The landscape of antibody binding affinity in SARS-CoV-2 Omicron BA.1 evolution

Alief Moulana[1]*, **Thomas Dupic**[1]*, **Angela M. Phillips**[1]*[†], **Jeffrey Chang**[2]*, **Anne A. Roffler**[3], **Allison J. Greaney**[4,5,6], **Tyler N. Starr**[4], **Jesse D. Bloom**[4,5,7], **Michael M. Desai**[1,2,8,9†]

[1]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge MA 02138,* [2]*Department of Physics, Harvard University, Cambridge, MA 02138,* [3]*Biological and Biomedical Sciences, Harvard Medical School, Boston MA 02115,* [4]*Basic Sciences Division and Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA 98109,* [5]*Department of Genome Sciences, University of Washington, Seattle, WA 98195,* [6]*Medical Scientist Training Program, University of Washington, Seattle, WA 98195,* [7]*Howard Hughes Medical Institute, Seattle, WA 98109,* [8]*NSF-Simons Center for Mathematical and Statistical Analysis of Biology, Harvard University, Cambridge MA 02138,* [9]*Quantitative Biology Initiative, Harvard University, Cambridge MA 02138.*
These authors contributed equally to this work.
[†]*angela.phillips@ucsf.edu, mdesai@oeb.harvard.edu*

ABSTRACT

The Omicron BA.1 variant of SARS-CoV-2 escapes convalescent sera and monoclonal antibodies that are effective against earlier strains of the virus. This immune evasion is largely a consequence of mutations in the BA.1 receptor binding domain (RBD), the major antigenic target of SARS-CoV-2. Previous studies have identified several key RBD mutations leading to escape from most antibodies. However, little is known about how these escape mutations interact with each other and with other mutations in the RBD. Here, we systematically map these interactions by measuring the binding affinity of all possible combinations of these 15 RBD mutations (2^{15} = 32,768 genotypes) to four monoclonal antibodies (LY-CoV016, LY-CoV555, REGN10987, and S309) with distinct epitopes. We find that BA.1 can lose affinity to diverse antibodies by acquiring a few large-effect mutations and can reduce affinity to others through several small-effect mutations. However, our results also reveal alternative pathways to antibody escape that do not include every large-effect mutation. Moreover, epistatic interactions are shown to constrain affinity decline in S309 but only modestly shape the affinity landscapes of other antibodies. Together with previous work on the ACE2 affinity landscape, our results suggest that escape of each antibody is mediated by distinct groups of mutations, whose deleterious effects on ACE2 affinity are compensated by another distinct group of mutations (most notably Q498R and N501Y).

1 In November 2021, the SARS-CoV-2 Omicron BA.1 variant emerged and quickly rose to high
2 frequency worldwide, in part due to its ability to escape preexisting immunity[1–4]. This immune
3 escape is mediated by mutations in the receptor binding domain (RBD) of the spike protein, which
4 is the major target of SARS-CoV-2 neutralizing antibodies[5–8]. Antibodies targeting the RBD can
5 bind different epitopes, and they have been grouped into several classes[9,10]. Some previous
6 SARS-CoV-2 variants which have a subset of the fifteen mutations found in the BA.1 RBD (e.g.
7 K417N, N501Y in Beta and T478K, N501Y in Delta) can evade some antibodies of certain epitope
8 classes but still bind to others[11–14]. In contrast, BA.1 can escape most antibodies that bind to very
9 distinct epitopes, including antibodies elicited by previously circulating variants[1,15,16].
10
11 Existing studies of SARS-CoV-2 immune escape have focused on measuring the effects of single
12 mutations (or, in some cases, of a small subset of mutations) on antibody escape in the context
13 of specific SARS-CoV-2 variants[15,17,18]. However, simultaneous escape of most antibodies is likely
14 to require multiple mutations, and it is unclear how these mutations might interact. A large body
15 of work has demonstrated that the specific combination of mutations in the BA.1 variant can evade
16 various antibodies of distinct epitopes[1,3,15,19]. However, the landscape on which this evolution
17 occurred is not well understood. Do mutations involved in escape from one antibody with a certain
18 epitope interfere with the effects of those involved in escaping others with different contact sites,
19 or are the effects largely independent? And how are these effects mediated by epistatic
20 interactions with other mutations in the RBD?
21
22 As we observed in previous work, several of these antibody-escape mutations also reduce affinity
23 to ACE2, suggesting that they were positively selected because they contribute to immune
24 escape[20–22]. Importantly, epistatic interactions between these mutations dramatically impact
25 ACE2 affinity and may also differentially impact the escape of antibodies with very different
26 epitopes[23,24]. For example, escape from some antibodies like S309 has been difficult to attribute
27 to specific mutations[22,25], perhaps because measurements have so far been limited to single
28 mutations. These observations suggest that we need to more comprehensively characterize the
29 role of epistasis and potential trade-offs to understand the simultaneous evolution of escape from
30 multiple antibodies of distinct epitopes and ACE2 binding affinity.
31
32 Here, to understand how immune pressure may have shaped the evolution of BA.1, we measured
33 the equilibrium binding affinities ($K_{D, app}$) of the spike protein RBD to four therapeutic monoclonal
34 antibodies (mAbs) with distinct RBD epitopes: LY-CoV016, LY-CoV555, REGN10987, and S309,
35 for all possible evolutionary intermediates between the ancestral Wuhan Hu-1 RBD and the BA.1
36 variant. This set of antibodies includes the primary epitopes generally covered by therapeutic
37 mAbs[10,16]. The first three antibodies are fully escaped by Omicron BA.1, while S309 has reduced
38 affinity. We find that for each antibody, only a few mutations significantly impact affinity, and these
39 mutations are largely (but not entirely) orthogonal between the four antibodies. Additionally, we
40 find that epistasis plays a limited role in determining affinity to antibodies that are fully escaped
41 by BA.1 but contributes substantially to the reduced affinity for the partially escaped antibody,
42 S309. Together, this work systematically characterizes how SARS-CoV-2 can evade distinct
43 RBD-targeted antibodies while maintaining ACE2 affinity.
44
45 **RESULTS**
46 In previous work, we generated a combinatorially complete library comprising all possible
47 intermediates between the ancestral SARS-CoV-2 Wuhan Hu-1 spike protein RBD and the
48 Omicron BA.1 variant[23]. The BA.1 RBD differs from Wuhan-1 by fifteen amino acid substitutions,
49 so this library contains 2^{15} variants containing all possible combinations of these fifteen mutations.
50 This RBD library is displayed on the surface of yeast, such that each yeast cell expresses a single
51 variant. Here, we use Tite-Seq (a high-throughput method that integrates flow cytometry and

52   sequencing[23,26–28]; see Supplementary Figure S1A) to measure the equilibrium binding affinities
53   of all 32,768 variants to four antibodies with different epitopes (LY-CoV016, LY-CoV555,
54   REGN10987, and S309). The resulting $K_{D, app}$ correlates between biological duplicates and with
55   isogenic measurements made by flow cytometry (Supplementary Figure S1A, Supplementary
56   Table S1).
57
58   Of the 32,768 variants in our library, we obtain $K_{D, app}$ for at least ~30,000 variants to each of the
59   mAbs (32,112 for Ly-CoV016, 30479 for REGN10987, 29892 for CoV555, and 32602 for S309)
60   after removing variants with poor titration curves ($r^2 < 0.8$ or $\sigma > 1$; see Methods). These $K_{D,app}$
61   range from 0.1 nM to 1 µM (which is our limit of detection and likely corresponds to nonspecific
62   binding), with 50% of the variants fully escaping LY-CoV016 (defined as having $K_{D,app}$ above the
63   limit of detection), 55% fully escaping LY-CoV555, 33% fully escaping REGN10897, and no
64   variants fully escaping S309 (Figure 1A; see https://desai-lab.github.io/wuhan_to_omicron/ for
65   an interactive data browser). Escape from LY-CoV016, LY-CoV555, and REGN10897 is
66   mediated by one or a few strong-effect mutations, with other mutations more subtly impacting
67   affinity (Figure 1B). In general, strong-effect mutations make substantial contact with the
68   corresponding antibody. Consistent with previous studies[14,16,17,29], these strong-effect mutations
69   are largely distinct for each antibody, which presumably reflects their non-overlapping footprints
70   on the RBD (Figure 1C) and suggests that evolution of escape from each antibody can be, to
71   some extent, orthogonal.
72
73   The picture is more complex for S309, where BA.1 has reduced affinity relative to Wuhan Hu-1,
74   but ~19% of variants have lower affinity than BA.1. These differences are not attributable to one
75   or two strong effect mutations (Figure 1A-B). In addition, although most mutations reduce affinity,
76   three mutations have small positive effects (on average across all backgrounds at the other loci):
77   S375F for LY-CoV016, E484A and N501Y for REGN10987 (Figure 1B). Intriguingly, each of these
78   mutations reduces affinity to at least one of the other antibodies, and N501Y significantly improves
79   binding to ACE2, suggesting a potential role for trade-offs (and/or epistasis that mitigates these
80   effects on specific backgrounds).
81
82   For each antibody, binding affinities generally decrease as the number of mutations increase
83   (Figure 1D-G). For LY-CoV016, LY-CoV555, and REGN10897 this trend is observed amongst
84   variants with and without the large-effect escape mutations (Figure 1D-F). For LY-CoV016,
85   K417N is sufficient for escape (Figure 1D, green), whereas both LY-CoV555 and REGN10987
86   require at least two mutations for complete escape. For LY-CoV555, both E484A and Q493R
87   decrease affinity drastically (1000 and 100-fold, respectively), but only the combination of both
88   mutations lead to complete escape. Complete escape from REGN10987 also requires two
89   mutations (N440K and G446S) but the individual effects of these mutations are more subtle
90   (reducing affinity by 10 and 5-fold, respectively). For S309, affinity declines after a few mutations
91   and in some backgrounds increases upon further mutation, suggesting that interactions between
92   these mutations are important in determining affinity (Figure 1G).
93
94   ***Mostly orthogonal large-effect mutations***
95   We first focused on analyzing how mutations and combinations of mutations lead to complete
96   escape (defined as $K_{D,app}$ above our limit of detection) for specific sets of antibodies. To do so, we
97   analyze the enrichment of specific mutations among non-binders (Figure 2A). We find a largely
98   orthogonal set of one or two mutations are enriched among variants that do not bind each
99   antibody: almost all variants that do not bind LY-CoV016 contain K417N, almost all variants that
100   do not bind REGN10987 contain G446S and many also contain N440K, and E484A and Q493R
101   are highly enriched among variants that do not bind LY-CoV555. This suggests that the RBD can

