

1 An engineered decoy receptor for SARS-CoV-2 broadly binds 2 protein S sequence variants

3 Kui K. Chan¹, Timothy J.C. Tan², Krishna K. Narayanan² and Erik Procko²

4 ¹Orthogonal Biologics, Champaign IL 61821

5 ²Department of Biochemistry and Cancer Center at Illinois, University of Illinois, Urbana IL 61801

6 Correspondence: procko@illinois.edu

7 ABSTRACT

8 The spike S of SARS-CoV-2 recognizes ACE2 on the host cell membrane to initiate entry. Soluble decoy
9 receptors, in which the ACE2 ectodomain is engineered to block S with high affinity, potentially neutralize
10 infection and, due to close similarity with the natural receptor, hold out the promise of being broadly
11 active against virus variants without opportunity for escape. Here, we directly test this hypothesis. We
12 find an engineered decoy receptor, sACE2_{v2.4}, tightly binds S of SARS-associated viruses from
13 humans and bats, despite the ACE2-binding surface being a region of high diversity. Saturation
14 mutagenesis of the receptor-binding domain (RBD) followed by in vitro selection, with wild type ACE2
15 and the engineered decoy competing for binding sites, failed to find S mutants that discriminate in favor
16 of the wild type receptor. Variant N501Y in the RBD, which has emerged in a rapidly spreading lineage
17 (B.1.1.7) in England, enhances affinity for wild type ACE2 20-fold but remains tightly bound to
18 engineered sACE2_{v2.4}. We conclude that resistance to engineered decoys will be rare and that decoys
19 may be active against future outbreaks of SARS-associated betacoronaviruses.

20 INTRODUCTION

21 Zoonotic coronaviruses have crossed over from animal reservoirs multiple times in the past two
22 decades, and it is almost certain that wild animals will continue to be a source of devastating outbreaks.
23 Unlike ubiquitous human coronaviruses responsible for common respiratory illnesses, these zoonotic
24 coronaviruses with pandemic potential cause serious and complex diseases, in part due to their tissue
25 tropisms driven by receptor usage. Severe Acute Respiratory Syndrome Coronaviruses 1 (SARS-CoV-1)
26 and 2 (SARS-CoV-2) engage angiotensin-converting enzyme 2 (ACE2) for cell attachment and entry (*1-
27 7*). ACE2 is a protease responsible for regulating blood volume and pressure that is expressed on the
28 surface of cells in the lung, heart and gastrointestinal tract, among other tissues (*8, 9*). The ongoing spread
29 of SARS-CoV-2 and the disease it causes, COVID-19, has had a crippling toll on global healthcare
30 systems and economies, and effective treatments and vaccines are urgently needed.

31 As SARS-CoV-2 becomes endemic in the human population, it has the potential to mutate and
32 undergo genetic drift and recombination. To what extent this will occur as increasing numbers of people
33 are infected and mount counter immune responses is unknown, but already a variant in the viral spike
34 protein S (D614G) has rapidly emerged from multiple independent events and affects S protein stability
35 and dynamics (*10, 11*). Another S variant (D839Y) became prevalent in Portugal, possibly due to a
36 founder effect (*12*). SARS-CoV-2 has a moderate mutation rate estimated at 10^{-3} substitutions per site per
37 year (*13*). However, the virus has undergone rapid mutation and adaptation after infecting mink in
38 Denmark, from which it then crossed back to humans (*14*), causing Danish authorities to order 17 million
39 farmed mink culled to pre-emptively prevent the possible emergence of vaccine-resistant variants.
40 Additionally, large changes in coronavirus genomes have frequently occurred in nature from
41 recombination events, especially in bats where co-infection levels can be high (*15, 16*). Recombination of
42 MERS-CoVs has been documented in camels (*17*), there are reported cases of recombination between co-
43 circulating SARS-CoV-2 variants (*18*), and SARS-CoV-2 itself may have emerged through

44 recombination of coronavirus genomes (19). This will all have profound implications for the current
45 pandemic's trajectory, the potential for future coronavirus pandemics, and whether drug or vaccine
46 resistance in SARS-CoV-2 emerges and becomes widespread.

47 The viral spike is a vulnerable target for neutralizing monoclonal antibodies that are progressing
48 through clinical trials, yet in tissue culture escape mutations in the spike rapidly emerge to all antibodies
49 tested (20). Deep mutagenesis of the isolated receptor-binding domain (RBD) by yeast surface display
50 has easily identified mutations in S that retain high expression and ACE2 affinity, yet are no longer bound
51 by monoclonal antibodies and confer resistance (21). This has motivated the development of cocktails of
52 non-competing monoclonals (20, 22), inspired by lessons learned from the treatment of HIV-1 and Ebola,
53 to limit the possibilities for the virus to escape. Notably, drug maker Eli Lilly has a monoclonal
54 monotherapy (LY-CoV555) in advanced trials (NCT04427501) where the selection of resistant virus
55 variants in patients has occurred. A trial update added an arm with a second monoclonal (LY-CoV016)
56 and the company has not reported putative resistance variants in patients receiving the cocktail thus far.
57 However, even the use of monoclonal cocktails does not address future coronavirus spill overs from wild
58 animals that may be antigenically distinct. Indeed, large screening efforts were required to find
59 antibodies from recovered SARS-CoV-1 patients that cross-react with SARS-CoV-2 (23), indicating
60 antibodies have confined capacity for interacting with variable epitopes on the spike surface, and are
61 unlikely to be broad and pan-specific for all SARS-related viruses.

62 An alternative protein-based antiviral to monoclonal antibodies is to use soluble ACE2 (sACE2) as a
63 decoy to compete for receptor-binding sites on the viral spike (6, 24-27). In principle, the virus has
64 limited potential to escape sACE2-mediated neutralization without simultaneously decreasing affinity for
65 the native ACE2 receptor, rendering the virus less virulent. Wild type sACE2 is currently in a phase II
66 clinical trial (28) and multiple groups have now engineered sACE2 to create high affinity decoys for
67 SARS-CoV-2 that rival matured monoclonal antibodies for potent neutralization of infection (27, 29, 30).
68 In our group, deep mutagenesis was used to identify a large number of mutations in ACE2 that increase
69 affinity for S (27). These mutations were dispersed across the interface and also at distal sites where they
70 are predicted to enhance folding of the virus-recognized conformation. A combination of three mutations,
71 called sACE2_{2.v2.4}, increases affinity 35-fold and binds SARS-CoV-2 S (K_D 600 pM) with affinity
72 comparable to the best monoclonal antibodies (27). Even tighter apparent affinities are reached through
73 avid binding to trimeric spike expressed on a membrane. Despite engineering being focused exclusively
74 on SARS-CoV-2 affinity, sACE2_{2.v2.4} potently neutralized authentic SARS-CoV-1 and -2 infection in
75 tissue culture, suggesting it's close resemblance to the wild type receptor confers broad activity against
76 ACE2-utilizing betacoronaviruses generally. Soluble ACE2_{2.v2.4} is dimeric and monodisperse without
77 aggregation, catalytically active, highly soluble, stable after storage at 37°C for days, and well expressed
78 at levels greater than the wild type protein. Due to both its high activity and favorable properties for
79 manufacture, sACE2_{2.v2.4} is a genuine drug candidate for preclinical development.

80 Engineered, high affinity decoy receptors, while very similar to natural ACE2, nonetheless have
81 mutations present at or near the interaction surface. There is therefore an opportunity for viral spike
82 variants to discriminate between an engineered decoy and wild type receptors, providing a route towards
83 resistance. Here, we show that the engineered decoy sACE2_{2.v2.4} binds broadly and tightly to the RBDs
84 of diverse SARS-associated betacoronaviruses that use ACE2 for entry. We further fail to find mutations
85 within the RBD, which directly contacts ACE2 and is where possible escape mutations will most likely
86 reside, that redirect specificity towards the wild type receptor. We conclude that resistance to an
87 engineered decoy receptor will be rare, and sACE2_{2.v2.4} targets common attributes for affinity to S in
88 SARS-associated viruses.

