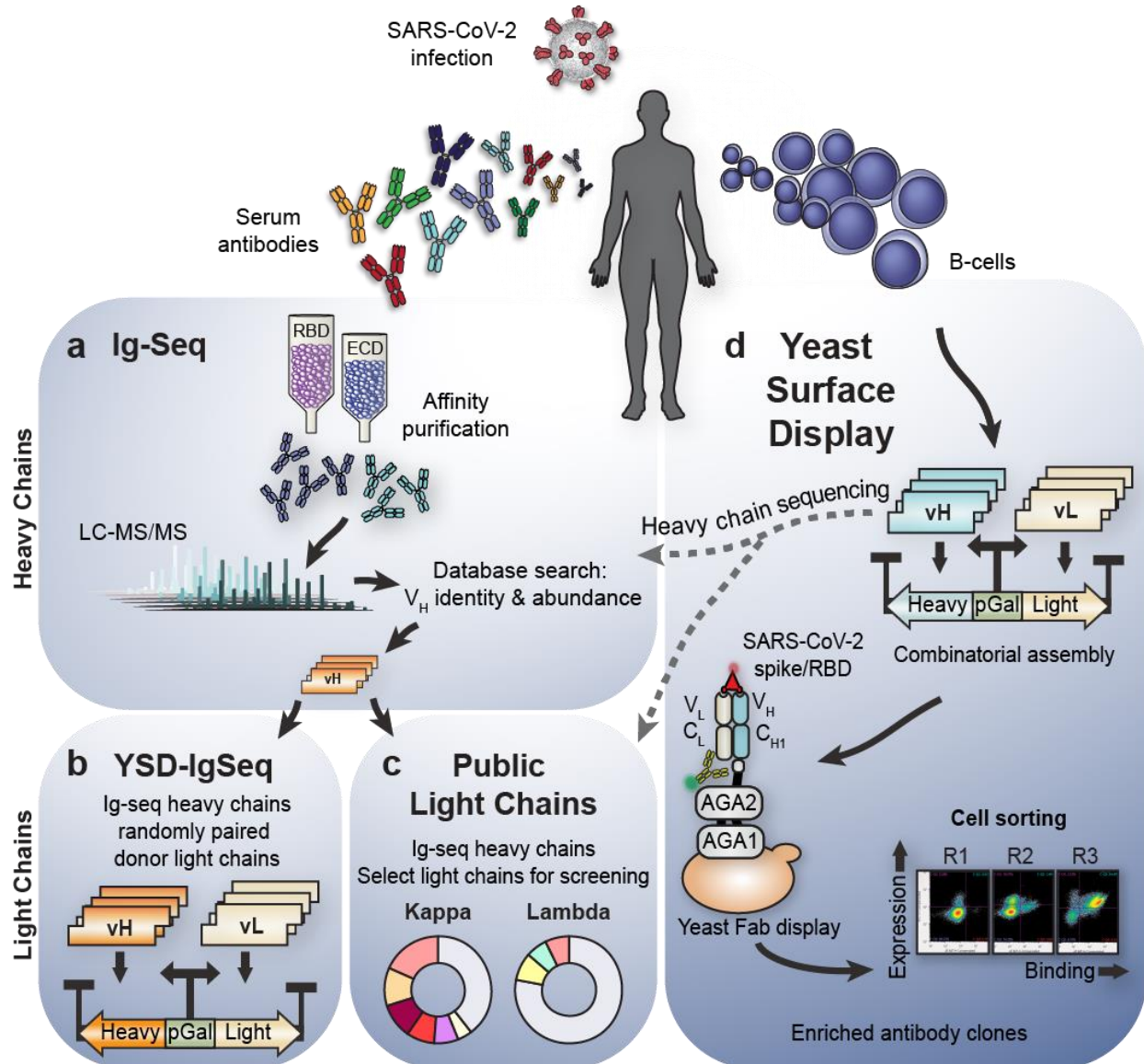
846 **Negative stain EM for spike-IgG complexes.** To investigate mAb N3-1 binding to spike proteins,  
847 purified SARS-CoV-2 Wuhan-Hu-1 S-HexaPro was incubated with 1.2-fold molar excess of IgG  
848 N3-1 in 2 mM Tris pH 8.0, 200 mM NaCl and 0.02% NaN<sub>3</sub> on ice for 10 min. The spike-IgG  
849 complexes were at a concentration of 0.05 mg/mL in 2 mM Tris pH 8.0, 200 mM NaCl and 0.02%  
850 NaN<sub>3</sub> prior to deposition on a CF-400-CU grid (Electron Microscopy Sciences) that was plasma  
851 cleaned for 30 sec in a Solarus 950 plasma cleaner (Gatan) with a 4:1 ratio of O<sub>2</sub>/H<sub>2</sub> and stained  
852 using methylamine tungstate (Nanoprobes). Grids were imaged at a magnification of 92,000X  
853 (corresponding to a calibrated pixel size of 1.63 Å/pix) in a Talos F200C TEM microscope  
854 equipped with a Ceta 16M detector. The CTF-estimation, particle picking and 2D classification  
855 were all performed in cisTEM<sup>44</sup>.