3

102 evolve to independently escape antibodies with each distinct epitope and mutations can to some
103 extent act independently on binding to each antibody.
104
105 To analyze this further, we calculate the percentage of genetic backgrounds on which each
106 mutation leads to complete escape from a specific antibody (i.e. that mutation converts a variant
107 with measurable $K_{D,app}$ to a $K_{D,app}$ above our limit of detection). We see that for each antibody, one
108 or two mutations abrogate binding. These sets of mutations are largely orthogonal among
109 antibodies (Figure 2B), consistent with the enrichment analysis (Figure 2A). Specifically, K417N
110 always abrogates binding to LY-CoV016, G446S and N440K often abrogate binding to
111 REGN10987, and E484A and Q493R often abrogate binding to LY-CoV555.
112
113 However, we note that this orthogonality is not complete. For example, G446S is slightly enriched
114 among LY-CoV555 binders, while mutation E484A is slightly depleted among variants that do not
115 bind REGN10987 (Figure 2A). Consistent with this, G446S sometimes abrogates binding to LY-
116 CoV555 (Figure 2B). In addition, some apparently smaller-effect mutations can be involved in
117 abolishing binding to multiple antibodies. For example, S375F is weakly enriched among variants
118 that do not bind REGN10987 and LY-CoV555 and often abrogates binding to these two
119 antibodies, with G496S, Q498R, and N501Y also playing a role.
120
121 To summarize how these different mutations can act individually or in combination to lead to
122 antibody escape, we inferred a decision tree to classify variants as binders or non-binders. To do
123 so, for each antibody we calculate the mutation that maximally partitions the variants into binders
124 or non-binders. If this partitioning is not perfect, we then calculate the second mutation that
125 maximally partitions the variants conditional on each possible state of the first site. We then
126 proceed to further partition variants based on additional mutations in the same way, until the
127 variants are perfectly partitioned or no further mutations can significantly improve the partitioning
128 (see Methods). We show the corresponding decision trees for LY-CoV016, REGN10987, and LY-
129 CoV555 in Figures 2C, D, and E respectively. As expected, the tree associated with LY-CoV016
130 is very simple: the mutation K417N perfectly partitions the variants into binders and non-binders.
131 In contrast, the trees for REGN10987 and LY-CoV555 have more complex structures, reflecting
132 the fact that it is possible to abrogate affinity to these antibodies via multiple distinct combinations
133 of mutations. For example, variants can escape REGN10987 by acquiring G446S and N440K
134 (100%), or alternatively, with S375F and G446S (77%, as additional mutations are also required).
135 For LY-CoV555, different sets of mutations can lead to escape (e.g. Q493R and G446S or E484A
136 and S375F), and G496S can help or hinder escape depending on the presence of other mutations.
137 Some of the mutations resulting in LY-CoV555 escape partially overlap with those for
138 REGN10987 (i.e. they are not fully orthogonal), suggesting that selection pressure from one
139 antibody could promote subsequent escape of another.
140
Inference of epistatic affinity landscapes
142 In addition to large-effect mutations which lead to complete escape of specific antibodies, a variety
143 of other sites contribute to more subtle but potentially important changes in binding affinities. To
144 analyze these subtle effects as well as the large-effect mutations leading to escape, we defined
145 a linear model for $-\log(K_{D,\,app})$ as the sum of single (additive) mutational effects plus interaction
146 terms up to a specified order (note that because $-\log(K_{D,\,app})$ is proportional to the free energy of
147 binding, we expect it to behave additively in the absence of epistatic interactions). Because non-
148 binding variants have $-\log(K_{D,\,app})$ beyond our limit of detection, we fit a Tobit model (a class of
149 regression model capable of handling truncated measurements, see Methods for details) using
150 maximum likelihood with an L2-norm Lasso regularization. Specifically, we partition our data into
151 training (90%) and test (10%) sets and use the training dataset to fit epistatic coefficients to a
152 linear model truncated at each order (e.g. truncating to first-order yields additive mutational

153   effects, second-order includes both additive effects and pairwise terms, and so on). We then
154   evaluate performance (as the coefficient of variation) of each model on the held-out test dataset,
155   and compare the model performance using -log($K_{D, app}$) for each of the antibodies and ACE2
156   (Figure 3A).

157

158   We find that adding epistatic interactions improves the predictive power of the model for all four
159   antibodies as well as for binding to ACE2, though the optimal order varies (Figure 3A). This
160   indicates that epistasis does play a significant role in all cases (up to second order for
161   REGN10987, to third order for LY-CoV555, and to fourth or higher order for LY-CoV016, S309,
162   and ACE2). The additive, pairwise, and higher-order coefficients resulting from these models are
163   summarized in Supplementary Figure S2. In general, we find many strong interactions across
164   several positions in each antibody, involving both the sites that strongly determine escape variants
165   for that antibody (e.g. between N440K and G446S for REGN10987) as well as others.

166

167   Notably, we find that the higher-order epistasis plays a much stronger role in determining affinity
168   for ACE2 than for the three antibodies fully escaped by BA.1 (Figure 3B). This reflects the impact
169   of a few strong-effect mutations in determining affinity for LY-CoV555, LY-CoV016, and
170   REGN10987, and the role of compensatory epistasis in determining ACE2 affinity. In other words,
171   while epistasis is relevant for all measured phenotypes, antibody escape is more simply
172   determined by the additive effects of individual mutations, while maintaining ACE2 affinity involves
173   more complex epistatic interactions.

174

175   High-order epistatic interactions are also important in determining affinity to S309. In Figure 3C
176   we highlight four neighboring mutations (G339D, S371L, S373P, and S375F) which interact non-
177   additively to produce the reduction in affinity observed in BA.1 relative to Wuhan Hu-1. Each of
178   these mutations weakly reduces affinity on their own, and specific combinations of these
179   mutations can reduce affinity by up to two orders of magnitude, but the reduction in affinity
180   resulting from all four mutations is less than some sets of three mutations. These patterns emerge
181   from a complex set of high-order epistatic interactions among the mutations. For example, S371L
182   reduces affinity on the Wuhan Hu-1 background but increases affinity on the background
183   containing G339D, S373P, and S375F (and without S375F, S371L increases affinity in the
184   presence of S373P if G339D is absent but not if it is present). Thus, some variants lacking S371L
185   evade S309 more effectively than BA.1, and interestingly we note that this mutation is absent in
186   BA.2 and BA.3 and replaced instead by S371F.

187

188   ***Tradeoffs between antibody and ACE2 affinities***
189   In previous work, we found that antibody escape mutations (as defined in earlier studies) typically
190   reduce ACE2 affinity, suggesting that viral evolution is constrained by a tradeoff between immune
191   evasion and the ability to enter host cells. Consistent with this, we find here that variants that
192   escape one or more antibodies (as defined by the data reported in this work) but have few
193   additional mutations have reduced ACE2 affinity relative to Wuhan Hu-1. However, as additional
194   BA.1 mutations are accumulated, the ACE2 binding affinity tends to increase until it exceeds the
195   Wuhan Hu-1 value even in the presence of multiple antibody escape mutations (Figure 4A). This
196   suggests that the evolution of the BA.1 variant is driven both by immune escape and the need for
197   compensatory mutations that mitigate the negative effects of the escape mutations on ACE2
198   binding.

199

200   The strength of this tradeoff and the potential importance of compensatory evolution is distinct
201   between the different antibodies (Figure 4B,C). For example, escape from LY-CoV016 or LY-
202   CoV555 reduces ACE2 binding affinity in the absence of compensatory mutations (Q498R and
203   N501Y), but not in their presence (Figure 4B). In contrast, REGN10987 escape does not strongly

204   reduce affinity to ACE2, whether or not Q498R and N501Y are present. However, this tradeoff is
205   likely relevant overall, as escaping all three antibodies substantially reduces ACE2 affinity in the
206   absence of Q498R and N501Y, while the reduction in ACE2 affinity is minimal in their presence.
207   Consistent with this general picture, the frequency of most escape mutations (G446S, E484A,
208   and Q493R) is higher across the SARS-CoV-2 phylogeny in the presence of compensatory
209   mutations (Figure 4D), though we note that this is not universally true (e.g. the frequency of N440K
210   is only slightly higher in the presence of compensatory mutations, and the frequency of K417N is
211   lower with the compensatory mutations).