89 RESULTS

90 **An engineered decoy receptor broadly binds RBDs from SARS-associated CoVs with tight affinity**

91 The affinities of the decoy receptor sACE2₂.v2.4 were determined for purified RBDs from the S
92 proteins of five coronaviruses from *Rhinolophus* bat species (isolates LYRa11, Rs4231, Rs7327, Rs4084
93 and RsSHC014) and two human coronaviruses, SARS-CoV-1 and SARS-CoV-2. These viruses fall
94 within a common clade of betacoronaviruses that have been experimentally validated to use human ACE2
95 as an entry receptor (7). They share close sequence identity within the RBD core while variation is
96 highest within the functional ACE2 binding site (Figures 1 and S1), possibly due to a co-evolutionary
97 'arms race' with polymorphic ACE2 sequences in ecologically diverse bat species (31). Affinity was
98 measured by biolayer interferometry (BLI), with sACE2₂ (a.a. S19-G732) fused at the C-terminus with
99 the Fc moiety of human IgG1 immobilized to the sensor surface and monomeric 8his-tagged RBD (Figure
100 S2) used as the soluble analyte. This arrangement excludes avidity effects, which otherwise cause
101 artificially tight (picomolar) apparent affinities whenever dimeric sACE2₂ in solution is bound to
102 immobilized RBD decorating an interaction surface. Wild type sACE2₂ bound all the RBDs with
103 affinities ranging from 16 nM for SARS-CoV-2 to 91 nM for LYRa11, with median affinity 60 nM
104 (Table 1). The measured affinities for the RBDs of SARS-CoV-1 and SARS-CoV-2 are comparable to
105 published data (4, 27, 32-34). Engineered sACE2₂.v2.4 displayed large increases in affinity for all the
106 RBDs, with K_Ds ranging from 0.4 nM for SARS-CoV-2 to 3.5 nM for isolate Rs4231, with median
107 affinity less than 2 nM (Table 1). The approximate 35-fold affinity increase of the engineered decoy
108 applies universally to coronaviruses in the test panel and the molecular basis for affinity enhancement
109 must therefore be grounded in common attributes of RBD/ACE2 recognition.

110 **A deep mutational scan of the RBD in the context of full-length S reveals residues in the ACE2** 111 **binding site are mutationally tolerant**

112 To explore potential sequence diversity in S of SARS-CoV-2 that may act as a 'reservoir' for drug
113 resistance, the mutational tolerance of the RBD was evaluated by deep mutagenesis (35). Saturation
114 mutagenesis was focused to the RBD (a.a. C336-L517) of full-length S tagged at the extracellular N-
115 terminus with a c-myc epitope for detection of surface expression. The spike library, encompassing 3,640
116 single amino acid substitutions, was transfected in human Expi293F cells under conditions where cells
117 typically acquire no more than a single sequence variant (36, 37). The culture was incubated with wild
118 type, 8his-tagged, dimeric sACE2₂ at a sub-saturating concentration (2.5 nM). Bound sACE2₂-8h and
119 surface-expressed S were stained with fluorescent antibodies for flow cytometry analysis (Figure 2A).
120 Compared to cells expressing wild type S, the library was poorly expressed, indicating many mutations
121 are deleterious for folding and expression. A cell population was clearly discernable expressing S
122 variants that bind ACE2 with decreased affinity (Figure 2B). After gating for c-myc-positive cells
123 expressing S, cells with high and low levels of bound sACE2₂ were collected by fluorescence-activated
124 cell sorting (FACS), called the ACE2-High and ACE2-Low populations, respectively (Figure 2C). Both
125 the expression and sACE2₂ binding signals decreased over minutes to hours during sorting, possibly due
126 to shedding of the S1 subunit. Cells were therefore collected and pooled from three separate FACS
127 experiments for a combined 8 hours sort time.

128 Transcripts in the sorted cells were Illumina sequenced and compared to the naive plasmid library to
129 determine an enrichment ratio for each amino acid substitution (38). Mutations in S that express and bind
130 ACE2 tightly are selectively enriched in the ACE2-High sort (Figure S3); mutations that express but have
131 reduced ACE2 binding are selectively enriched in the ACE2-Low sort; and mutations that are poorly
132 expressed are depleted from both sorted populations. Positional conservation scores were calculated by
133 averaging the log₂ enrichment ratios for each of the possible amino acids at a residue position. By adding
134 conservation scores for both the ACE2-High and ACE2-Low sorts we derive a score for surface
135 expression, which shows that the hydrophobic RBD core is tightly conserved for folding and trafficking

136 of the viral spike (Figure 3A). By comparison, residues on the exposed RBD surface are mutationally
137 permissive for S surface expression. This matches the mutational tolerance of proteins generally.

138 For tight ACE2 binding (i.e. S variants in the ACE2-High population), conservation increases for
139 RBD residues at the ACE2 interface, yet mutational tolerance remains high (Figure 3C). The sequence
140 diversity observed among natural betacoronaviruses, which display high diversity at the ACE2 binding
141 site, is therefore replicated in the deep mutational scan, which predicts the SARS-CoV-2 spike tolerates
142 substantial genetic diversity at the receptor-binding site for function. From this accessible sequence
143 diversity SARS-CoV-2 might feasibly mutate to acquire resistance to monoclonal antibodies or
144 engineered decoy receptors targeting the ACE2-binding site.

145 There are two ‘hot spot’ regions for interactions at the interface that determine receptor affinity and
146 species adaptation, centered around ACE2 residues K31 and K353 (39). These regions are also the
147 locations for substitutions to ACE2 residues T27, L79, and N330 in the engineered sACE2₂.v2.4 decoy.
148 Mutations to RBD residues at these sites tend to be weakly depleted for high ACE2 binding (Figure S4),
149 but are much more tolerant of mutations than structural positions buried in the RBD core. To highlight a
150 few residues, S-Y505 that packs against the hydrocarbon chain of ACE2-K353 is notably more conserved
151 than most other interfacial residues, with the exceptions of partially buried residues like S-R403 and S-
152 Y453 that likely have additional structural roles; S-Y489 contacting ACE2-T27 and ACE2-K31 has an
153 overall weak preference for aromatic amino acids; and S-G485 on a loop packed against ACE2-L79 has
154 highest tolerance for polar substitutions, possibly to maintain the loop conformation and solubility (Figure
155 S4). Some mutations are found to be highly enriched for ACE2 binding, including small hydrophobic
156 amino acids for S-Q493 and aromatic amino acids for S-N501 that are both anticipated to increase local
157 hydrophobic or aromatic ring packing. This is consistent with observations from yeast surface display of
158 RBD mutants (40). That these mutations predicted to increase ACE2 binding are not enriched in
159 circulating SARS-CoV-2 variants suggests the affinity of the virus for its receptor is already sufficient for
160 high transmission and peak fitness (40).

161 **Comparison to a deep mutational scan of the isolated RBD by yeast surface display**

162 Two deep mutational scans have been reported for the isolated RBD displayed on the surface of
163 yeast (40, 41). We compare our data, from a selection of full-length S expressed in human cells, to the
164 publicly accessible Starr et al data set (40). Important residues within the RBD for surface expression of
165 full-length spike in human cells are closely correlated with data from yeast surface display of the isolated
166 RBD (Figure 3B), with the exception of a notable region. The surface of the RBD opposing the ACE2-
167 binding site (e.g. V362, Y365, and C391) is free to mutate for yeast surface display, but its sequence is
168 constrained in our experiments; this region of the RBD is buried by connecting structural elements to the
169 global fold of an S subunit in the closed-down conformation (this is the dominant conformation for S
170 subunits and is inaccessible to receptor binding; Figure S5) (2, 4, 42, 43). While some mutations might
171 have allosteric effects, we note that among substitutions of V362, Y365, and C391, mutations tend to be
172 lightly depleted from the ACE2-Low and heavily depleted from the ACE2-High sorted populations. No
173 mutations at these positions were selectively enriched in just the ACE2-High sort, as might be expected if
174 a mutation favored the open-up conformation through allosteric mechanisms. This is consistent with
175 mutations at these residues reducing ACE2 interactions through defects in folding and decreased surface
176 expression. Compared to single mutations that destabilize the RBD in the closed-down conformation,
177 more extensive engineering with multiple stabilizing mutations has been shown to shift the
178 conformational equilibrium of S subunits to the open-up state (44).

179 We used targeted mutagenesis to individually test alanine substitutions to all the cysteines in the
180 RBD (Figure S5). We found all cysteine-to-alanine mutations severely diminish S surface expression in

181 Expi293F cells, including C391A and C525A on the RBD 'backside' that were neutral in the yeast display
182 scan (40). These differences demonstrate that there are tighter sequence constraints on the RBD in the
183 context of a full spike expressed at a human cell membrane, yet overall we consider the yeast display and
184 the human cell data sets to closely agree.