856 **Cryo-EM sample preparation and data collection.** Purified SARS-CoV-2 S (HexaPro variant)  
857 at 0.2 mg/mL was incubated with 5-fold molar excess of Fab N3-1 in 2 mM Tris pH 8.0, 200 mM  
858 NaCl and 0.02% NaN<sub>3</sub> at RT for 30 min. The sample was then deposited on plasma-cleaned  
859 UltrAuFoil 1.2/1.3 grids before being blotted for 4 sec with -3 force in a Vitrobot Mark IV and  
860 plunge-frozen into liquid ethane. Purified SARS-CoV-2 S with three RBDs covalently trapped in  
861 the down conformations (HexaPro-RBD-down variant, S383C/D985C<sup>45-47</sup>) at 0.2 mg/mL,  
862 complexed with 2-fold molar excess of Fab A7V3, was deposited on plasma-cleaned UltrAuFoil  
863 1.2/1.3 grids before being blotted for 3 sec with -4 force in a Vitrobot Mark IV and plunge-frozen  
864 into liquid ethane. For the HexaPro-N3-1 sample, 3,203 micrographs were collected from a single  
865 grid. For the HexaPro-RBD-down-A7V3 sample, 3,636 micrographs were collected from a single  
866 grid. FEI Titan Krios equipped with a K3 direct electron detector (Gatan) was used for imaging.

867 Data were collected at a magnification of 22,500x, corresponding to a calibrated pixel size of 1.07  
868 Å/pix. A full description of the data collection parameters can be found in **Extended Data Tables**  
869 **3-4** .

870 **Cryo-EM data processing.** Gain reference- and motion-corrected micrographs processed by  
871 Warp<sup>48</sup> were imported into cryoSPARC v2.15.0<sup>49</sup>, which was used to perform CTF correction,  
872 micrograph curation, particle picking, and particle curation via iterative rounds of 2D classification.  
873 The final global reconstructions were then obtained via ab initio reconstruction, iterative rounds  
874 of heterogeneous refinement, and subsequently non-uniform homogeneous refinement of final  
875 classes with C1 symmetry. For the HexaPro-RBD-down-A7V3 sample, C3 symmetry was  
876 attempted in the initial refinement process. Given the low occupancy of the Fabs on a trimeric  
877 spike, C1 symmetry was used for the final runs of heterogeneous and homogeneous refinement.  
878 To better resolve the Fab-spike interfaces, both datasets were subjected to particle subtraction  
879 and focused refinement. Finally, both global and focused maps were sharpened using  
880 DeepEMhancer<sup>50</sup>. For A7V3-NTD model building, we used an NTD from the 4A8 complexed spike  
881 structure (PDB ID: 7C2L<sup>25</sup>) and a homologous Fab structure (PDB ID: 6IEK) as an initial model to  
882 build into map density using UCSF ChimeraX<sup>51</sup>. For N3-1-RBDs model building, we used one  
883 RBD-up and one RBD-down from S-HexaPro (PDB ID: 6XKL<sup>52</sup>) and two homologous Fab  
884 structures (PDB ID: 5BV7 and 5ITB) as an initial model to build into map density via UCSF  
885 ChimeraX. Both models were built further and iteratively refined using a combination of Coot<sup>53</sup>,  
886 Phenix<sup>54</sup>, and ISOLDE<sup>55</sup>. The detailed workflows of cryo-EM data processing and data validation  
887 can be found in **Extended Data Figures 10, 11 and 14**.

