The sequence of human ACE2 is suboptimal for binding the S spike protein of SARS coronavirus 2

Erik Procko

Department of Biochemistry, University of Illinois, Urbana IL 61801
Email: coronavirus-research@illinois.edu

SUMMARY. The rapid and escalating spread of SARS coronavirus 2 (SARS-CoV-2) poses an immediate public health emergency. The viral spike protein S binds ACE2 on host cells to initiate molecular events that release the viral genome intracellularly. Soluble ACE2 inhibits entry of both SARS and SARS-2 coronaviruses by acting as a decoy for S binding sites, and is a candidate for therapeutic and prophylactic development. Using deep mutagenesis, variants of ACE2 are identified with increased binding to the receptor binding domain of S. Mutations are found across the interface, in the N90-glycosylation motif, and at buried sites where they are predicted to enhance folding and presentation of the interaction epitope. When single substitutions are combined, large increases in binding can be achieved. The mutational landscape offers a blueprint for engineering high affinity proteins and peptides that block receptor binding sites on S to meet this unprecedented challenge.

In December, 2019, a novel zoonotic betacoronavirus closely related to bat coronaviruses spilled over to humans at the Huanan Seafood Market in the Chinese city of Wuhan (*1*, *2*). The virus, called SARS-CoV-2 due to its similarities with the severe acute respiratory syndrome (SARS) coronavirus responsible for a smaller outbreak nearly two decades prior (*3*, *4*), has since spread human-to-human rapidly across the world, precipitating extraordinary containment measures from governments (*5*). Stock markets have fallen, travel restrictions have been imposed, public gatherings canceled, and large numbers of people are self-isolating. These events are unlike any experienced in generations. Symptoms of coronavirus disease 2019 (COVID-19) range from mild to dry cough, fever, pneumonia and death, and SARS-CoV-2 is devastating among the elderly and other vulnerable groups (*6*, *7*).

The S spike glycoprotein of SARS-CoV-2 binds angiotensin-converting enzyme 2 (ACE2) on host cells (*2*, *8-13*). S is a trimeric class I viral fusion protein that is proteolytically processed into S1 and S2 subunits that remain noncovalently associated in a prefusion state (*8*, *11*, *14*). Upon engagement of ACE2 by a receptor binding domain (RBD) in S1 (*15*), conformational rearrangements occur that cause S1 shedding, cleavage of S2 by host proteases, and exposure of a fusion peptide adjacent to the S2' proteolysis site (*14*, *16-18*). Favorable folding of S to a post-fusion conformation is coupled to host cell/virus membrane fusion and cytosolic release of viral RNA. Atomic contacts with the RBD are restricted to the protease domain of ACE2 (*19*, *20*), and soluble ACE2 (sACE2) in which the neck and transmembrane domains are removed is sufficient for binding S and neutralizing infection (*12*, *21-24*). In principle, the virus has limited potential to escape sACE2-

41 mediated neutralization without simultaneously decreasing affinity for native ACE2
42 receptors, thereby attenuating virulence. Furthermore, fusion of sACE2 to the Fc region of
43 human immunoglobulin can provide an avidity boost while recruiting immune effector
44 functions and increasing serum stability, an especially desirable quality if intended for
45 prophylaxis (*23*, *25*), and sACE2 has proven safe in healthy human subjects (*26*) and
46 patients with lung disease (*27*). Recombinant sACE2 has now been rushed into a clinical
47 trial for COVID-19 in Guangdong province, China (Clinicaltrials.gov #NCT04287686), and
48 peptide derivatives of ACE2 are also being explored as cell entry inhibitors (*28*).

49 Since human ACE2 has not evolved to recognize SARS-CoV-2 S, it was hypothesized that
50 mutations may be found that increase affinity for therapeutic and diagnostic applications.
51 The coding sequence of full length ACE2 with an N-terminal c-myc epitope tag was
52 diversified to create a library containing all possible single amino acid substitutions at 117
53 sites spanning the entire interface with S and lining the substrate-binding cavity. S binding
54 is independent of ACE2 catalytic activity (*23*) and occurs on the outer surface of ACE2 (*19*,
55 *20*), whereas angiotensin substrates bind within a deep cleft that houses the active site
56 (*29*). Substitutions within the substrate-binding cleft of ACE2 therefore act as controls that
57 are anticipated to have minimal impact on S interactions, yet may be useful for engineering
58 out substrate affinity to enhance *in vivo* safety. It is important to note though that
59 catalytically active protein may have desirable effects for replenishing lost ACE2 activity in
60 COVID-19 patients in respiratory distress (*30*, *31*).