213   Although antibody escape mutations do tend to reduce ACE2 affinity, antibody binding affinity
214   (but not complete escape) is not strongly correlated with ACE2 affinity (Figure 4C). The details of
215   this relationship vary by antibody. For LY-CoV016 and LY–CoV555, there is a weak overall
216   positive correlation (i.e. lower antibody affinity also tends to correspond to reduced ACE2 affinity).
217   However, this correlation is dominated by the variants that lack the compensatory mutations at
218   sites 498 and 501; in the presence of Q498R and N501Y the correlation largely disappears. For
219   REGN10987, there is a similar weak overall positive correlation, which is less dependent on
220   Q498R and N501Y. Finally, for S309, there is essentially no correlation regardless of whether
221   compensatory mutations are present. In other words, for this antibody it appears possible to
222   evolve reduced affinity (at least to the extent seen in BA.1) without compromising ACE2 binding.
223   We also note that while compensatory mutations Q498R and N501Y largely drive the variance in
224   ACE2 affinity, they minimally impact antibody binding affinities.

DISCUSSION
227   Overall, we find that BA.1 escape from LY-CoV016, LY-CoV555, and REGN10987 is driven by a
228   relatively small set of mutations: K417N for LY-CoV016, N440K and G446S for REGN10987,
229   E484A and Q493R for LY-CoV555. These mutations have largely orthogonal effects on affinity to
230   the three antibodies, suggesting that the evolution of escape to each can occur independently, as
231   might be expected given the distinct epitopes they target[9,16,18].

233   However, despite these largely orthogonal effects of large-effect mutations on antibody escape,
234   we do observe limited trade-offs between LY-CoV016 and LY-CoV555 and between LY-CoV555
235   and REGN10987, with three mutations (S375F, K417N, and E484A) improving affinity to one
236   antibody and reducing affinity to another. These positive effects are modest compared to the
237   reductions in affinity caused by other mutations. In fact, for all antibodies studied here, outside of
238   a few large-effect mutations that abrogate or nearly abrogate binding, most mutations weakly
239   impact binding affinity and, even collectively, are insufficient to abrogate it.

241   In contrast to the orthogonality of antibody escape, trade-offs between binding ACE2 and
242   escaping antibodies are much stronger. While the mutations with a small effect on antibody
243   escape are mostly uncorrelated to ACE2 affinity, the strong-effect mutations substantially reduce
244   ACE2 affinity. Thus, ACE2 affinity is lower for variants that escape a larger number of antibodies,
245   unless compensatory mutations are acquired, suggesting that these compensatory mutations
246   potentiated the establishment of the antibody escape mutations[23,30].

248   We also find that epistatic interactions are important in determining antibody affinity. This is
249   particularly true for S309. Prior to this work, the reduced affinity of S309 to Omicron could not be
250   attributed to specific mutations[16,25]. Here, we find that this ambiguity can be resolved by examining
251   higher-order interactions between mutations, as the reduction in affinity is attributable to a fourth-
252   order epistatic interaction. This finding suggests that the potential for future SARS-CoV-2 lineages
253   to escape S309 and similar antibodies could depend on epistatic interactions between emerging
254   mutations.

255
256   We note that our study focuses only on binding affinities, which may not always perfectly reflect
257   viral escape from antibody neutralization[31,32]. In particular, some of the binding affinities we
258   measure could be too weak to be physiologically relevant, and mutations may impact
259   neutralization without impacting binding affinity significantly. However, because neutralization
260   cannot occur in the absence of affinity, our measurements are likely to be relevant for
261   understanding the reduced sensitivity of BA.1 to these antibodies. We also note that practical
262   constraints limit us to studying four antibodies. This limits the generalizability of our results,
263   particularly in light of previous structural studies which have revealed more epitopes bound by
264   mAbs.
265
266   In spite of these limitations, our binding affinity landscapes reveal that BA.1 can escape diverse
267   antibodies by acquiring a few large-effect mutations and can reduce affinity to others by
268   accumulating several small-effect mutations. For the first three antibodies, one or two mutations
269   are sufficient for total escape. However, in some cases, additional mutations can restore affinity,
270   and in others, specific combinations of large and small-effect mutations can abrogate affinity.
271   Thus, despite the seemingly simple landscape of antibody escape, there are alternative, more
272   intricate pathways that can abrogate affinity. In contrast, for the S309 antibody, four mutations
273   drive the decline in affinity yet are also involved in higher-order epistatic interactions that
274   counteract this decline. This epistasis results in an affinity threshold, beyond which additional
275   mutations do not reduce affinity.
276
277   Predicting the future evolution of the Omicron lineage will require determining how these affinity
278   landscapes translate to immune evasion, how antibody affinity landscapes vary within a class or
279   epitope group, and how mutations beyond this set may further enhance immune evasion. For
280   example, neutralization assays with minimally mutated genotypes would confirm whether the
281   strong-effect mutations are indeed sufficient for escape. Further, assessing affinity landscapes
282   for additional antibodies with similar epitopes would reveal how the landscape structure varies
283   within such a group, and whether there are general features that we can extrapolate to
284   unmeasured sequences. Finally, integrating these combinatorial libraries with saturating
285   mutagenesis approaches would reveal how the evolvability of this lineage changes over time, and
286   what additional mutations – such as those in BA.2, BA.4, or BA.5 – might confer further immune
287   escape. Looking beyond the Omicron lineage, such approaches could provide more general
288   insight into how mutations in SARS-CoV-2 may result in host-range expansion or antigenic
289   evolution.

290 **METHODS**
291 ***Yeast display plasmid, strains, and library production***
292 We used the same library and strains as produced in Moulana et al[23]. In brief, to generate clonal
293 yeast strains for the Wuhan Hu-1 and Omicron BA.1 variants, we cloned the corresponding
294 RBD gblock into a pETcon vector via Gibson Assembly. We then extracted and transformed
295 Sanger-verified plasmids into the AWY101 yeast strain (kind gift from Dr. Eric Shusta)[33] as
296 described in Gietz and Schiestl[34]. To produce the RBD variant library, we employed a Golden
297 Gate combinatorial assembly strategy. We constructed full RBD sequences from five sets of
298 dsDNA fragments of roughly equal size. Each set contains versions of the fragments that differ
299 by the mutations included. Following bacterial transformation of this Golden Gate assembly
300 product, we extracted and transformed the library into AWY101 yeast strain, from which we
301 inoculated and froze a library containing obtained ~1.2 million colonies.
302
303 ***High-throughput binding affinity assay (Tite-Seq)***
304 We performed Tite-seq assay as previously described[20,23,26,27], with two replicates for each
305 antibody (LY-CoV016, LY-CoV555, REGN10987, and S309 [Genscript, Gene-to-Antibody
306 service]) assay on different days, for a total of eight assays.
307
308 Briefly, we thawed yeast RBD library and the Wuhan Hu-1 and Omicron BA.1 strains by
309 inoculating the corresponding glycerol stocks in SDCAA (6.7 g/L YNB without amino acid [VWR
310 #90004-150], 5 g/L ammonium sulfate [Sigma-Aldrich #A4418], 2% dextrose [VWR #90000–
311 904], 5 g/L Bacto casamino acids [VWR #223050], 1.065 g/L MES buffer [Cayman Chemical,
312 Ann Arbor, MI, #70310], 100 g/L ampicillin [VWR # V0339])at 30°C for 20 hr. The cultures were
313 then induced in SGDCAA (6.7 g/L YNB without amino acid [VWR #90004-150], 5 g/L
314 ammonium sulfate [Sigma-Aldrich #A4418], 2% galactose [Sigma-Aldrich #G0625], 0.1%
315 dextrose [VWR #90000–904], 5 g/L Bacto casamino acids [VWR #223050], 1.065 g/L MES
316 buffer [Cayman Chemical, Ann Arbor, MI, #70310], 100 g/L ampicillin [VWR # V0339]), and
317 rotated at room temperature for 16–20 hr.
318
319 Following overnight induction, we pelleted, washed (with 0.01% PBSA [VWR #45001–130;
320 GoldBio, St. Louis, MO, #A-420–50]), and incubated the cultures with monoclonal antibody at a
321 range of concentrations (10^{-6} to 10^{-12} with 0.75-log increments for CoV555, 10^{-7} to 10^{-12} with 0.5-
322 log increments for S309, 10^{-6} to $10^{-12.7}$ with 0.75-log increments for REGN10987, 10^{-6} to 10^{-12}
323 with 0.75-log increments for SB6). The yeast-antibody mixtures were incubated at room
324 temperature for 20 hr. The cultures were then pelleted washed twice with PBSA and
325 subsequently labeled with PE-conjugated goat anti-human IgG (1:100, Jackson
326 ImmunoResearch Labs #109-115-098) and FITC-conjugated chicken anti-cMmyc (1:100,
327 Immunology Consultants Laboratory Inc., Portland, OR, #CMYC-45F). The mixtures were
328 rotated at 4°C for 45 min and then washed twice in 0.01% PBSA.
329
330 Sorting, recovery, and sequencing library preparation followed Moulana, et al[23]. In short, we
331 sorted ~1.2 million yeast cells per concentration, gated by FSC vs SSC and then by expression
332 (FITC) and/or binding fluorescence (PE) on a BD FACS Aria Illu. The machine was equipped
333 with 405 nm, 440 nm, 488 nm, 561 nm, and 635 nm lasers, and an 85 micron fixed nozzle.
334 Sorted cells were then pelleted, resuspended in SDCAA, and rotated at 30°C until late-log phase
335 (OD600 = 0.9-1.4). The cultures were then pelleted and stored at -20°C for at least six hours
336 prior to extraction using Zymo Yeast Plasmid Miniprep II (Zymo Research # D2004), following
337 the manufacturer's protocol. The sequencing amplicon libraries were then prepared by a two-
338 step PCR as previously described[23,27,35]. In brief, we added to the amplicon unique molecular
339 identifies (UMI), inline indices, and partial Illumina adapters through a 7-cycle PCR which
340 amplifies the RBD sequence in the plasmid. We then used the cleaned product from the first