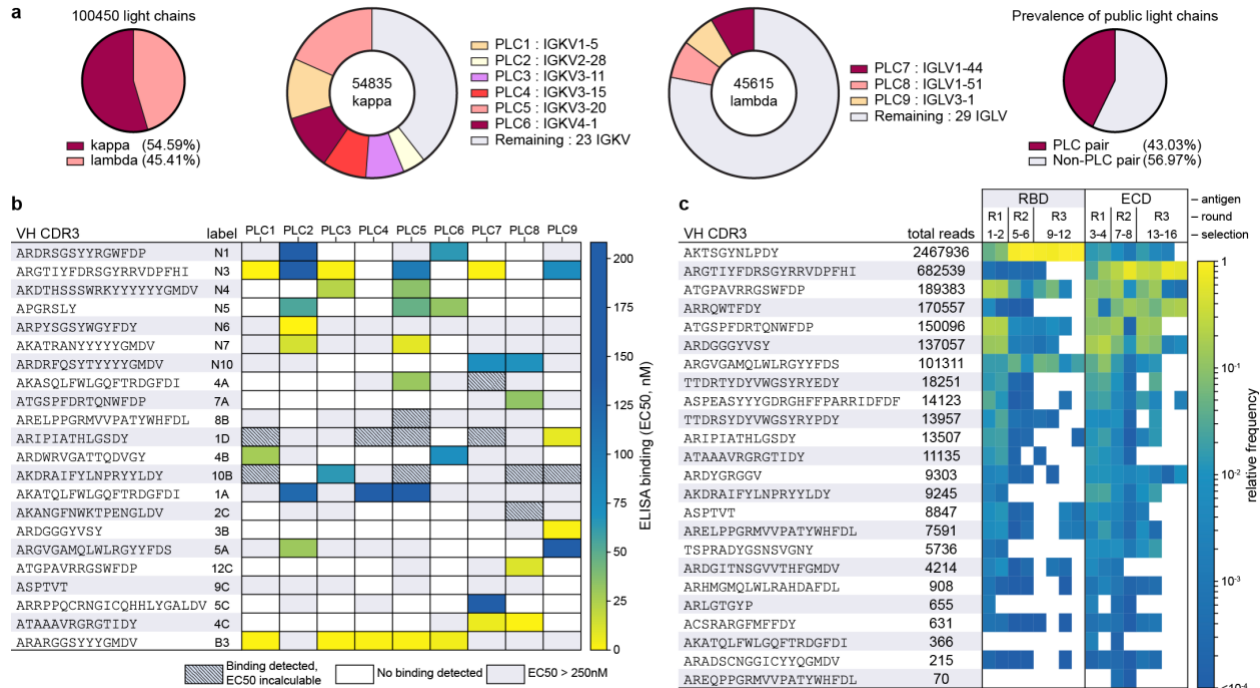
888



889

890 **Figure 1. Overview of complementary strategies for discovering convalescent patient-derived anti-**  
 891 **SARS-CoV-2 neutralizing antibodies.**

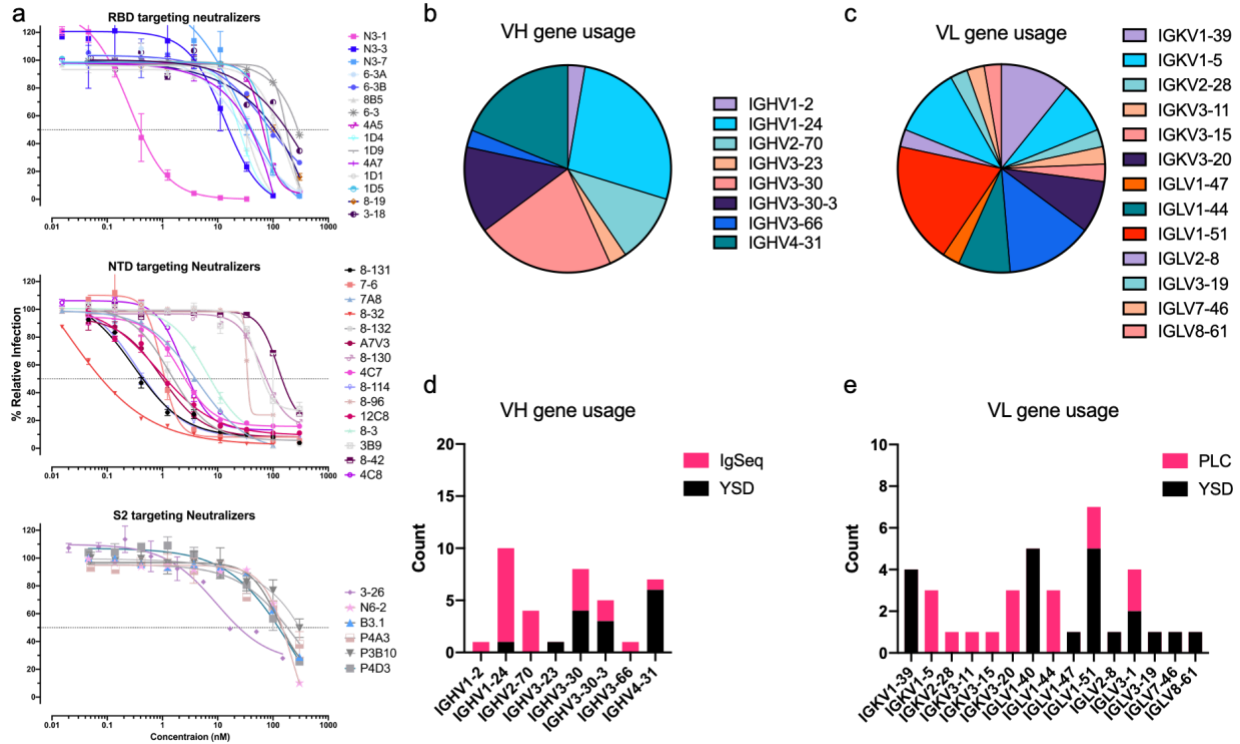
892 Serum antibodies and B-cells are isolated from patient blood post infection. **a**, IgG is antigen-enriched, digested, and  
 893 analyzed with tandem mass spectrometry. Searching against a database of donor B cell receptor sequences yields  
 894 antigen-specific heavy chain candidates. **b**, To find productive VH-VL pairs, the proteomically identified VH are  
 895 randomly combined with donor VL, generating a large library of Fab amplicons for selection in yeast. **c**, Additional  
 896 productive VH-VL are recovered by matching each proteomic VH candidate against a fixed set of nine public light  
 897 chains significantly overrepresented in prior repertoire studies. All combinations are screened for binding. **d**, Yeast  
 898 surface display provides an independent method to recover VH or VHVL candidates. VH and VL sequences are  
 899 amplified from donor B-cells, and pairs are randomly assembled into large Fab libraries and selected using  
 900 fluorescent cell sorting. Yeast pools are then sequenced after each selection to track individual clone dynamics round  
 901 over round. Promising candidates are expressed as IgG for further characterization.