185 For binding to dimeric sACE2₂, we note that interface residues were more tightly conserved in the
186 Starr et al data set (Figure 3D), possibly a consequence of three differences between the deep mutagenesis
187 experiments. First, our selections for ACE2 binding of S variants at the plasma membrane appears to
188 primarily reflect mutational effects on surface expression, which is almost certainly more stringent in
189 human cells. Yeast permit many poorly folded proteins to leak to the cell surface (45). Second, the yeast
190 selections were conducted at multiple sACE2 concentrations from which apparent K_D changes were
191 computed (40); the Starr et al data in this regard is very comprehensive. Due to the long sort times
192 required for our human cell libraries where only a small fraction of cells express spike, we sorted at a
193 single sACE2₂ concentration that cannot accurately capture a range of different binding affinities
194 quantitatively. Third, dimeric sACE2₂ may geometrically complement trimeric S densely packed on a
195 human cell membrane, such that avidity masks the effects of affinity-reducing mutations. Nonetheless,
196 there is overall agreement that ACE2 binding often persists following mutations to the RBD surface, and
197 our data simply suggests mutational tolerance may be even greater than that already observed by Starr et
198 al.

199 **A screen for S variants that preferentially bind wild type ACE2 over the engineered decoy**

200 Having shown that the ACE2-binding site of SARS-CoV-2 protein S tolerates many mutations, we
201 asked whether mutations might therefore be found that confer resistance to the engineered decoy
202 sACE2₂.v2.4. Resistance mutations are anticipated to lose affinity for sACE2₂.v2.4 while maintaining
203 binding to the wild type receptor, and are most likely to reside in the RBD where physical contacts are
204 made. Similar reasoning formed the foundation of a deep mutagenesis-based selection of the isolated
205 RBD by yeast surface display to find escape mutations to monoclonal antibodies, and the results were
206 predictive of escape mutations in pseudovirus growth selections (21).

207 To address whether escape mutations from the engineered decoy might be found in the RBD, we
208 repurposed the S protein library for a specificity selection. Cells expressing the library, encoding all
209 possible substitutions in the RBD, were co-incubated with wild type sACE2₂ fused to the Fc region of
210 IgG1 and 8his-tagged sACE2₂.v2.4 at concentrations where both proteins bind competitively (27). It was
211 immediately apparent from flow cytometry of the Expi293F culture expressing the S library that there
212 were cells expressing S variants shifted towards preferential binding to sACE2₂.v2.4, but no significant
213 population with preferential binding to the wild type receptor (Figures 4A and 4B). Cells expressing S
214 variants that might preferentially bind sACE2₂(WT)-IgG1 or sACE2₂.v2.4 were gated and collected by
215 FACS (Figure 4C), followed by deep sequencing of S transcripts to determine enrichment ratios. There
216 was close agreement between two independent replicate experiments (Figures 4D-4G). Most RBD
217 mutations were depleted following sorting, consistent with deleterious effects on S folding and
218 expression.

219 Soluble ACE2₂.v2.4 has three mutations from wild type ACE2: T27Y buried within the RBD
220 interface, and L79T and N330Y at the interface periphery (Figure 5A). A substantial number of mutations
221 in the RBD of S were selectively enriched for preferential binding to sACE2₂.v2.4 (Figure 5B, upper-left
222 quadrant). While sACE2₂.v2.4-specificity mutations could be found immediately adjacent to the sites of
223 engineered mutations in ACE2 (in particular mutations to S-F486 adjacent to ACE2-L79 and S-T500
224 adjacent to ACE2-N330), major hot spots for sACE2₂.v2.4-specificity mutations were also mapped to
225 RBD loop 498-506, contacting the region where the ACE2- α 1 helix packs against a β -hairpin motif
226 (Figure 5A). By comparison, there were no hot spots in the RBD for sACE2₂(WT)-specificity mutations.

227 Indeed, only a small number of mutations were selectively enriched for preferential binding to wild type
228 receptor (Figure 5B), and the abundance of these putative wild type-specific mutations barely rose above
229 the expected level of noise in the deep mutagenesis data. In this competition assay, S binding to wild type
230 sACE2₂ is therefore more sensitive to RBD mutations than S binding to engineered sACE2₂.v2.4.

231 To determine whether the potential wild type ACE2-specific mutations found by deep mutagenesis
232 are real as opposed to false predictions due to data noise, we tested 24 mutants of S selectively enriched in
233 the wild type-specific gate by targeted mutagenesis (blue data points in Figure 5B). Only minor shifts
234 towards binding wild type sACE2₂ were observed (Figure S6). Two S mutants were investigated further
235 in sACE2₂ titration experiments, N501W and N501Y, which both retained high receptor binding and
236 displayed small shifts towards wild type sACE2₂ in the competition experiment. N501 of S is located in
237 the 498-506 loop and its substitution to large aromatic side chains might alter the loop conformation to
238 cause steric strain with nearby ACE2 mutation N330Y in sACE2₂.v2.4. After titrating the concentrations
239 of 8his-tagged sACE2₂(WT) and sACE2₂.v2.4 and measuring bound protein to S-expressing cells by flow
240 cytometry, it was found S-N501W and S-N501Y do show enhanced specificity for wild type sACE2₂, but
241 the effect is small and sACE2₂.v2.4 remains the stronger binder (Figure 5C).

242 Dimeric sACE2₂ binds avidly to S protein on a membrane surface; avid interactions are also
243 observed between sACE2₂ and spikes on authentic SARS-CoV-2 in infection assays (27). We used BLI
244 kinetics measurements, in which immobilized sACE2₂-IgG1 interacts with monomeric RBD, to determine
245 how the observed changes in avid sACE2₂ binding to S-expressing cells translate to changes in affinity.
246 Both N501W and N501Y mutants of SARS-CoV-2 RBD displayed increased affinity for wild type ACE2
247 and engineered ACE2.v2.4, with larger affinity gains in favor of the wild type receptor (Table 1). This
248 aligns with the flow cytometry data indicating a small shift in specificity towards wild type ACE2, but not
249 enough to escape the engineered decoy. By comparison, multiple independent escape mutations are
250 readily found in S of SARS-CoV-2 that diminish the efficacy of monoclonal antibodies by many orders of
251 magnitude (20, 21).

252 Finally, 8 representative mutations to S predicted from the deep mutational scan to increase
253 specificity towards sACE2₂.v2.4 (purple data points in Figure 5B) were cloned and 7 were found to have
254 large shifts towards preferential sACE2₂.v2.4 binding in the competition assay (Figure S7). These S
255 mutations were Y449K/Q/S, L455G/R/Y, and G504K. The basis for why the mutations increase
256 specificity towards engineered sACE2₂.v2.4 is ambiguous, since RBD residues Y449, L455, and G504
257 are not in direct contact with engineered sites of the receptor. BLI kinetics between immobilized sACE2₂-
258 IgG1 and monomeric RBD as the analyte showed reduced affinity of a representative mutant, RBD-
259 Y449K, to both wild type and engineered sACE2₂ (Table 1). However, affinity changes in the picomolar
260 range for sACE2₂.v2.4 are hidden during avid binding to full-length S-Y449K at the cell surface, whereas
261 avid binding of wild type sACE2₂ to S-Y449K (with affinity measured by BLI in the moderate nanomolar
262 range) is substantially reduced. This finding might explain why the competition selection found many
263 mutations that shift specificity towards engineered ACE2, as mutations causing small decreases in affinity
264 may have larger effects on avid binding of the weaker-bound wild type receptor.

265 Overall, validation by targeted mutagenesis confirms that the selection can successfully find
266 mutations in S with altered specificity. The inability to find mutations in the RBD that impart high
267 specificity for the wild type receptor means such mutations are rare or may not even exist, at least within
268 the receptor-binding domain where direct physical contacts with receptors occur. We cannot exclude
269 mutations elsewhere having long-range conformational effects. Engineered, soluble decoy receptors
270 therefore live up to their promise as broad therapeutic candidates against which a virus cannot easily
271 escape.