8

341   PCR in the second PCR to append Illumina i5 and i7 indices accordingly (see
342   https://github.com/desai-lab/compensatory_epistasis_omicron/tree/main/Supplementary_Files
343   for primer sequences). The products were then cleaned using 0.85x Aline beads, verified using
344   1% agarose gel, quantified on Spectramax i3, pooled, and verified on Tapestation 5000HS and
345   1000HS. Final library was quantitated by Qubit fluorometer and sequenced on Illumina Novaseq
346   SP, supplemented with 10% PhiX.
347

348   ***Sequence data processing***
349   Following Moulana et al.[23], we processed raw demultiplexed sequencing reads to identify and
350   extract the indexes and mutational sites. Briefly, for each antibody, we utilized a snakemake
351   pipeline (https://github.com/desai-lab/omicron_ab_landscape) to parse through all fastq files and
352   group the reads according to inline indices, UMIs, and sequence reads. We accepted
353   sequences based on criteria previously determined (10% bp mismatches) and converted
354   accepted sequences into binary genotypes ('0' for Wuhan Hu-1 allele or '1' for Omicon BA.1
355   allele at each mutation position). Reads containing errors at mutation sites were removed.
356   Finally, the pipeline collated genotype counts based on distinct UMIs from all samples into a
357   single table.
358

359   We fit the binding dissociation constants $K_{D,app}$ for each genotype as previously described[23,27].
360   Briefly, using sequencing and flow cytometry data, we calculated the mean log-fluorescence of
361   each genotype s at each concentration c, as follows:

362
$$\bar{F}_{s,c} = \sum_b F_{b,c}\, p_{b,s|c},$$

363   where $F_{b,c}$ is the mean log-fluorescence of bin b at concentration c, and $p_{b,s|c}$ is the inferred
364   proportion of cells from genotype s that are sorted into bin b at concentration c, which is
365   estimated from the read counts as:

366
$$p_{b,s|c} = \frac{\frac{R_{b,s,c}}{\sum_s R_{b,s,c}} C_{b,c}}{\sum_b \left(\frac{R_{b,s,c}}{\sum_s R_{b,s,c}} C_{b,c} \right)} ,$$

367   Here, $R_{b,s,c}$ represents the number of reads from genotype s that are found in bin b at
368   concentration c, and $C_{b,c}$ refers to the number of cells sorted into bin b at concentration c.
369

370   We then computed the uncertainty for the mean log-fluorescence:

371
$$\delta \bar{F}_{s,c} = \sqrt{\sum_b \left(\delta F_{b,c}^2\, p_{b,s|c}^2 + F_{b,c}^2 \delta p_{b,s|c}^2 \right)}$$

372   where $\delta F_{b,c}$ is the spread of the log fluorescence of cells sorted into bin b at concentration c.
373   The error in $p_{b,s|c}$ emerges from the sampling error, which can be approximated as a Poisson
374   process, such that:

375
$$\delta p_{b,s|c} = \frac{p_{b,s|c}}{\sqrt{R_{b,s,c}}}.$$

376

377   Finally, we inferred the binding dissociation constant ($K_{D,s}$) for each variant by fitting the
378   logarithm of Hill function to the mean log-fluorescence $\bar{F}_{s,c}$, as a function of concentrations c:

379
$$\bar{F}_{s,c} = log_{10}\left(\frac{c}{c+K_{D,s}} A_s + B_s \right) ,$$

380   where A_s is the increase in fluorescence at antibody saturation, and B_s is the background
381   fluorescence level. The fit was performed using the *curve_fit* function in the Python package
382   *scipy.optimize*. Across all genotypes, we imposed bounds on the values of A_s to be 10^2-10^6, B_s
383   to be 1-10^5, and $K_{D,s}$ to be 10^{-14}-10^{-5}. We then averaged the inferred $K_{D,s}$ values across the two
384   replicates for each antibody after removing values with poor fit ($r^2 < 0.8$ or standard error > 1).
385   Variants were defined as non-binders if the difference between the maximum and the minimum

386 of their estimated log-fluorescence over all concentrations was lower than 1 (in log-fluorescence
387 units). This value was set by measuring the distribution for known non-binders (see
388 Supplementary Figure S3).
389
390 *Isogenic measurements for validation*
391 We validated our high-throughput binding affinity method by measuring the binding affinities for
392 the Wuhan Hu-1 and Omicron BA.1 RBD variants. For each isogenic titration curve, we followed
393 the same labeling strategy as in Tite-seq, titrating each antibody at concentrations ranging from
394 10^{-12}-10^{-7} M (with increments of 0.5 for the first replicate and 1 for the second one) for isogenic
395 yeast strains that display only the sequence of interest. The mean log fluorescence was
396 measured using a BD LSR Fortessa cell analyzer. We directly computed the mean and
397 variances of these distributions for each concentration and used them to infer the value of $K_{D,app}$
398 using the formula shown above.
399
400 *Decision trees on loss-of-binding mutations*
401 To summarize mutations that drive the loss of binding (escape) for each antibody, we
402 constructed a decision tree using package rpart in R[36] with its default parameters. In brief, for
403 each antibody (except for S309 where every sequence binds the antibody), we first categorized
404 each genotype into a binary parameter with values 'binding' or 'non-binding'. Then, the function
405 rpart splits the tree based on any one of the fifteen mutations by minimizing the Gini impurity for
406 the binding parameter. The method continues to partition the tree if the cost complexity
407 parameter (cp) of the split does not drop below 0.01. This parameter is the sum of all
408 misclassifications (binding vs. non-binding) at every terminal node (analogous to residual sum of
409 squares in regression), added by the product between the number of splits (analogous to
410 degree of freedom) and a penalty term inferred through cross-validation performed by the rpart
411 algorithm. The tree is then presented in Figure 2 using `ggparty` package[37].
412
413 *Epistasis analysis*
414 We used a linear model where the effects of combinations of mutations sum to the phenotype of
415 a sequence. The logarithm of the binding affinity $\log K_{D,app}$ is proportional to change in free
416 energy. Thus, without epistatic interactions, the effects of mutations are expected to combine
417 additively[38,39]. However, some phenotypic values $\log K_{D,app}$ are unavailable in our dataset due to
418 the upper limit of the assay concentration. We are unable to precisely determine $K_{D,app}$ for the
419 low-affinity (or non-binding) variants, especially when the true -$\log K_{D,app} < 6.0$ (the highest log-
420 concentration used). To deal with this problem, we supplemented our linear model with a
421 boundary, following tobit left-censored model. In this model, the $-\log K_{D,app}$ phenotype is
422 measured as 6.0 for all values lower than 6.0. Thus, the full *K*-order model can be written as:

423
$$-\log K_{D,s}^* = \beta_0 + \sum_{i=1}^{K} \sum_{c \in C_i} \beta_c x_{c,s} + \epsilon_s$$

424 where C_i contains all $\binom{L}{i}$ combinations of size i of the mutations and $x_{c,s}$ equal to 1 if the
425 sequence s contains all the mutations in c and to 0 otherwise. Here, $-\log K_{D,s} = -\log K_{D,s}^*$ if
426 $-\log K_{D,s}^* > 6$ and $-\log K_{D,s} = 6$ if $-\log K_{D,s}^* \leq 6$. Then, following the tobit model approach, we
427 compute the likelihood function to infer coefficient parameters $\hat{\beta}_{MLE}$, given by:
428

429
$$\mathcal{L}(\beta, \sigma) = \prod_{j=1}^{N} \left(\frac{1}{\sigma} \varphi \left(\frac{y_j - (\beta_0 + \sum_{i=1}^{K} \sum_{c \in C_i} \beta_c x_{c,s})}{\sigma} \right) \right)^{I(y_j)} \left(1 \right.$$

430
$$\left. - \Phi \left(\frac{(\beta_0 + \sum_{i=1}^{K} \sum_{c \in C_i} \beta_c x_{c,s}) - 6}{\sigma} \right) \right)^{1 - I(y_j)} \right)$$

431
432 where $y_j = -\log K_{D,\text{app},j}$, and φ and Φ denote the standard normal cumulative distribution
433 function and probability density function, respectively. Moreover, note that $I(y) = \begin{cases} 0 \text{ if } y \leq 6 \\ 1 \text{ otherwise} \end{cases}$.
434 This optimization problem would include coefficients that are associated with the loss-of-binding
435 phenotypes. Consequently, by the model, these coefficients do not have lower bounds and the
436 optimization would have resulted in deflated coefficients offset by inflated higher-order
437 coefficients, or vice-versa. To resolve this problem, we add a lasso regularization term in the
438 form of $\epsilon \sum |\beta_c|$ to the likelihood, with $\epsilon = 0.01$. This term is small enough to reduce the
439 magnitude of constrained coefficients but act as intended on the non-constrained ones. To
440 maximize the log-likelihood function, which is a concave function, we used the optimize module
441 in the scipy package, with the BFGS gradient-descent method.
442
443 ***Structural analysis***
444 We used the reference structure of a 2.79 Å cryo-EM structure of Omicron BA.1 complexed with
445 (PDB ID: 7WPB). The contact surface area is determined by using ChimeraX[40] to measure the
446 buried surface area between ACE2 and each mutated residue in the RBD (*measure buriedarea*
447 function, default probeRadius of 1.4Å), whereas distance between α-carbons is measured using
448 PyMol[41].
449
450 ***Force directed layout***
451 The high-dimensional binding affinity landscape can be projected in two dimensions with a
452 force-directed graph layout approach (see https://desai-lab.github.io/wuhan_to_omicron/). Each
453 node corresponds to each sequence in the library, connected by edges to a neighbor that differs
454 in one single site. For each antibody, an edge between two sequences s and t is given the
455 weight:

456
$$w_{s,t} = \frac{1}{0.01 + |\log K_{D,s} - \log K_{D,t}|}$$

457
458 Additionally, we also constructed a different layout that includes affinities to all antibodies, where
459 the weight between two sequences depends on the sum over the antibodies of the difference
460 between their affinities:
461

462
$$w_{s,t} = \frac{1}{0.01 + \sum_{a \in A} |\log K_{D,a,s} - \log K_{D,a,t}|}$$

463
464 where A is the set of antibodies we used. In a force-directed representation, the edges pull
465 together the nodes they are attached to proportional to the weight given to each edge. In our
466 scenario, this means that nodes with a similar genotype (a few mutations apart) and a similar
467 phenotype (binding affinity or total binding affinity) will be close to each other in two dimensions.
468
469 Importantly this is not a "landscape" representation: the distance between two points is
470 unrelated to how easy it is to reach one genotype from another in a particular selection model.