902

903 **Figure 2. Public light chain and YSD-IgSeq abundance and screening.**

904 **a**, Nine PLCs emerge from analysis of 100,450 previously published paired VH-VL sequences. **b**, Screening with VHs  
 905 (H-CDR3 depicted) identified by Ig-Seq and YSD against the panel of nine PLCs to determine productive VHVL  
 906 pairings. IgG mAbs ELISA EC50s revealed that partnering VHs with PLCs can create low nM affinity binders. Grey  
 907 boxes indicate that binding was detected but at an affinity too low to determine an EC50 in the concentration range  
 908 tested. White boxes, no binding was detected. **c**, VH CDR3 read counts and relative abundance from the YSD-IgSeq  
 909 experiment shown for MinION reads of 1.7 kb to 2.1 kb. Heatmap values are CDR3 frequencies, or read counts  
 910 normalized within each respective sample.



911  
 912 **Figure 3. Non-cognate paired antibodies neutralize SARS-CoV-2 live virus.**  
 913 **a**, SARS-CoV-2 WA1 live virus neutralization assays. The graphs depict the neutralizers binned by their target spike  
 914 subdomain. **b**, VH gene usage of the neutralizing antibodies as a fraction of the whole. **c**, Stacked bar chart showing  
 915 neutralizer VH gene usage broken down by discovery method. **d-e**, Same analysis described in b-c but by VL gene.