272 DISCUSSION

273 The allure of soluble decoy receptors is that the virus cannot easily mutate to escape neutralization.
274 Mutations that reduce affinity of the soluble decoy will likely also decrease affinity for the wild type
275 receptor on host cells, thereby coming at the cost of diminished infectivity and virulence. However, this
276 hypothesis has not been rigorously tested, and since engineered decoy receptors differ from their wild
277 type counterparts, even if by just a small number of mutations, it is possible a virus may evolve to
278 discriminate between the two. Here, we show that an engineered decoy receptor for SARS-CoV-2
279 broadly binds with low nanomolar K_D to the spikes of SARS-associated betacoronaviruses that use ACE2
280 for entry, despite high sequence diversity within the ACE2-binding site. Mutations in S of SARS-CoV-2
281 that confer high specificity for wild type ACE2 were not found in a comprehensive screen of all
282 substitutions within the RBD. The engineered decoy receptor is therefore broad against zoonotic ACE2-
283 utilizing coronaviruses that may spill over from animal reservoirs in the future and against variants of
284 SARS-CoV-2 that may arise as the current COVID-19 pandemic rages on. These findings are highly
285 consistent with research of a high-affinity decoy receptor engineered for a different pathogen, the human
286 immunodeficiency virus (HIV). An IgG1 Fc-fused soluble decoy based on HIV receptors, called eCD4-
287 Ig, broadly neutralizes HIV-1, HIV-2, and related simian viruses, with single mutations in the HIV spike
288 protein unable to achieve full escape (46, 47). Together with the results reported here, these studies
289 collectively demonstrate that engineered decoy receptors can achieve exceptional breadth against virus
290 sequence variants. We argue it is unlikely that decoy receptors will need to be combined in cocktail
291 formulations, as is required for many monoclonal antibodies and possibly designed miniprotein binders to
292 prevent the rapid emergence of resistance (20, 48), facilitating manufacture and distribution. Our findings
293 give insight into how a potential therapeutic can achieve breadth with a low chance of virus resistance for
294 a family of highly infectious and deadly viruses.

295 MATERIALS AND METHODS

296 **Plasmids.** Residue numbers for constructs begin from the start methionine as amino acid (a.a.) 1. The
297 cloning of human codon-optimized, mature S from SARS-CoV-2 (GenBank Acc. No. YP_009724390.1;
298 a.a. V16-T1273) into the NheI-XhoI sites of pCEP4 (Invitrogen) with an N-terminal, extracellular c-myc
299 tag is described elsewhere (27). Soluble ACE2 (a.a. 1-732 encoding a dimer; wild type or engineered
300 variant sACE2_{v2.4}) fused to an 8his purification tag or to human IgG1-Fc (a.a. D221-K447; nG1m1
301 isoallotype; GenBank KY432415.1) and cloned in to the NheI-XhoI sites of pcDNA3.1(+) (Invitrogen) is
302 also previously described (27). The RBDs of SARS-CoV-1 (Urbani isolate; GenBank AAP13441.1; a.a.
303 T320-D518), SARS-CoV-2 (YP_009724390; a.a. T333-K529), LYRa11 (AHX37558.1; a.a. T324-D522),
304 Rs7327 (ATO98218.1; a.a. T321-D519), Rs4231 (ATO98157.1; a.a. T320-D518), Rs4084 (ATO98132.1;
305 a.a. T321-D519) and RsSHC014 (AGZ48806.1; a.a. T321-D519) were cloned with N-terminal influenza
306 HA leader peptides (sequence MKTIIALSYIFCLVFA) and C-terminal 8-his tags (sequence
307 GSGHHHHHHHH) into the NheI-XhoI sites of pcDNA3.1(+). These plasmids are deposited with
308 Addgene under accession numbers 145145 and 161821-161826. Mutations were made by overlap
309 extension PCR and verified by Sanger sequencing.

310 **Tissue Culture.** Expi293F cells (ThermoFisher) were grown in Expi293 Expression Medium
311 (ThermoFisher) at 125 rpm, 8 % CO₂, 37 °C.

312 **Recombinant Protein Production.** Plasmids (500 ng DNA per ml culture) and polyethylenimine (MW
313 25,000; Polysciences; 3 µg per ml culture) were mixed with OptiMEM (Gibco; 100 µl per ml culture),
314 incubated 20 minutes at room temperature and added to Expi293F cells at a density of 2×10^6 / ml.
315 Transfection Enhancers (ThermoFisher) were added 18-23 h post-transfection. Culture supernatant was
316 harvested 4-6 days later by two centrifugation steps ($800 \times g$ for 10 minutes to remove cells and $20,000 \times$
317 g for 20 minutes to remove debris). IgG1 Fc fused and 8his-tagged proteins were subsequently purified as
318 previously described (27) using KANEKA KanCapA 3G Affinity (Pall) and HisPur Ni-NTA (Thermo
319 Scientific) resins, respectively. Eluted proteins from affinity chromatography were then separated on a

320 Superdex 200 Increase 10/300 GL column (GE Healthcare Life Sciences) equilibrated with Dulbecco's
321 phosphate-buffered saline (PBS). Proteins from peak fractions were concentrated using centrifugal
322 ultrafiltration devices (Millipore) to final concentrations of ~1 mg/ml (RBD-8h proteins), ~10 mg/ml
323 (sACE2₂-8h proteins) and ~50 mg/ml (sACE2₂-IgG1 proteins). Concentrations were determined by
324 absorbance at 280 nm using calculated extinction coefficients. Reported concentrations for sACE2₂ are
325 based on monomeric subunits. Aliquots were snap frozen in liquid N₂ and stored at -80 °C.

326 **Biolayer Interferometry.** BLI kinetics were collected on an Octet RED96a and analyzed with a 1:1
327 binding model (global fit) using instrument software (Molecular Devices). IgG1 Fc-fused sACE2₂ (wild
328 type or engineered variant sACE2₂.v2.4) were immobilized at 100 nM for 10 minutes to anti-human IgG
329 Fc biosensors (Molecular Devices). The assay buffer was 10 mM HEPES pH 7.6, 150 mM NaCl, 3 mM
330 EDTA, 0.05% polysorbate 20, 0.5% non-fat dry milk (Bio-Rad). Loaded sensors were equilibrated for 30
331 s in buffer, then dipped in RBD-8h solutions for 60 s to measure association and transferred back to
332 buffer to measure dissociation over 300 s.

333 **Library Construction, FACS and Illumina Sequencing Analysis.** Using plasmid pCEP4-myc-S
334 encoding tagged, full length S of SARS-CoV-2, saturation mutagenesis was focused to residues C336-
335 L517 forming the RBD. Degenerate NNK codons were introduced at all RBD positions using overlap
336 extension PCR as previously described (49). Transient transfection conditions were used that typically
337 provide no more than a single coding variant per cell (36, 37). Expi293F cells at 2×10^6 / ml were
338 transfected with a mixture of 1 ng coding plasmid (i.e. library DNA) with 1.5 µg pCEP4-ΔCMV carrier
339 plasmid (described in (37)). The medium was replaced 2 h post-transfection and cells were collected 24 h
340 post-transfection for FACS. Cells were washed with ice-cold PBS supplemented with 0.2 % bovine serum
341 albumin (PBS-BSA).

342 For investigations of binding to wild type sACE2₂, the cells expressing the S library were resuspended in
343 2.5 nM sACE2₂(WT)-8h and incubated 30 minutes on ice. Cells were washed twice with PBS-BSA and
344 then co-stained for 20 minutes with anti-myc Alexa 647 (clone 9B11, 1/250 dilution; Cell Signaling
345 Technology) and anti-HIS-FITC (chicken polyclonal, 1/100 dilution; Immunology Consultants
346 Laboratory). Cells were again washed twice before sorting on a BD FACS Aria II at the Roy J. Carver
347 Biotechnology Center. Dead cells, doublets and debris were excluded by first gating on the main
348 population by forward/side scattering and then excluding DAPI positive cells. From the myc-S-positive
349 (Alexa 647) population, the 20 % of cells with the highest and 20% of cells with the lowest anti-HIS-
350 FITC fluorescence for bound sACE2₂(WT)-8h were collected (Figure 2C). Collection tubes were coated
351 overnight with fetal bovine serum prior to sorting and contained Expi293 Expression Medium.
352 Fluorescent signals decreased over time during FACS, and therefore transfected cultures were prepared
353 on three separate occasions for a combined total of 8 hours sort time. The total numbers of collected cells
354 were 57,000 and 72,700 for the ACE2-High and ACE2-Low gates, respectively. Collected cells were
355 centrifuged (500 × g, 300 s) and pellets were frozen at -80°C. Samples from the independent sorts were
356 pooled during extraction of total RNA.