471 Practically, after assigning all edge weights, we use the layout function *layout_drl* from the
472 Python package *iGraph*, with default settings, to obtain the layout coordinates for each variant.
473
474 ***Genomic data***
475 To analyze SARS-CoV-2 phylogeny, we used all complete RBD sequences from all SARS-CoV-
476 2 genomes deposited in the Global Initiative on Sharing All Influenza Data (GISAID)
477 repository[42–44] with the GISAID Audacity global phylogeny (EPI_SET ID: EPI_SET_20220615uq,
478 available on GISAID up to June 15, 2022, and accessible at
479 https://doi.org/10.55876/gis8.220615uq) . We pruned the tree to remove all sequences with
480 RBD not matching any of the possible intermediates between Wuhan Hu-1 and Omicron BA.1
481 and analyzed this tree using the python toolkit ete3[45]. We measured the frequency of each
482 mutation by counting how many times it emerges in the tree, normalized by the total
483 occurrences of other mutations. For frequency with Q498R and N501Y, we counted the
484 occurrence of each mutation only on branches that already contains Q498R and N501Y and
485 normalized similarly.
486
487 ***Statistical analyses and visualization***
488 All data processing and statistical analyses were performed using R v4.1.0[46] and python
489 3.10.0[47]. All figures were generated using ggplot2[48] and matplotlib[49].

ACKNOWLEDGEMENTS

We thank Zach Niziolek for flow cytometry assistance and all members of the Desai lab and Serafina Nieves for helpful discussions. T.D. acknowledges support from the Human Frontier Science Program Postdoctoral Fellowship, A.M.P. acknowledges support from the Howard Hughes Medical Institute Hanna H. Gray Postdoctoral Fellowship, J.C. acknowledges support from the National Science Foundation Graduate Research Fellowship, and M.M.D. acknowledges support from the NSF-Simons Center for Mathematical and Statistical Analysis of Biology at Harvard University, supported by NSF grant no. DMS-1764269, and the Harvard FAS Quantitative Biology Initiative, grant DEB-1655960 from the NSF and grant GM104239 from the NIH. J.D.B. acknowledges support from NIH/NIAID grant R01AI141707 and is an Investigator of the Howard Hughes Medical Institute. We gratefully acknowledge all data contributors, i.e. the Authors and their Originating laboratories responsible for obtaining the specimens, and their Submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative. Computational work was performed on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University.

AUTHOR CONTRIBUTIONS

Conceptualization: A.M., T.D., A.M.P., J.C., T.N.S., A.J.G., J.D.B., and M.M.D. Methodology: A.M., T.D., A.M.P., J.C., S.N., T.N.S., and A.J.G. Library design and production: A.M., T.D., A.M.P., J.C., and A.J.G. Experiments: A.M., T.D., A.M.P., J.C., and A.A.R. Validation: A.M., T.D., A.M.P., J.C., S.N., and T.N.S. Data analysis: A.M., T.D., A.M.P., J.C., S.N., and T.N.S. Supervision: A.M.P, J.D.B., and M.M.D. Funding acquisition: J.D.B. and M.M.D. Writing— original draft: A.M., T.D., A.M.P., J.C., and M.M.D. All the authors reviewed and edited the manuscript.

COMPETING INTERESTS

A.M.P. and M.M.D. have or have recently consulted for Leyden Labs. J.D.B. has or has recently consulted for Apriori Bio, Oncorus, Moderna, and Merck. J.D.B., A.J.G., and T.N.S. are inventors on Fred Hutch licensed patents related to viral deep mutational scanning. The other authors declare no competing financial interests.

MATERIALS AND CORRESPONDENCE

Correspondence and requests for materials should be addressed to M.M.D. (mdesai@oeb.harvard.edu).

DATA AND CODE AVAILABILITY STATEMENT

Raw sequencing reads have been deposited in the NCBI BioProject database under accession number PRJNA877045. All associated metadata are available at https://github.com/desai-lab/omicron_ab_landscape.

13

REFERENCES CITED

1. Cao, Y. *et al.* Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **602**, 657–663 (2022).
2. Ao, D. *et al.* SARS-CoV-2 Omicron variant: Immune escape and vaccine development. *MedComm (2020)* **3**, e126 (2022).
3. Planas, D. *et al.* Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **602**, 671–675 (2022).
4. Viana, R. *et al.* Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* **603**, 679–686 (2022).
5. Greaney, A. J. *et al.* Mapping mutations to the SARS-CoV-2 RBD that escape binding by different classes of antibodies. *Nat. Commun.* **12**, 4196 (2021).
6. Greaney, A. J. *et al.* Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe* **29**, 44-57.e9 (2021).
7. Iketani, S. *et al.* Antibody evasion properties of SARS-CoV-2 Omicron sublineages. *Nature* **604**, 553–556 (2022).
8. Dai, L. & Gao, G. F. Viral targets for vaccines against COVID-19. *Nat. Rev. Immunol.* **21**, 73–82 (2021).
9. Barnes, C. O. *et al.* Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell* **182**, 828-842.e16 (2020).
10. Barnes, C. O. *et al.* SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).
11. Liu, C. *et al.* Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. *Cell* **184**, 4220-4236.e13 (2021).
12. Zhou, D. *et al.* Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* **184**, 2348-2361.e6 (2021).
13. Greaney, A. J. *et al.* The SARS-CoV-2 Delta variant induces an antibody response largely focused on class 1 and 2 antibody epitopes. *PLoS Pathog.* **18**, e1010592 (2022).
14. Greaney, A. J. *et al.* Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **29**, 463-476.e6 (2021).
15. Dejnirattisai, W. *et al.* SARS-CoV-2 Omicron-B.1.1.529 leads to widespread escape from neutralizing antibody responses. *Cell* **185**, 467-484.e15 (2022).
16. Cameroni, E. *et al.* Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* **602**, 664–670 (2022).
17. Starr, T. N. *et al.* Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **371**, 850–854 (2021).
18. Starr, T. N. *et al.* SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. *Nature* **597**, 97–102 (2021).
19. Chakraborty, C., Sharma, A. R., Bhattacharya, M. & Lee, S.-S. A detailed overview of immune escape, antibody escape, partial vaccine escape of SARS-CoV-2 and their emerging variants with escape mutations. *Front. Immunol.* **13**, 801522 (2022).
20. Starr, T. N. *et al.* Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **182**, 1295-1310.e20 (2020).
21. Mannar, D. *et al.* SARS-CoV-2 Omicron variant: Antibody evasion and cryo-EM structure of spike protein-ACE2 complex. *Science* **375**, 760–764 (2022).
22. McCallum, M. *et al.* Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement. *Science* **375**, 864–868 (2022).
23. Moulana, A. *et al.* Compensatory epistasis maintains ACE2 affinity in SARS-CoV-2 Omicron BA.1. *bioRxiv* (2022) doi:10.1101/2022.06.17.496635.

24. Starr, T. N. *et al.* Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution. *Science* **377**, 420–424 (2022).

25. Case, J. B. *et al.* Resilience of S309 and AZD7442 monoclonal antibody treatments against infection by SARS-CoV-2 Omicron lineage strains. *Nat. Commun.* **13**, 3824 (2022).

26. Adams, R. M., Mora, T., Walczak, A. M. & Kinney, J. B. Measuring the sequence-affinity landscape of antibodies with massively parallel titration curves. *Elife* **5**, (2016).

27. Phillips, A. M. *et al.* Binding affinity landscapes constrain the evolution of broadly neutralizing anti-influenza antibodies. *Elife* **10**, (2021).

28. Starr, T. N. *et al.* ACE2 binding is an ancestral and evolvable trait of sarbecoviruses. *Nature* **603**, 913–918 (2022).

29. Windsor, I. W. *et al.* Antibodies induced by an ancestral SARS-CoV-2 strain that cross-neutralize variants from Alpha to Omicron BA.1. *Sci. Immunol.* **7**, eabo3425 (2022).

30. Javanmardi, K. *et al.* Antibody escape and cryptic cross-domain stabilization in the SARS-CoV-2 Omicron spike protein. *Cell Host Microbe* (2022) doi:10.1016/j.chom.2022.07.016.

31. Steckbeck, J. D. *et al.* Kinetic rates of antibody binding correlate with neutralization sensitivity of variant simian immunodeficiency virus strains. *J. Virol.* **79**, 12311–12320 (2005).

32. Culp, T. D., Spatz, C. M., Reed, C. A. & Christensen, N. D. Binding and neutralization efficiencies of monoclonal antibodies, Fab fragments, and scFv specific for L1 epitopes on the capsid of infectious HPV particles. *Virology* **361**, 435–446 (2007).

33. Wentz, A. E. & Shusta, E. V. A novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins. *Appl. Environ. Microbiol.* **73**, 1189–1198 (2007).