61 The ACE2 library was transiently expressed in human Expi293F cells under conditions that
62 typically yield no more than one coding variant per cell, providing a tight link between
63 genotype and phenotype (*32*, *33*). Cells were then incubated with a subsaturating dilution
64 of medium containing the RBD of SARS-CoV-2 fused C-terminally to superfolder GFP
65 (sfGFP: (*34*)) (Fig. 1A). Levels of bound RBD-sfGFP correlate with surface expression levels
66 of myc-tagged ACE2 measured by dual color flow cytometry. Compared to cells expressing
67 wild type ACE2 (Fig. 1C), many variants in the ACE2 library fail to bind RBD, while there
68 appeared to be a smaller number of ACE2 variants with higher binding signals (Fig. 1D).
69 Cells expressing ACE2 variants with high or low binding to RBD were collected by
70 fluorescence-activated cell sorting (FACS), referred to as "nCoV-S-High" and "nCoV-S-Low"
71 sorted populations, respectively. During FACS, fluorescence signal for bound RBD-sfGFP
72 continuously declined, requiring the collection gates to be regularly updated to 'chase' the
73 relevant populations. This is consistent with RBD dissociating over hours during the
74 experiment. Reported affinities of RBD for ACE2 range from 1 to 15 nM (*8*, *10*).

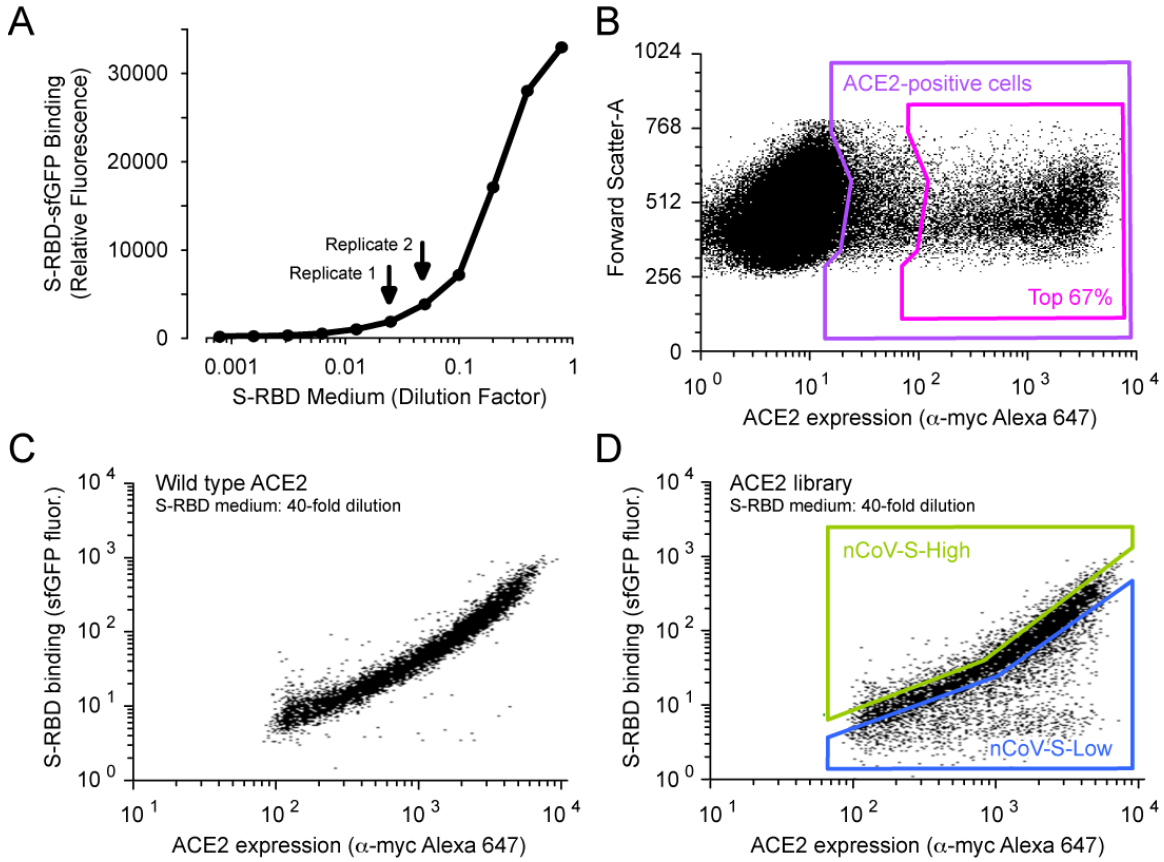


Figure 1. A selection strategy for ACE2 variants with high binding to the RBD of SARS-CoV-2 S.