357 The competition selection was performed similarly, with the exception that cells expressing the S library
358 were incubated for 30 minutes in a mixture of 20 nM sACE2₂.v2.4-8h and 25 nM sACE2₂(WT)-IgG1.
359 After washing twice, bound proteins were stained for 30 minutes with anti-human IgG-APC (clone
360 HP6017, 1/250 dilution; BioLegend) and anti-HIS-FITC (chicken polyclonal, 1/100 dilution;
361 Immunology Consultants Laboratory). Cells were washed twice and sorted. After gating for the main
362 population of viable cells as described above, the 20 % of cells with the highest FITC-relative-to-APC
363 and highest APC-relative-to-FITC signals were collected (Figure 4C). The total numbers of collected
364 cells were 53,950 (Replicate 1: sACE2₂(WT)-specific gate), 42,860 (Replicate 1: sACE2₂.v2.4-specific
365 gate), 41,420 (Replicate 2: sACE2₂(WT)-specific gate), and 34,730 (Replicate 2: sACE2₂.v2.4-specific
366 gate).

367 Total RNA was extracted from the collected cells using a GeneJET RNA purification kit (Thermo
368 Scientific). First strand cDNA was synthesized with Accuscript (Agilent) primed with a gene-specific
369 oligonucleotide. The region of S scanned by saturation mutagenesis was PCR amplified as 3 overlapping
370 fragments that together span the full RBD sequence. Following a second round of PCR, primers added
371 adapters for annealing to the Illumina flow cell and sequencing primers, together with barcodes for
372 experiment identification. The PCR products were sequenced on an Illumina NovaSeq 6000 using a
373 2×250 nt paired end protocol. Data were analyzed using Enrich (38), where the frequencies of S variants
374 in the transcripts of the sorted populations were compared to their frequencies in the naive plasmid
375 library. Log₂ enrichment ratios for all the individual mutations were calculated and normalized by
376 subtracting the log₂ enrichment ratio for the wild type sequence across the same PCR-amplified fragment.
377 Conservation scores at residue positions were calculated by averaging the log₂ enrichment ratios for all
378 non-synonymous mutations at the residue.

379 **Flow Cytometry Analysis of SARS-CoV-2 S Mutants.** Expi293F cells at 2.0×10^6 cells/ml were
380 transfected with plasmid DNA (300 ng per ml of culture for measuring myc-S surface expression and
381 sACE₂ competition binding, 500 ng per ml for titration experiments) encoding myc-S variants using
382 Expifectamine (ThermoFisher) according to the manufacturer's directions. At 24 h post-transfection, cells
383 were washed with PBS-BSA. To detect surface expressed myc-S, cells were incubated with anti-myc
384 Alexa 647 (clone 9B11, 1/250 dilution; Cell Signaling Technology) on a rocker at 4°C for 30 minutes. To
385 measure competitive binding of wild type and engineered receptors, cells were instead incubated with 25
386 nM sACE₂(WT)-IgG1 and 20 nM sACE₂.v2.4-8h for 30 minutes at 4°C, washed twice, and stained
387 with anti-human IgG-APC (clone HP6017, 1/250 dilution; BioLegend) and anti-HIS-FITC (chicken
388 polyclonal, 1/100 dilution; Immunology Consultants Laboratory) secondary antibodies for 20 minutes at
389 4°C. Finally, in titration experiments, transfected cells were incubated for 30 minutes at 4°C with 1/3
390 serial dilutions of sACE₂(WT)-8h or sACE₂.v2.4-8h, followed by two washes and a 30 minute
391 incubation with anti-myc Alexa 647 (clone 9B11, 1/250 dilution) and anti-HIS-FITC (chicken polyclonal,
392 1/100 dilution). For all experiments, cells were washed twice prior to analysis on an Accuri C6 Flow
393 Cytometer (BD Biosciences) and data were processed with FCS Express (De Novo Software).
394 Quantification of myc-S surface expression is detailed in Figure S5.

395 **Reagent and Data Availability.** Plasmids for RBD protein expression are deposited with Addgene
396 (numbers 145145 and 161821-161826). Illumina sequencing data are deposited in NCBI's Gene
397 Expression Omnibus (GEO) under series accession number GSE159372.

398 ACKNOWLEDGEMENTS

399 The Roy J. Carver Biotechnology Center at the University of Illinois assisted with flow cytometry and
400 Illumina sequencing. E.P. designed the study and completed deep mutagenesis. K.K.C. purified proteins,
401 tested BLI kinetics. K.K.C., K.K.N. and T.J.C.T. prepared samples for flow cytometry. This work was in
402 part supported by NIH award R01AI129719 to E.P. The University of Illinois has filed a provisional
403 patent for engineered decoy receptors and E.P. and K.K.C. are co-founders of Orthogonal Biologics, Inc.

404

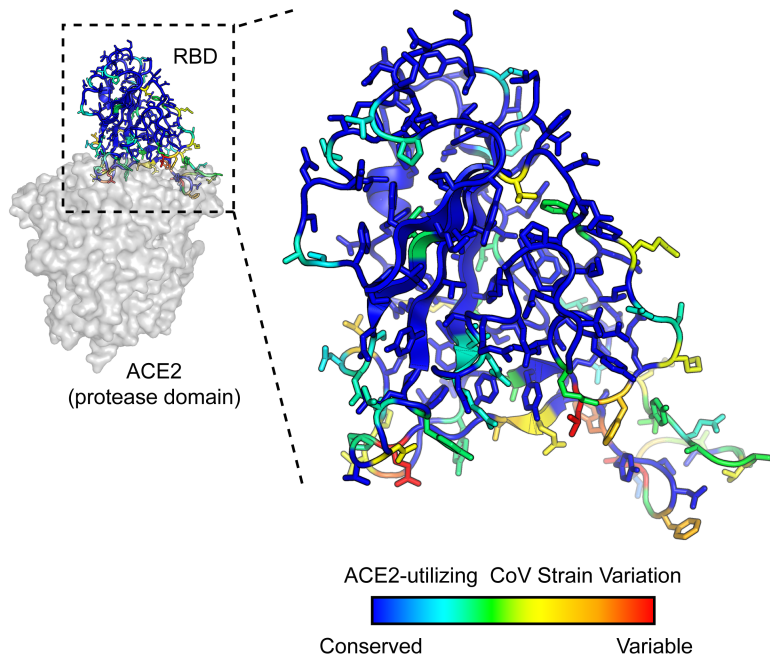
405 TABLES

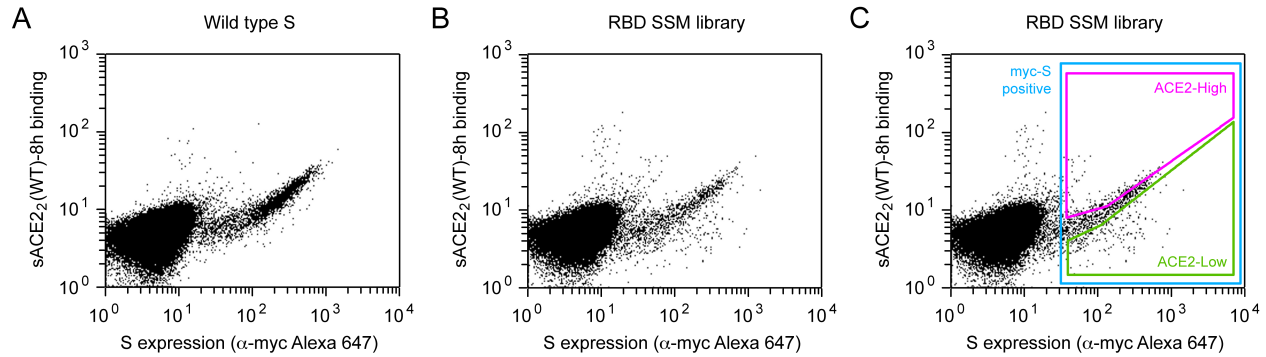
CoV strain ^a	Wild type sACE2 ₂ -IgG1 ^b				sACE2 ₂ .v2.4-IgG1			
	k _{on} (M ⁻¹ s ⁻¹)	k _{off} (s ⁻¹)	K _D (nM)	χ ^{2c}	k _{on} (M ⁻¹ s ⁻¹)	k _{off} (s ⁻¹)	K _D (nM)	χ ²
LYRa11	8.7 × 10 ⁵	7.9 × 10 ⁻²	91	0.12	1.4 × 10 ⁶	2.5 × 10 ⁻³	1.8	0.10
Rs7327	6.4 × 10 ⁵	4.0 × 10 ⁻²	63	0.25	9.8 × 10 ⁵	1.8 × 10 ⁻³	1.9	0.11
Rs4231	3.2 × 10 ⁵	2.2 × 10 ⁻²	69	0.04	4.5 × 10 ⁵	1.6 × 10 ⁻³	3.5	0.10
Rs4084	2.9 × 10 ⁵	2.5 × 10 ⁻²	85	0.24	4.8 × 10 ⁵	1.5 × 10 ⁻³	3.1	0.10
RsSHC014	8.8 × 10 ⁵	2.6 × 10 ⁻²	29	0.20	1.6 × 10 ⁶	2.0 × 10 ⁻³	1.3	0.29
SARS-1	6.6 × 10 ³	1.2 × 10 ⁻⁴	58	0.03	3.0 × 10 ³	5.6 × 10 ⁻⁶	2.1	0.03
SARS-2	1.4 × 10 ⁶	8.1 × 10 ⁻³	16	0.25	6.6 × 10 ⁵	2.8 × 10 ⁻⁴	0.4	0.09
SARS-2 (Y449K)	2.0 × 10 ⁶	9.0 × 10 ⁻²	46	0.67	4.3 × 10 ⁶	4.0 × 10 ⁻³	0.9	0.71
SARS-2 (N501W)	2.4 × 10 ⁶	5.4 × 10 ⁻³	2.3	0.43	3.3 × 10 ⁶	2.8 × 10 ⁻⁴	0.1	0.23
SARS-2 (N501Y)	2.2 × 10 ⁶	1.8 × 10 ⁻³	0.8	0.15	3.6 × 10 ⁵	1.1 × 10 ⁻⁴	< 0.1	0.24