34. Gietz, R. D. & Schiestl, R. H. Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 35–37 (2007).

35. Nguyen Ba, A. N. *et al.* High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. *Nature* **575**, 494–499 (2019).

36. Therneau, T., Atkinson, B., & Ripley, B. *Rpart: Recursive Partitioning*. (2013).

37. Borkovec, M. *et al. ggparty: "ggplot" Visualizations for the "partykit" Package*. (2019).

38. Wells, J. A. Additivity of mutational effects in proteins. *Biochemistry* **29**, 8509–8517 (1990).

39. Olson, C. A., Wu, N. C. & Sun, R. A comprehensive biophysical description of pairwise epistasis throughout an entire protein domain. *Curr. Biol.* **24**, 2643–2651 (2014).

40. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

41. Schrodinger, L. L. C. *The PyMOL Molecular Graphics System*. (2015).

42. Khare, S. *et al.* GISAID's role in pandemic response. *China CDC Wkly* **3**, 1049–1051 (2021).

43. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global Chall.* **1**, 33–46 (2017).

44. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Euro Surveill.* **22**, (2017).

45. Huerta-Cepas, J., Serra, F. & Bork, P. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–1638 (2016).

46. R Core Team. *R: A language and environment for statistical computing*. (2017).

47. Van Rossum, G. & Drake, F. L. *Python 3 Reference Manual*. (CreateSpace, 2009).

48. Wickham, H. *Ggplot2*. (Springer International Publishing, 2016).

49. Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **9**, 90–95 (2007).
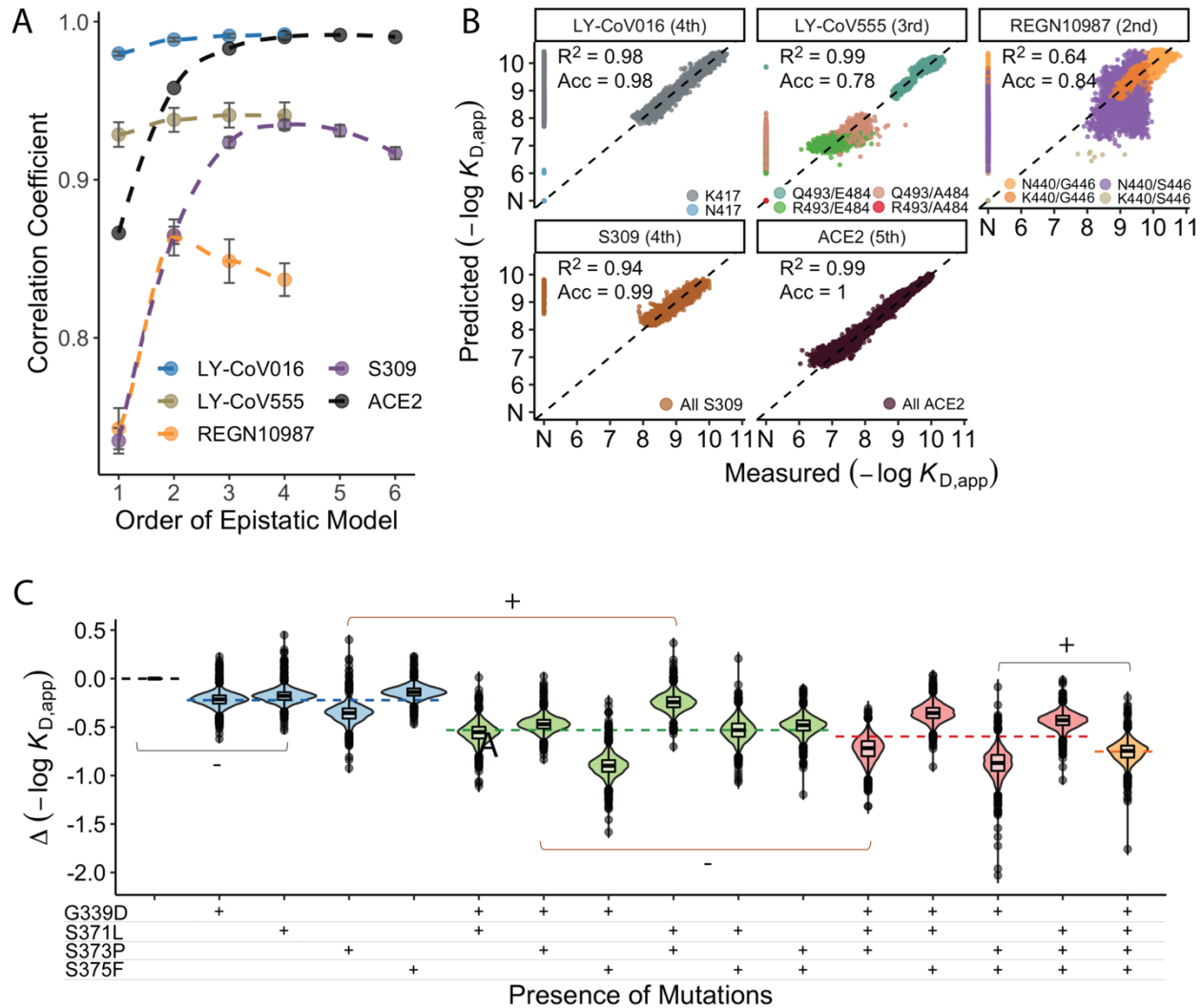
FIGURES
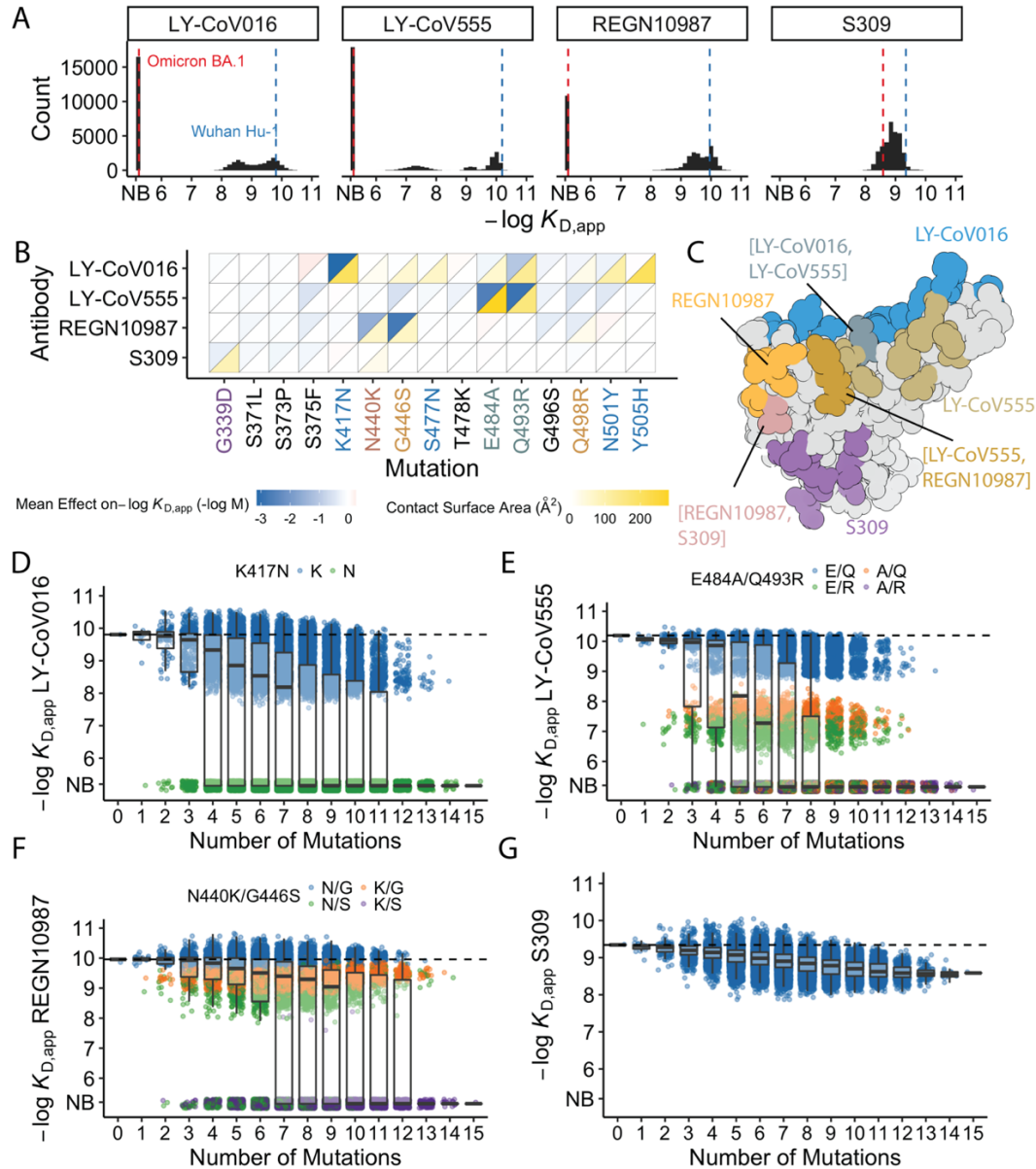
Figure 1. Antibody affinity landscape. (**A**) Binding affinities to one antibody from each class (LY-CoV016, LY-CoV555, REGN10987, and S309, from classes 1-4, respectively) across all N=32,768 RBD genotypes tested. Binding affinities are shown as -log$K_{D,app}$; vertical blue and red dashed lines indicate the -log$K_{D,app}$ for Wuhan Hu-1 and Omicron BA.1, respectively. 'NB' denotes non-binding (escape). (**B**) Mean effect of each mutation on antibody and ACE2 affinity (defined as the change in -log$K_{D,app}$ resulting from mutation) plotted with contact surface area between each residue and each antibody. Mutations are colored by footprint highlighted in (**C**). (**C**) Structure of SARS-CoV-2 BA.1 RBD with each antibody footprint annotated (PDB ID 7KMG, 6WPT, 7C01, 6XDG). Residues with overlapping footprints are colored and labeled accordingly. (**D-G**) Distribution of binding affinities to different antibodies grouped by number of Omicron BA.1 mutations. Binding affinity of the Wuhan Hu-1 variant is indicated by horizontal dashed lines. Variants with antibody escape mutations of interest are colored as noted in each key. NB denotes non-binding (escape).