916

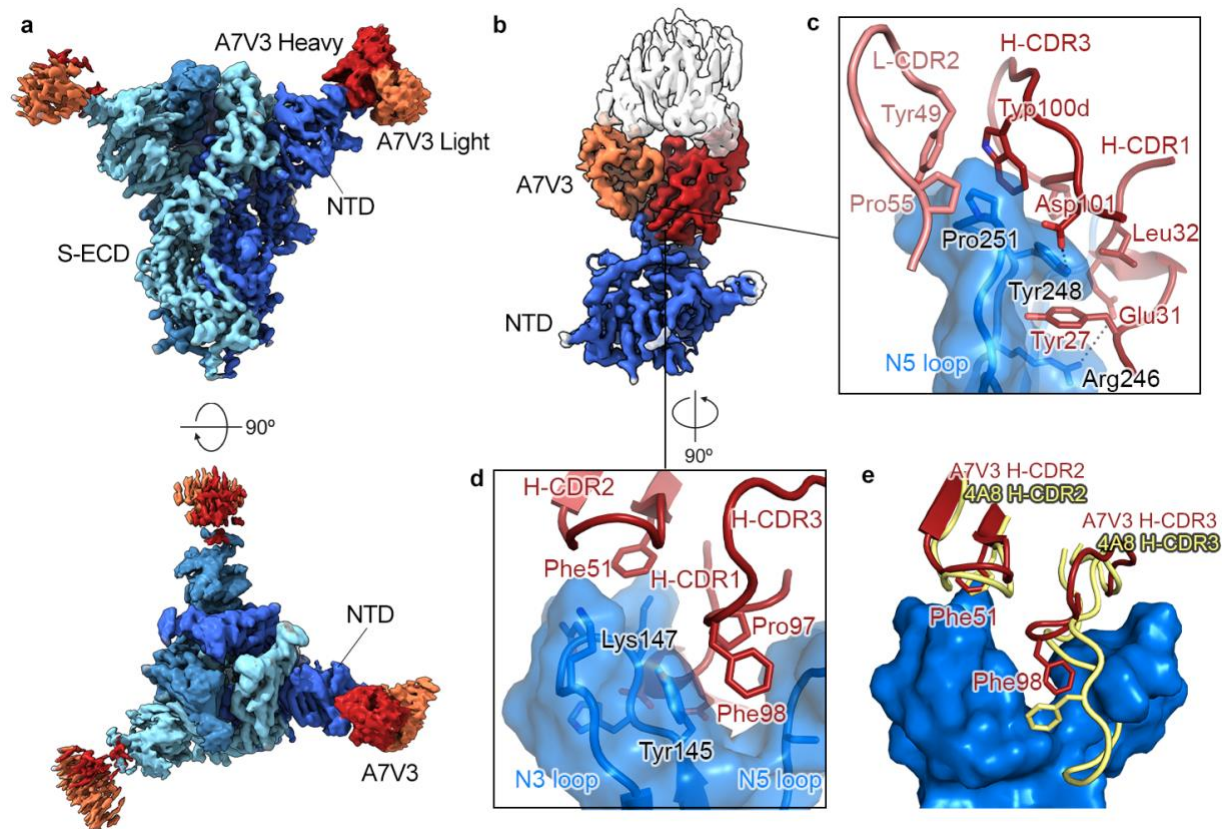
917

918



919

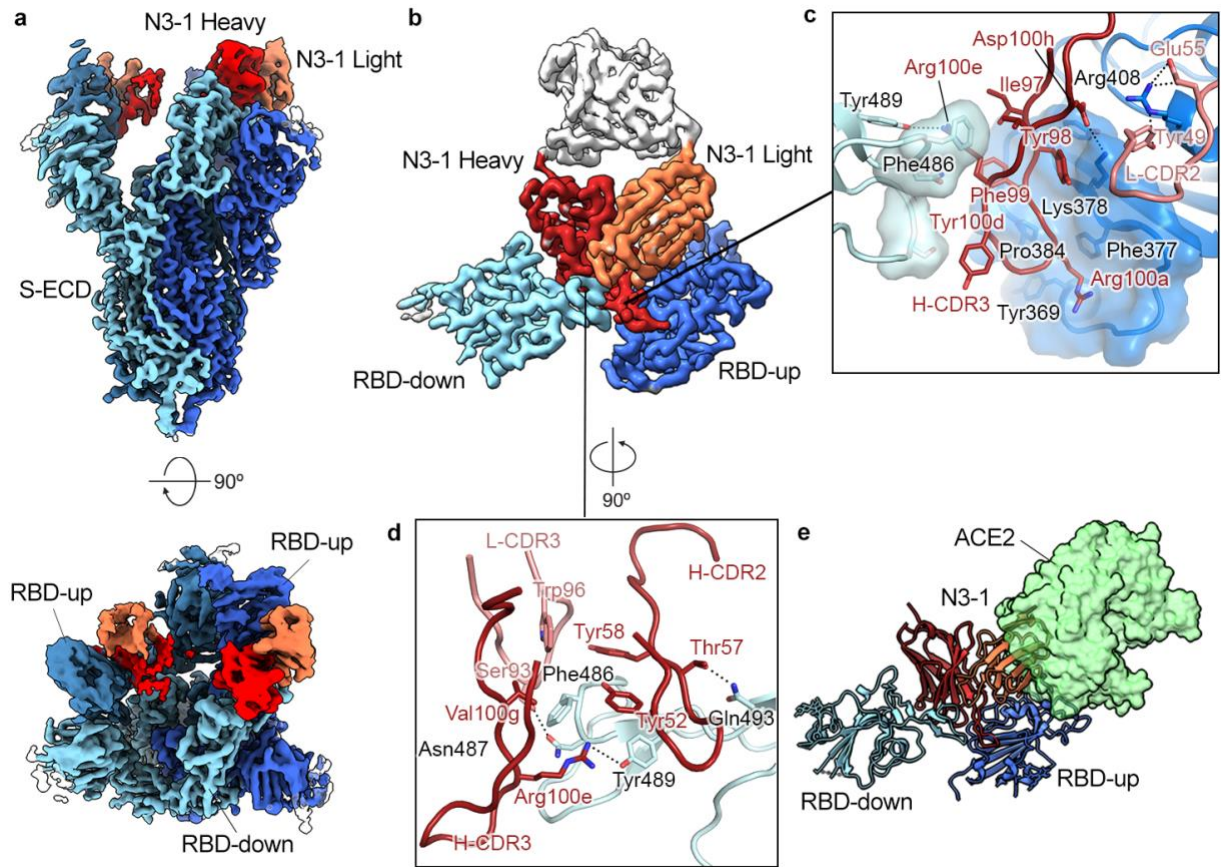
920



921

922 **Figure 4. NTD-directed mAb A7V3 exhibits a common binding mode.**

923 **a**, Cryo-EM structure of A7V3 bound to SARS-CoV-2 S at a global resolution of 3.0 Å. Side view and top-down views  
924 of the complex are shown in the upper and the lower panel, respectively. Each protomer is depicted in steel blue, royal  
925 blue and sky blue. The heavy chain of A7V3 is colored firebrick, and the light chain is colored coral. **b**, Focused map  
926 of A7V3 bound to NTD reveals a major binding site on the N3 and N5 loops. **c**, The tip of the N5 loop is surrounded by  