^a Purified RBDs at 5 to 7 concentrations were used as the soluble analytes.
^b IgG1-Fc fused sACE2₂ was immobilized to anti-human IgG Fc capture biosensors.
^c χ² values represent the goodness of curve fitting. Acceptable values were considered to be less than 3.

406
407

408 **FIGURE LEGENDS**

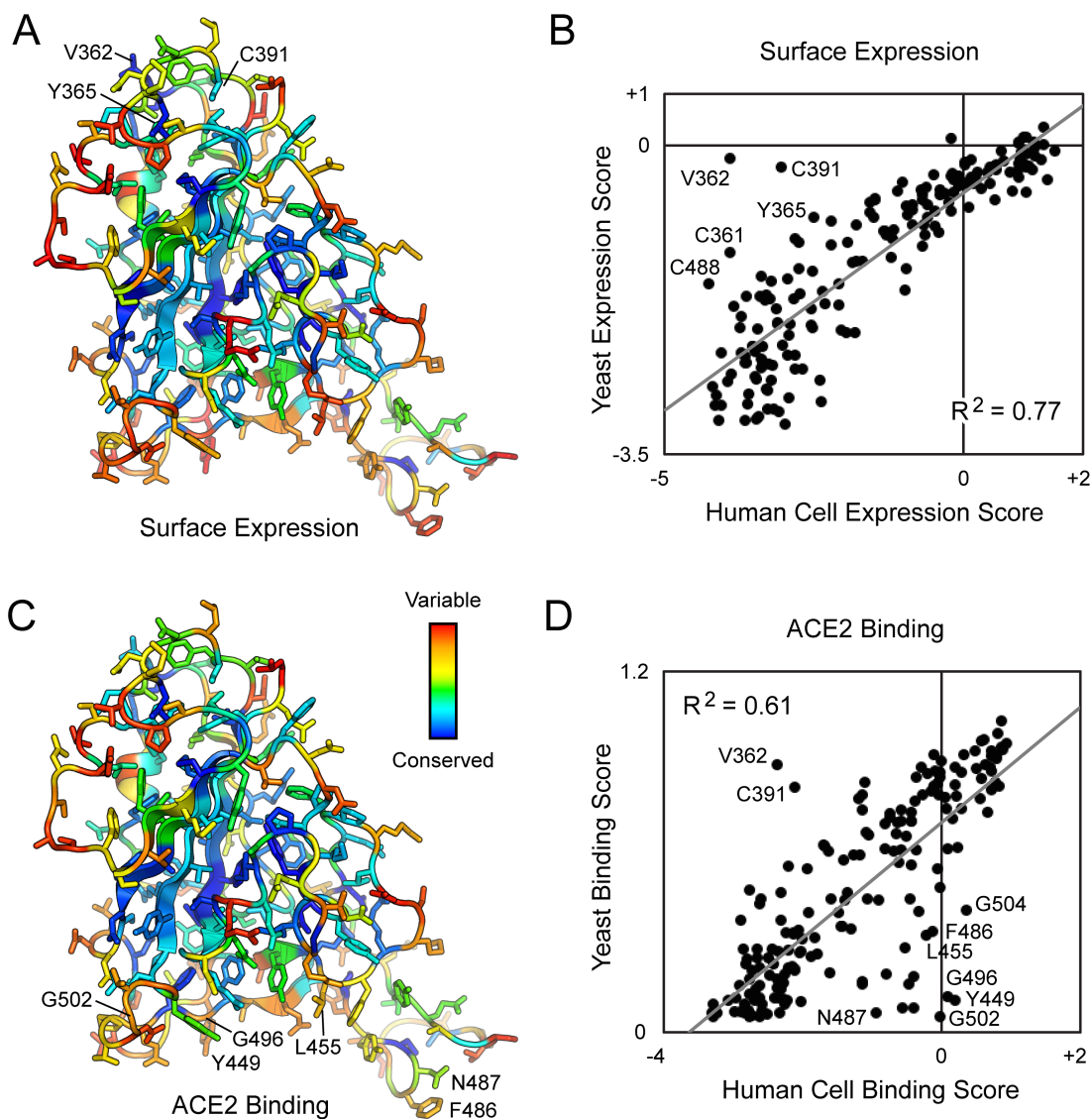




414

415 **Figure 2. FACS selection for variants of S with high or low binding signal to ACE2.** (A) Flow
416 cytometry analysis of Expi293F cells expressing full-length S of SARS-CoV-2 with an N-terminal c-myc
417 tag. Staining for the myc-epitope is on the x-axis while the detection of bound sACE2-8h (2.5 nM) is on
418 the y-axis. S plasmid was diluted 1500-fold by weight with carrier DNA so that cells typically express no
419 more than one coding variant; under these conditions most cells are negative. (B) Flow cytometry of cells
420 transfected with the RBD single site-saturation mutagenesis (SSM) library shows cells expressing S
421 variants with reduced sACE2-8h binding. (C) Gating strategy for FACS. S-expressing cells positive for
422 the c-myc epitope were gated (blue) and the highest ("ACE2-High") and lowest ("ACE2-Low") 20% of
423 cells with bound sACE2-8h relative to myc-S expression were collected.