16
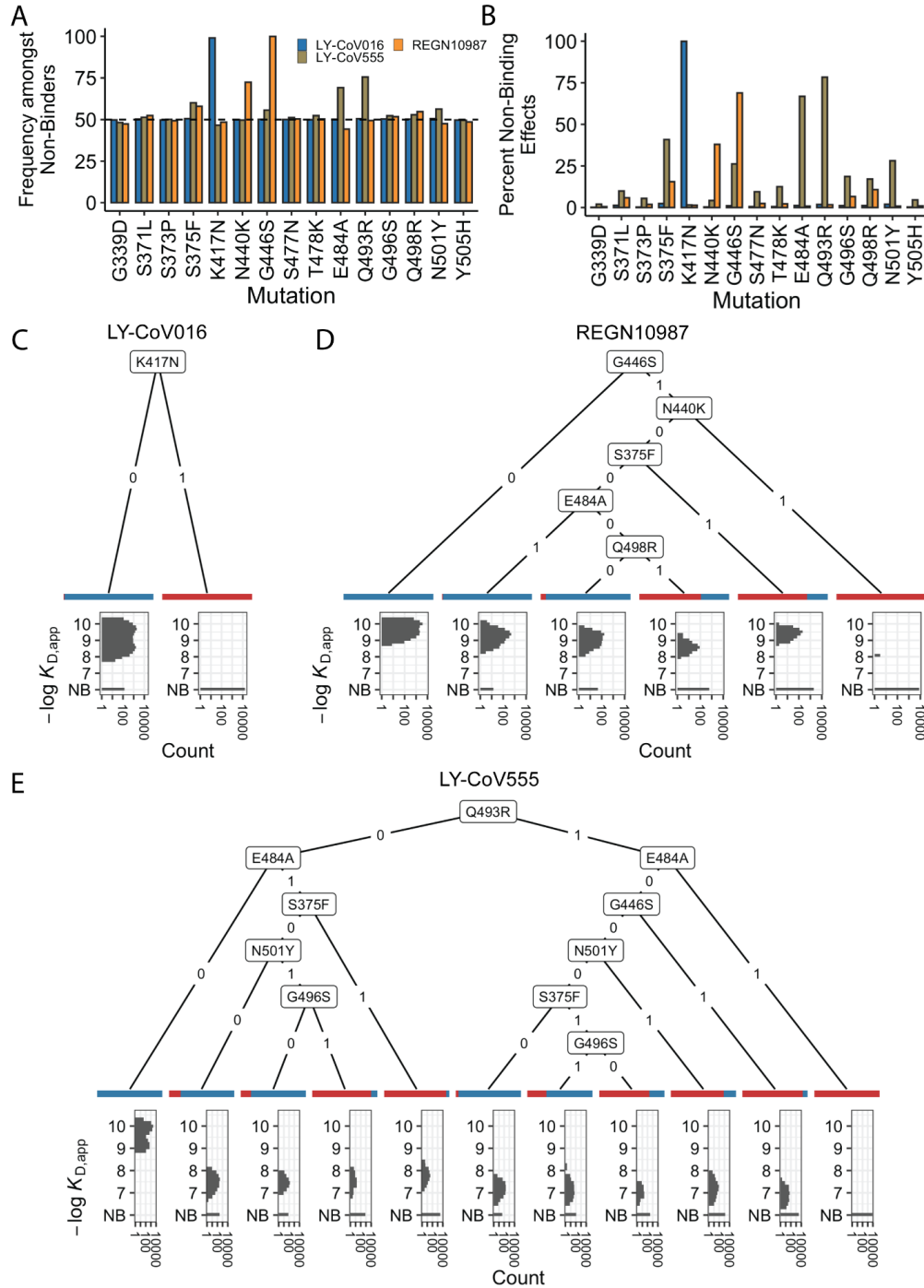
Figure 2. Escape mutations and genotypes. (**A**) Fraction of antibody-escaping genotypes with each mutation (**B**) Fraction of variants for which a given mutation confers antibody escape. Effects are colored as in (A). (**C-E**) Decision trees of escape phenotype for each antibody modeled as a function of the mutations present. Each leaf is annotated by the proportion of the genotypes that escape the corresponding antibody (red: escape, blue: does not) and by corresponding affinity distribution. NB denotes non-binding.

Figure 3. Epistastic effects on antibody binding. (**A**) Correlation coefficients between the measured values of -log($K_{D,app}$) and the model estimate for various orders of epistatic models. Correlations are computed on the hold-out subset averaged over 10 folds of cross-validation. zoomed-in version for orders 3 to 6. (**B**) Binding affinities predicted by complete coefficients from the optimum epistasis model are compared to the measured binding affinities for each antibody. Points are colored by mutations present in the genotypes, 'N' corresponds to nonbinding genotypes. The accuracy measures the quality of the binary classification between binders and non-binders and the coefficient of determination R^2 refer to the correlation between inferred and measured binding affinities, excluding non-binders. (**C**) Effects of mutations G339D, S371L, S373P, and S375F on S309 affinity grouped by the presence of each mutation. Each violin color corresponds to the number of mutations considered. Dashed line color denotes the average effect for each group represented by the violin color.
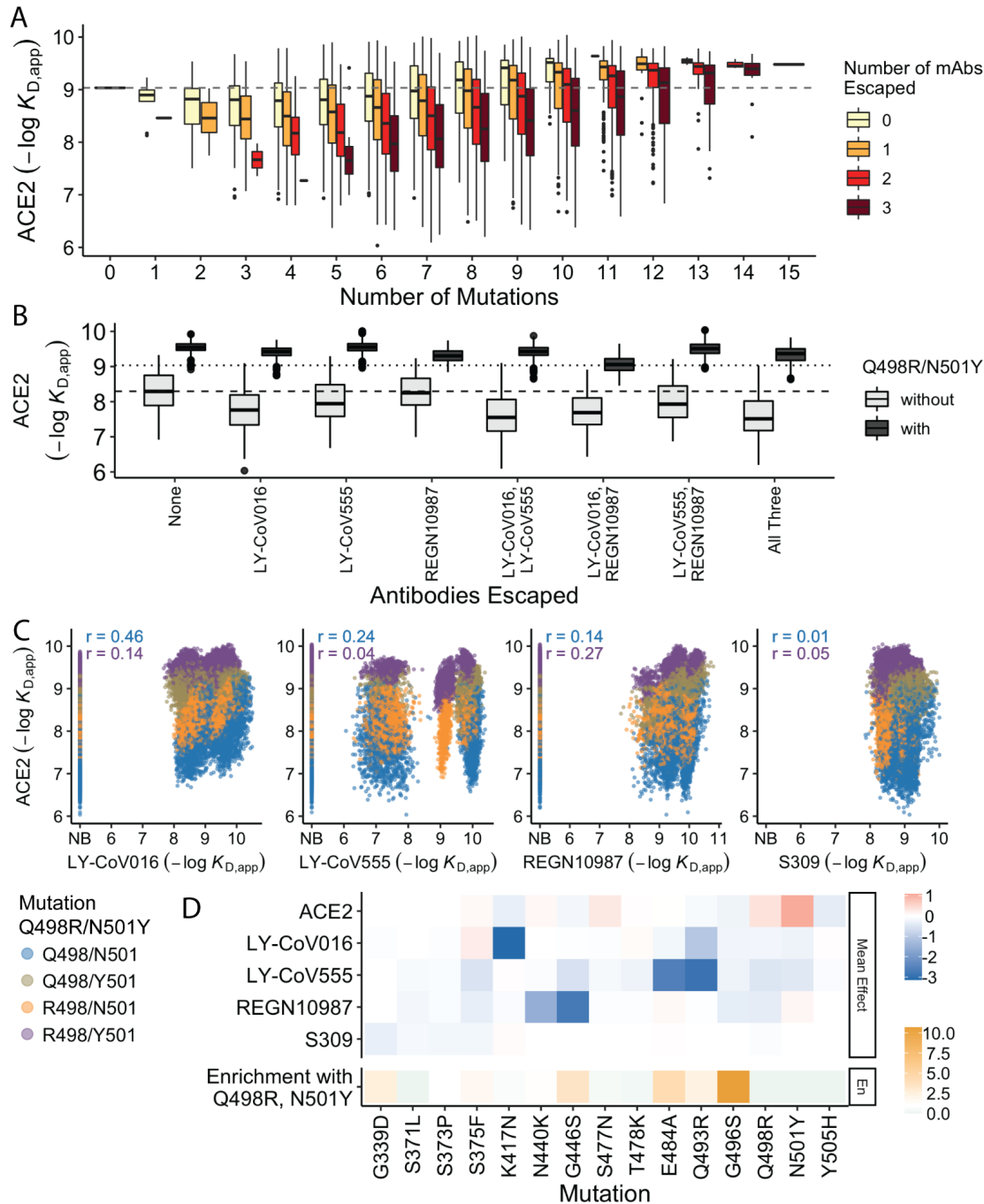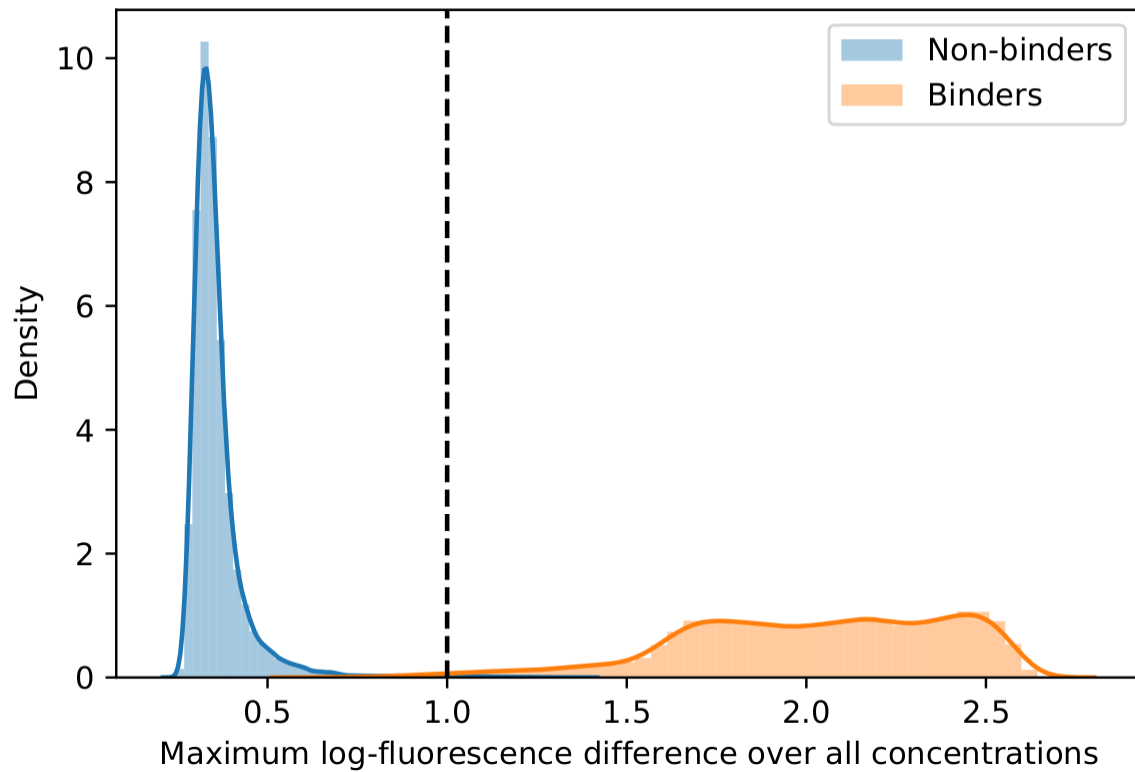
Figure 4. Trade-offs and comparison with ACE2 affinity. (**A**) Distribution of ACE2 binding affinity grouped by number of BA.1 mutations and the number of monoclonal antibodies escaped. (**B**) ACE2 affinity distribution grouped by antibodies escaped and the presence of compensatory mutations (N501Y and Q498R). (**C**) Affinities to monoclonal antibodies plotted as a function of the ACE2 affinity for all genotypes. Points are colored by presence of Q498R and N501Y. **(D)** Mean effect of each mutation on antibody affinity and on ACE2 affinity (red-blue colormap) compared to the enrichment of their frequency with Q498R and N501Y (orange colormap). The enrichment score is defined as the normalized frequency a mutation emerged on a branch on which mutations Q498R and N501Y appear divided by the normalized frequency it emerged on any intermediate background between Wuhan and BA.1.