927 the hydrophobic residues from L-CDR2 and H-CDR3. Arg246 and Tyr248 contact Glu31 and Asp101 via polar  
928 interactions. **d**, Pro97 and Phe98 in H-CDR3 insert into a groove walled by the N3 and N5 loops. Conserved Phe51 in  
929 H-CDR2 forms a pi-cation interaction with Lys147 **e**, Superimposed structure of 4A8-NTD complex (PDB ID: 7C2L)  
930 with A7V3-bound NTD. The molecular surface of the N3 and N5 loops is shown in blue. Unlike 4A8, the relatively short  
931 H-CDR3 from A7V3 barely contacts the N3 loop.  
932  
933



934

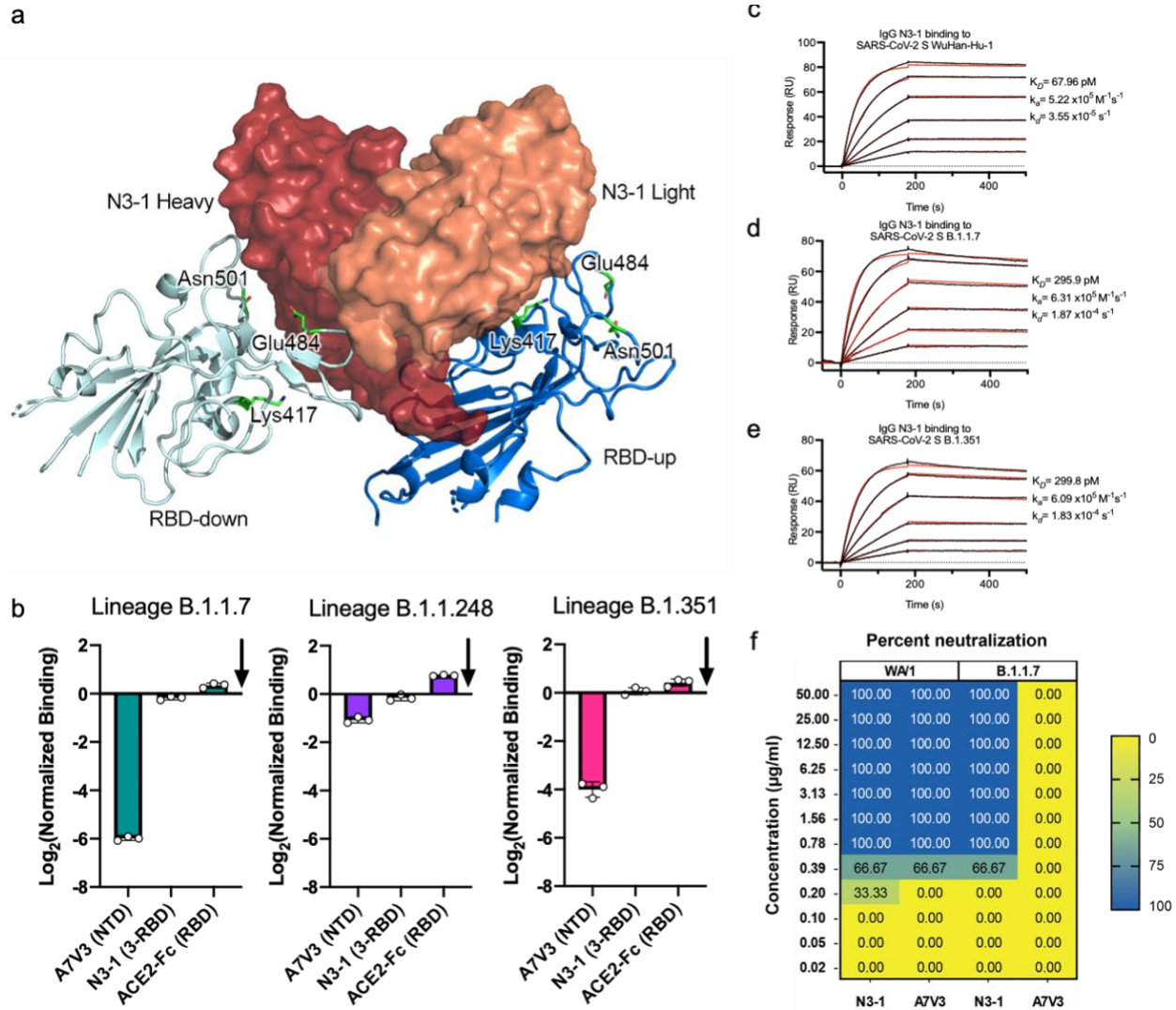
935 **Figure 5. RBD-directed mAb N3-1 exhibits a unique binding mode by recognizing two distinct**  
 936 **epitopes.**