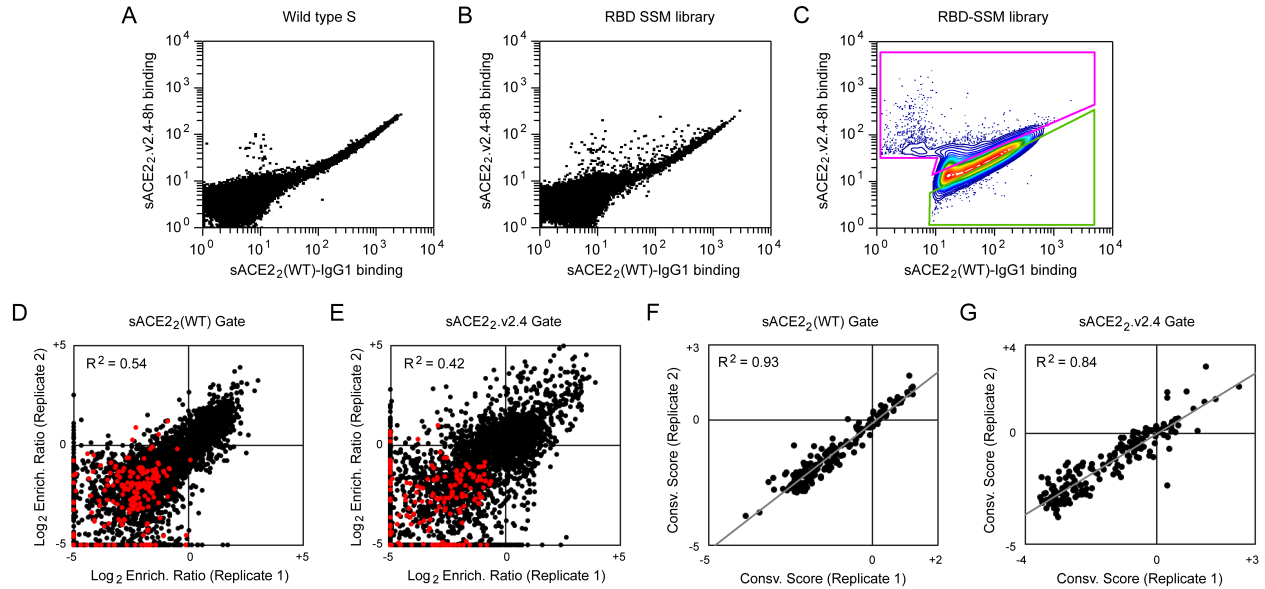
424



425

426 **Figure 3. Deep mutagenesis reveals that the ACE2-binding site of SARS-CoV-2 tolerates many**
427 **mutations.** (A) Positional scores for surface expression are mapped to the structure of the SARS-CoV-2
428 RBD (PDB 6M17, oriented as in Figure 1). Blue residues in the protein core are highly conserved in the
429 FACS selection for surface S expression (judged by depletion of mutations from the ACE2-High and
430 ACE2-Low gates), while surface residues in red tolerate mutations. (B) Correlation plot of expression
431 scores from mutant selection in human cells of full-length S (x-axis) versus the conservation scores (mean
432 of the \log_2 enrichment ratios at a residue position) from mutant selection in the isolated RBD by yeast
433 display (y-axis). Notable outliers are indicated. (C) Conservation scores from the ACE2-High gated cell
434 population are mapped to the RBD structure, with residues colored from low (blue) to high (red)
435 mutational tolerance. (D) Correlation plot of RBD conservation scores for high ACE2 binding from deep
436 mutagenesis of S in human cells (x-axis) versus deep mutagenesis of the RBD on the yeast surface (mean
437 of $\Delta K_{D app}$; y-axis).

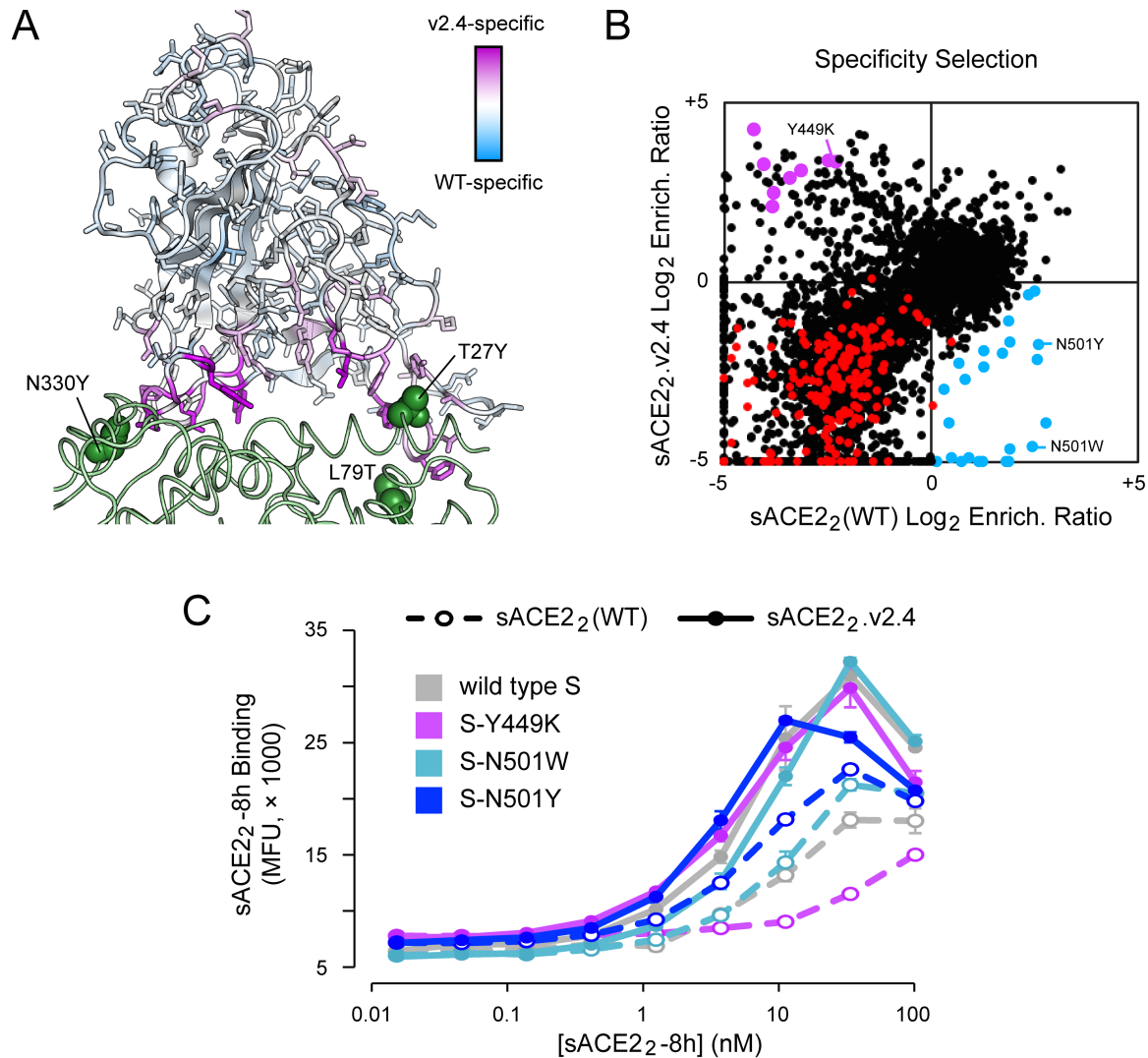
438



439

440 **Figure 4. A competition-based selection to identify RBD mutations within S of SARS-CoV-2 that**
441 **preferentially bind wild type or engineered ACE2 receptors. (A)** Expi293F cells were transfected with
442 wild type myc-S and incubated with competing sACE2₂(WT)-IgG1 (25 nM) and sACE2₂.v2.4-8h (20
443 nM). Bound protein was detected by flow cytometry after immuno-staining for the respective epitope
444 tags. **(B)** As in A, except cells were transfected with the RBD SSM library. A population of cells
445 expressing S variants with increased specificity towards sACE2₂.v2.4 is apparent (cells shifted to the
446 upper-left of the main population). **(C)** Gates used for FACS of cells expressing the RBD SSM library.
447 After excluding cells without bound protein, the top 20% of cells for bound sACE2₂.v2.4-8h (magenta
448 gate) and for bound sACE2₂(WT)-IgG1 (green gate) were collected. **(D-E)** Agreement between log₂
449 enrichment ratios from two independent FACS selections for cells expressing S variants with increased
450 specificity for (D) sACE2₂(WT) or (E) sACE2₂.v2.4. R² values are calculated for nonsynonymous
451 mutations (black). Nonsense mutations are red. **(F-G)** Conservation scores are calculated from the mean
452 of the log₂ enrichment ratios for all nonsynonymous substitutions at a given residue position. Correlation
453 plots show agreement between conservation scores for two independent selections for cells within the (D)
454 sACE2₂(WT) or (E) sACE2₂.v2.4 specific gates.

455



456

457 **Figure 5. Mutations within the RBD that confer specificity towards wild type ACE2 are rare.** (A)
 458 The SARS-CoV-2 RBD is colored by specificity score (the difference between the conservation scores for
 459 cells collected in the sACE2₂(WT) and sACE2₂.v2.4 specific gates). Residues that are hot spots for
 460 mutations with increased specificity towards sACE2₂(WT) are blue or towards sACE2₂.v2.4 are purple.
 461 The contacting surface of ACE2 is shown as a green ribbon, with sites of mutations in sACE2₂.v2.4
 462 labeled and shown as green spheres. (B) Log₂ enrichment ratios for mutations in S expressed by cell
 463 populations collected in the sACE2₂(WT) (x-axis) and sACE2₂.v2.4 specific gates. Data are the
 464 mean from two independent sorting experiments. S mutants in blue were predicted to have increased
 465 specificity for sACE2₂(WT) and were tested by targeted mutagenesis in Figure S6. S mutants in purple
 466 were predicted to have increased specificity for sACE2₂.v2.4 and were tested by targeted mutagenesis in
 467 Figure S7. Other nonsynonymous mutations are black. Nonsense mutations are red. (C) Wild type myc-S
 468 (grey) and three variants, Y449K (purple), N501W (light blue), and N501Y (dark blue), were expressed
 469 in Expi293F cells and tested by flow cytometry for binding to sACE2₂(WT)-8h (dashed lines) or
 470 sACE2₂.v2.4-8h (solid lines).