19
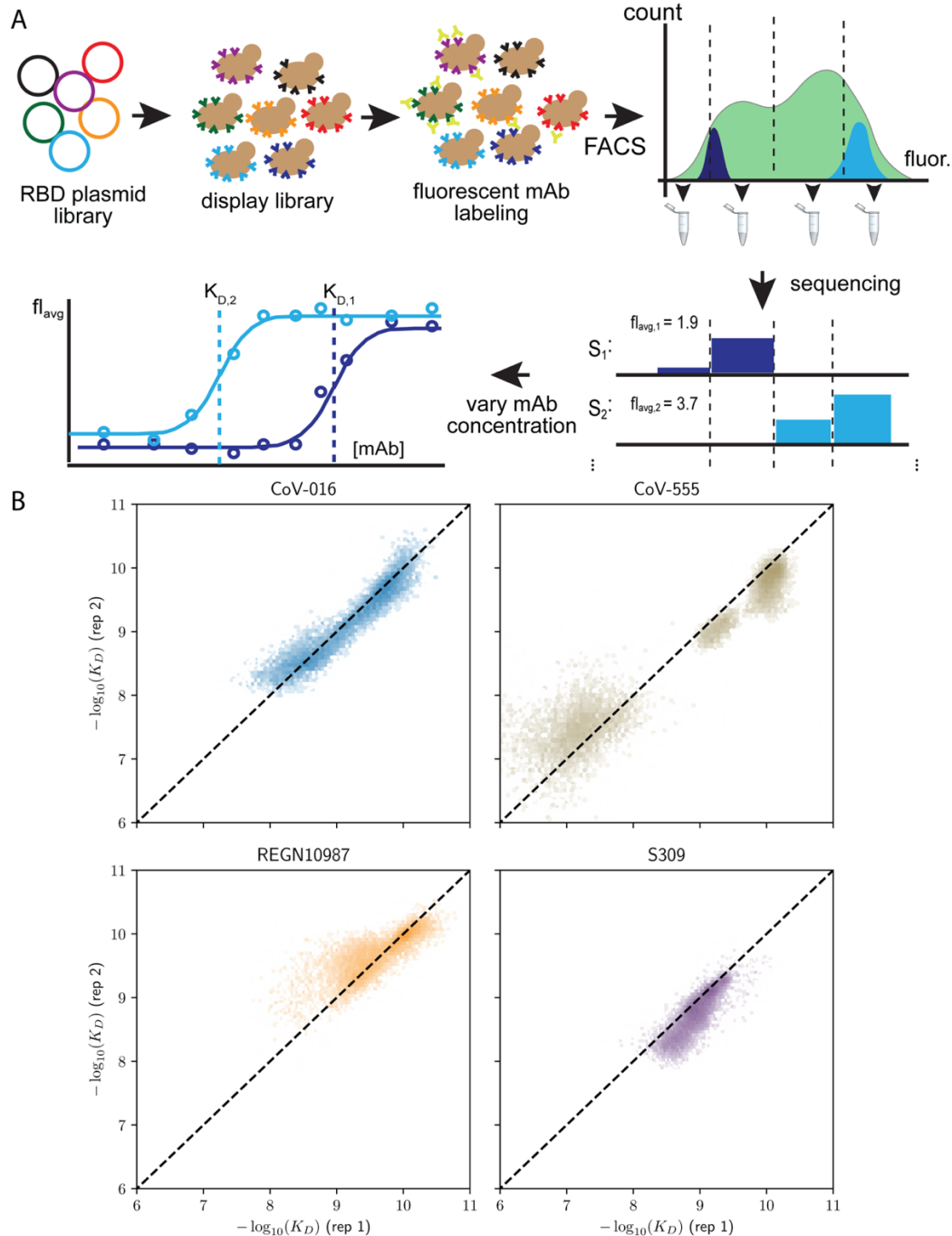
Supplementary Figure S1. Schematic overview of the Tite-Seq method and reproducibility of dissociation constants. (**A**) The plasmid library containing RBD sequences is transformed into a standard yeast display strain AWY101. The library is then incubated with soluble, fluorescent monoclonal antibody (either one of the four mAbs used in the study) and sorted by flow cytometry into bins based on mAb fluorescence. Deep sequencing of each bin results in the mean bin (B_{avg}) estimate of each RBD variant across varying mAb concentrations to produce a titration curve. Apparent equilibrium dissociation constants are them inferred by fitting B_{avg} to the mAb concentration. (**B**) Correlation of $-\log(K_{D,app})$ between the first and second biological replicates.

Supplementary Figure S2. Epistatic coefficients. (**A**) First-order effects in fourth order epistatic interaction models. (**B**) For each antibody, higher-order interaction coefficient of a mutation (shown at bottom of heat map plot) is calculated by summing over all third- and fourth-order interaction coefficients involving the mutation.

Supplementary Figure S3. Distribution of maximum log-fluorescence difference. To determine whether a genotype is a non-binder, the maximum difference of log fluorescence across concentrations of the genotype is computed. The dashed line represents the threshold below which a genotype with a certain maximum difference is considered non-binding.

Strain	Antibody	Isogenic -log $K_{D,app}$	TiteSeq -log $K_{D,app}$
Omicron BA.1	LY-CoV016	NB	NB
Omicron BA.1	LY-CoV555	NB	NB
Omicron BA.1	REGN10987	NB	NB
Omicron BA.1	S309	8.81	8.59 ± 0.16
Wuhan Hu-1	LY-CoV016	10.52 ± 0.24	9.80 ± 0.02
Wuhan Hu-1	LY-CoV555	10.01 ± 0.33	10.19 ± 0.24
Wuhan Hu-1	REGN10987	10.42 ± 0.49	9.96 ± 0.05
Wuhan Hu-1	S309	9.33 ± 0.27	9.34 ± 0.09

Supplementary Table S1. Isogenic validation of binding affinities. The $K_{D,app}$ inferred from Isogenic measurements (see Methods) shown with those inferred via Tite-Seq measurement. NB denotes non-binding and standard deviations between replicates are also shown.