937 **a**, Cryo-EM structure of N3-1-bound SARS-CoV-2 spike at a global resolution of 2.8 Å. Side view and top-down views  
 938 of the complex are shown in the upper and the lower panel, respectively. Each protomer is depicted in steel blue, royal  
 939 blue and sky blue. The heavy chain of N3-1 is colored firebrick, and the light chain is colored coral. **b**, Focused map of  
 940 N3-1 bound to RBDs in the up and down conformations with five CDRs involved in the binding interface. **c**, One face  
 941 of H-CDR3 contacts a conserved hydrophobic pocket (transparent royal blue surface) on RBD-up. Arg408 forms  
 942 multiple polar interactions with Tyr49 and Glu55. The other face of H-CDR3 contacts the ACE2-binding site on RBD-  
 943 down. H-CDR1 contacts the epitopes on RBD-up, but it is omitted for clarity. **d**, The epitope on RBD-down is centered  
 944 on Phe486, which fits into a hydrophobic surface formed by Trp96, Tyr58, Tyr52, Arg100e and Val100g (clockwise).  
 945 Arg100e also forms a cation- $\pi$  interaction with Phe486 and a hydrogen bond with Tyr489. **e**, Superimposed crystal  
 946 structure of RBD-ACE2 complex (PDB ID: 6M0J) with N3-1 bound RBDs. The molecular surface of ACE2 is shown in  
 947 transparent pale green. The light chain of N3-1 heavily clashes with ACE2. The ACE2-binding site on RBD-down is  
 948 completely blocked by H-CDR2, H-CDR3 and L-CDR3 **[d]**.  
 949

950

951

952



953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964

**Figure 6. N3-1 binds to circulating spike variants.**

**a**, Model of N3-1 bound to RBD-up and RBD-down. Mutation sites in B.1.1.7, B.1.1.248, and B.1.351 are highlighted in green. **b**, The B.1.1.7, B.1.1.248, and B.1.351 spike variants were tested with A7V3, N3-1, and chimeric human ACE2-Fc using mammalian display of spike variants and compared to SARS-CoV-2 Wuhan-Hu-1 S-HexaPro-D614G (baseline arrow). **c-e**, Binding of IgG N3-1 to SARS-CoV-2-S WuHan-Hu-1 [**c**], variants B.1.1.7 [**d**] and B.1.351 [**e**] were assessed by surface plasmon resonance using an NTA sensor chip. Binding data are shown as black lines. For [**c-e**], the best fit was achieved using a 1:1 binding model and shown as red lines. **f**, Neutralization of B.1.1.7 and WA/1 by N3-1 and A7V3, where neutralization is measured as the percent of replicates showing complete neutralization at a given concentration (see methods).

965 **Ethics statement**

966 The acquisition of blood specimens from convalescent individuals was approved by the  
967 University of Texas at Austin Institutional Review Board (protocol 2020-03-085; Breadth of  
968 serum antibody immune responses prior to, or following, patient recovery in asymptomatic and  
969 non-severe COVID-19). Informed consent was obtained from all participants.