471 **REFERENCES**

- 472 1. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin.
473 *Nature*. **579**, 270–273 (2020).
- 474 2. A. C. Walls *et al.*, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein.
475 *Cell* (2020), doi:10.1016/j.cell.2020.02.058.
- 476 3. Y. Wan, J. Shang, R. Graham, R. S. Baric, F. Li, Receptor recognition by novel coronavirus from
477 Wuhan: An analysis based on decade-long structural studies of SARS. *J. Virol.* (2020),
478 doi:10.1128/JVI.00127-20.
- 479 4. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation.
480 *Science*, eabb2507 (2020).
- 481 5. M. Hoffmann *et al.*, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked
482 by a Clinically Proven Protease Inhibitor. *Cell* (2020), doi:10.1016/j.cell.2020.02.052.
- 483 6. W. Li *et al.*, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus.
484 *Nature*. **426**, 450–454 (2003).
- 485 7. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for
486 SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol.* **11**, 1860 (2020).
- 487 8. L. Samavati, B. D. Uhal, ACE2, Much More Than Just a Receptor for SARS-COV-2. *Front. Cell.*
488 *Infect. Microbiol.* **10**, 752 (2020).
- 489 9. F. Jiang *et al.*, Angiotensin-converting enzyme 2 and angiotensin 1–7: novel therapeutic targets.
490 *Nature Reviews Cardiology.* **11**, 413–426 (2014).
- 491 10. L. Zhang *et al.*, The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and
492 increases infectivity. *bioRxiv*, 2020.06.12.148726 (2020).
- 493 11. B. Korber *et al.*, Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases
494 Infectivity of the COVID-19 Virus. *Cell*. **182**, 812–827.e19 (2020).
- 495 12. V. Borges *et al.*, On the track of the D839Y mutation in the SARS-CoV-2 Spike fusion peptide:
496 emergence and geotemporal spread of a highly prevalent variant in Portugal. *medRxiv*,
497 2020.08.10.20171884 (2020).
- 498 13. D. S. Candido *et al.*, Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science*. **369**,
499 1255–1260 (2020).
- 500 14. B. B. Oude Munnink *et al.*, Transmission of SARS-CoV-2 on mink farms between humans and
501 mink and back to humans. *Science*, eabe5901 (2020).
- 502 15. S. Su *et al.*, Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends*
503 *Microbiol.* **24**, 490–502 (2016).
- 504 16. M. F. Boni *et al.*, Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for
505 the COVID-19 pandemic. *Nat Microbiol* (2020).
- 506 17. J. S. M. Sabir *et al.*, Co-circulation of three camel coronavirus species and recombination of
507 MERS-CoVs in Saudi Arabia. *Science*. **351**, 81–84 (2016).
- 508 18. D. VanInsberghe, A. Neish, A. C. Lowen, K. Koelle, Identification of SARS-CoV-2 recombinant
509 genomes. *bioRxiv*, 2020.08.05.238386 (2020).
- 510 19. X. Li *et al.*, Emergence of SARS-CoV-2 through recombination and strong purifying selection. *Sci*
511 *Adv.* **6**, eabb9153 (2020).

- 512 20. A. Baum *et al.*, Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape
513 seen with individual antibodies. *Science*, eabd0831 (2020).
- 514 21. A. J. Greaney *et al.*, Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding
515 domain that escape antibody recognition. *Cell Host & Microbe* (2020),
516 doi:10.1016/j.chom.2020.11.007.
- 517 22. M. A. Tortorici *et al.*, Ultrapotent human antibodies protect against SARS-CoV-2 challenge via
518 multiple mechanisms. *Science*, eabe3354 (2020).
- 519 23. D. Pinto *et al.*, Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV
520 antibody. *Nature* (2020), doi:10.1038/s41586-020-2349-y.
- 521 24. H. Hofmann *et al.*, Susceptibility to SARS coronavirus S protein-driven infection correlates with
522 expression of angiotensin converting enzyme 2 and infection can be blocked by soluble receptor.
523 *Biochem. Biophys. Res. Commun.* **319**, 1216–1221 (2004).
- 524 25. C. Lei *et al.*, Neutralization of SARS-CoV-2 spike pseudotyped virus by recombinant ACE2-Ig.
525 *Nat Commun.* **11**, 2070 (2020).
- 526 26. V. Monteil *et al.*, Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using
527 Clinical-Grade Soluble Human ACE2. *Cell* (2020), doi:10.1016/j.cell.2020.04.004.
- 528 27. K. K. Chan *et al.*, Engineering human ACE2 to optimize binding to the spike protein of SARS
529 coronavirus 2. *Science*. **4**, eabc0870 (2020).
- 530 28. A. Zoufaly *et al.*, Human recombinant soluble ACE2 in severe COVID-19. *Lancet Respir Med.* **8**,
531 1154–1158 (2020).
- 532 29. A. Glasgow *et al.*, Engineered ACE2 receptor traps potently neutralize SARS-CoV-2. *Proc Natl*
533 *Acad Sci USA.* **117**, 28046 (2020).
- 534 30. Y. Higuchi *et al.*, High affinity modified ACE2 receptors prevent SARS-CoV-2 infection. *bioRxiv*,
535 2020.09.16.299891 (2020).
- 536 31. H. K. Frank, D. Enard, S. D. Boyd, Exceptional diversity and selection pressure on SARS-CoV
537 and SARS-CoV-2 host receptor in bats compared to other mammals. *bioRxiv.* **5**, 562 (2020).
- 538 32. J. Shang *et al.*, Structural basis of receptor recognition by SARS-CoV-2. *Nature.* **382**, 1199
539 (2020).
- 540 33. R. N. Kirchdoerfer *et al.*, Stabilized coronavirus spikes are resistant to conformational changes
541 induced by receptor recognition or proteolysis. *Sci Rep.* **8**, 15701 (2018).
- 542 34. W. Li *et al.*, Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2.
543 *EMBO J.* **24**, 1634–1643 (2005).
- 544 35. D. M. Fowler, S. Fields, Deep mutational scanning: a new style of protein science. *Nat. Methods.*
545 **11**, 801–807 (2014).
- 546 36. J. D. Heredia *et al.*, Mapping Interaction Sites on Human Chemokine Receptors by Deep
547 Mutational Scanning. *J. Immunol.* **200**, ji1800343–3839 (2018).
- 548 37. J. Park *et al.*, Structural architecture of a dimeric class C GPCR based on co-trafficking of sweet
549 taste receptor subunits. *Journal of Biological Chemistry.* **294**, 4759–4774 (2019).
- 550 38. D. M. Fowler, C. L. Araya, W. Gerard, S. Fields, Enrich: software for analysis of protein function
551 by enrichment and depletion of variants. *Bioinformatics.* **27**, 3430–3431 (2011).
- 552 39. K. Wu, G. Peng, M. Wilken, R. J. Geraghty, F. Li, Mechanisms of host receptor adaptation by
553 severe acute respiratory syndrome coronavirus. *J. Biol. Chem.* **287**, 8904–8911 (2012).

- 554 40. T. N. Starr *et al.*, Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals
555 Constraints on Folding and ACE2 Binding. *Cell*. **182**, 1295–1310.e20 (2020).
- 556 41. T. W. Linsky *et al.*, De novo design of ACE2 protein decoys to neutralize SARS-CoV-2. *Science*,
557 abe0075 (2020).
- 558 42. Y. Cai *et al.*, Distinct conformational states of SARS-CoV-2 spike protein. *Science*. **369**, 1586
559 (2020).
- 560 43. H. Yao *et al.*, Molecular architecture of the SARS-CoV-2 virus. *Cell* (2020).
- 561 44. R. Henderson *et al.*, Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct.*
562 *Mol. Biol.* **27**, 925–933 (2020).
- 563 45. G. J. Rocklin *et al.*, Global analysis of protein folding using massively parallel design, synthesis,
564 and testing. *Science*. **357**, 168–175 (2017).
- 565 46. M. R. Gardner *et al.*, AAV-expressed eCD4-Ig provides durable protection from multiple SHIV
566 challenges. *Nature*. **519**, 87–91 (2015).
- 567 47. C. H. Fellingner *et al.*, eCD4-Ig Limits HIV-1 Escape More Effectively than CD4-Ig or a Broadly
568 Neutralizing Antibody. *J. Virol.* **93**, e00443–19 (2019).
- 569 48. L. Cao *et al.*, De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science*,
570 eabd9909 (2020).
- 571 49. E. Procko *et al.*, Computational design of a protein-based enzyme inhibitor. *J. Mol. Biol.* **425**,
572 3563–3575 (2013).
- 573