970

971 **Data availability**

972 The sequences of neutralizing antibodies have been deposited in GenBank with accession  
973 numbers. Molecular coordinates for A7V3 and N3-1 Fab complexes with SARS-CoV-2 trimeric  
974 spike protein have been deposited to the Protein Data Bank. Structural data are presented in  
975 Figs. 4-5, Extended Data Tables 3-4, and Extended Data Figs. 10-12 and 14.

976

977 **Author Contributions**

978 Conceptualization: JG, CH, AH, ECG, DRB, JSM, and JDG; Methodology: JG, CH, AH, ECG,  
979 FB, NW, KJ, AH, RR, MJJ, SLW, ZLN, JL, TSS, SVK, VK, RAH, IF, DRB, JSM, and JDG;  
980 Investigation: JG, CH, AH, ECG, FB, NW, AA, KJ, AH, WNV, JAC, ACM, RR, MJJ, SLW, ZLN,  
981 JL, RAH, SVK, VK, IJF, JMM, DRB, JSM, and JDG; Data Analysis and Interpretation: JG, CH,  
982 AH, DRB, NW, AH, JMD, IJF, EMM, JSM, and JDG; Data Curation: JG, CH, AH, ECG, NW,  
983 DRB, JSM, AH, JDM, JSM, and JDG; Original Draft: JG, CH, AH, ECG, DRB, JMM, JSM, VK,  
984 EMM, and JDG; Review & Editing: JG, CH, AH, ECG, DRB, ADE, EMM, JMM, IJF, GG, JSM,  
985 and JDG; Funding: JMD, IJF, GCI, VK, SVK, GG, ADE, JSM and JDG.

986

987 **Acknowledgements**

988 We would like to thank the many frontline responders, who have sacrificed tirelessly for the  
989 health and treatment of COVID-19 patients. In particular, we would like to extend our gratitude

990 to Daniel Billick for acquiring blood samples from each of the patients represented here. We  
991 thank Dr. Kathryn Stockbauer for careful editing of the manuscript. We thank Samantha Nagel  
992 for her tremendous efforts. We also would like to thank Dr. Sasha Dickinson from the Sauer  
993 Structural Biology Laboratory at the University of Texas at Austin for his assistance with  
994 microscope data collection. The Sauer Structural Biology Laboratory is supported by the  
995 University of Texas College of Natural Sciences and by award RR160023 from the Cancer  
996 Prevention and Research Institute of Texas (CPRIT). Funding for USAMRIID was provided  
997 through the CARES Act with programmatic oversight from the Military Infectious Diseases  
998 Research Program—project 14066041. Opinions, conclusions, interpretations, and  
999 recommendations are those of the authors and are not necessarily endorsed by the U.S. Army.  
1000 The mention of trade names or commercial products does not constitute endorsement or  
1001 recommendation for use by the Department of the Army or the Department of Defense.  
1002 Molecular graphics and analyses were performed with UCSF Chimera, developed by the  
1003 Resource for Biocomputing, Visualization, and Informatics at the University of California, San  
1004 Francisco, with support from NIH P41-GM103311. This research was funded in part by: the  
1005 Army Research Laboratory's TRANSFORME Essential Research Program (JDG, DRB, RR,  
1006 THSS); a Cooperative Agreement (W911NF-17-2-0091) between ARL and UT Austin to ADE,  
1007 EMM, and GG; a grant from DTRA (HDTRA12010011) awarded to JDG and ADE; a National  
1008 Institutes of Health (NIH)/National Institute of Allergy and Infectious Diseases (NIAID) grant  
1009 awarded to JSM (R01 AI127521); financial assistance award 70NANB20H037 to ZLN, SLW,  
1010 and JDG from the US Department of Commerce, National Institutes of Standards and  
1011 Technology (NIST) via the National Institute for Innovation in Manufacturing Biopharmaceuticals  
1012 (NIIMBL); a Welch grant (F-1016) awarded to IJF; a Welch grant (F-1515) awarded to EMM;  
1013 funding support from the Huck Institutes of Life Sciences and the Pennsylvania Agricultural

1014 Experiment Station (to VK and SK); an NIH grant (R01 AI158177-01) to SK; and an NIH grant  
1015 (R35 GM122480) to EMM. GCI expended discretionary funds from his MBS-20 account.

1016

1017

1018 **Competing Interest Statement**

1019 JL, ADE, EMM, GG, and DRB declare competing financial interests in the form of provisional  
1020 and granted patent applications relevant to Ig-Seq. JG, CH, ECG, AH, DRB, EMM, JSM, GCI,  
1021 ADE, GG, and JDG have filed provisional applications for the discovery of neutralizing  
1022 antibodies. JG, DRB, ECG, AH, and JDG have filed applications for additional methods relevant  
1023 to this work.

1024