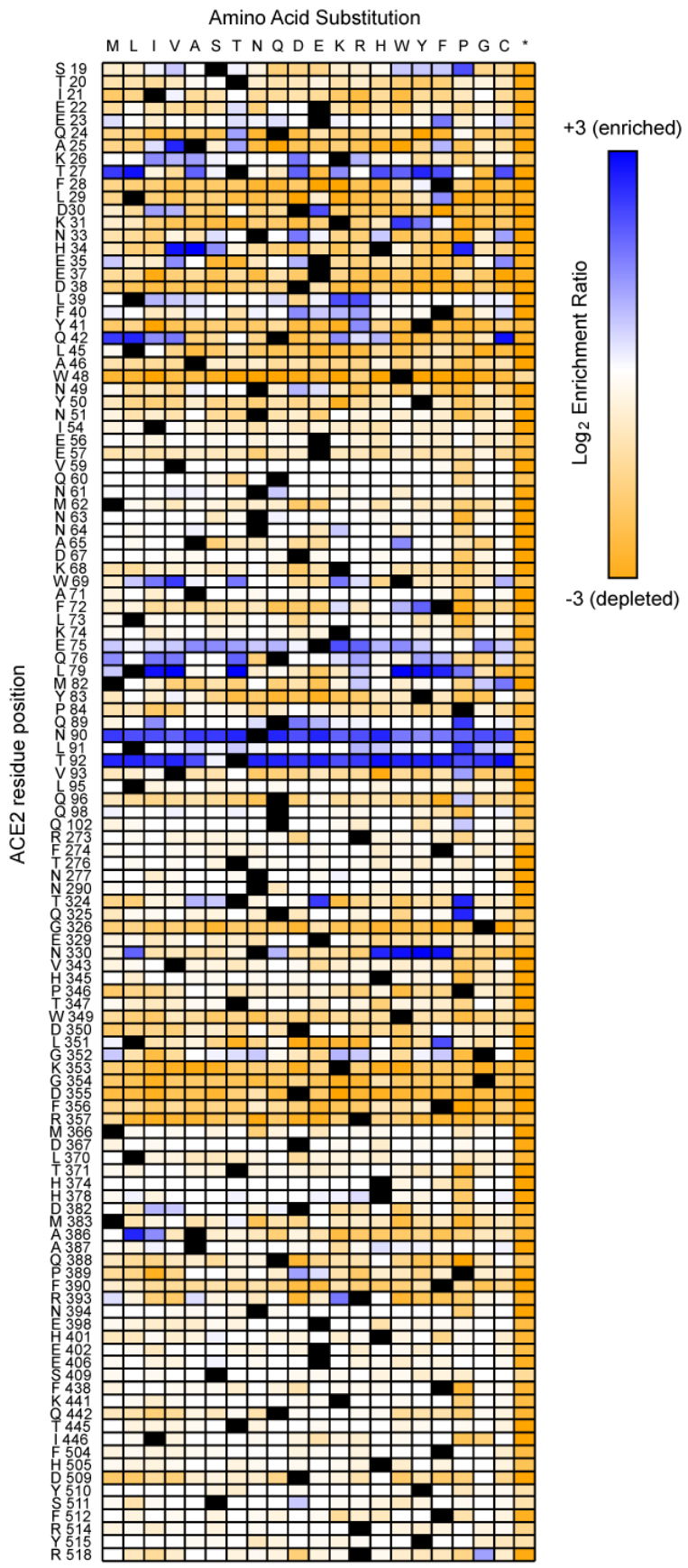
(A) Media from Expi293F cells secreting the SARS-CoV-2 RBD fused to sfGFP was collected and incubated at different dilutions with Expi293F cells expressing myc-tagged ACE2. Bound RBD-sfGFP was measured by flow cytometry. The dilutions of RBD-sfGFP-containing medium used for FACS selections are indicated by arrows.

(B-C) Expi293F cells were transfected with wild type ACE2 plasmid diluted with a large excess of carrier DNA. It has been previously shown that under these conditions, cells typically acquire no more than one coding plasmid and most cells are negative. Cells were incubated with RBD-sfGFP-containing medium and co-stained with fluorescent anti-myc to detect surface ACE2 by flow cytometry. During analysis, the top 67% (magenta gate) were chosen from the ACE2-positive population (purple gate) (B). Bound RBD was subsequently measured relative to surface ACE2 expression (C).

(D) Expi293F cells were transfected with an ACE2 single site-saturation mutagenesis library and analyzed as in B. During FACS, the top 15% of cells with bound RBD relative to ACE2 expression were collected (nCoV-S-High sort, green gate) and the bottom 20% were collected separately (nCoV-S-Low sort, blue gate).

Figure 2. A mutational landscape of ACE2 for high binding signal to the RBD of SARS-CoV-2 S.

Log_2 enrichment ratios from the nCoV-S-High sorts are plotted from ≤ -3 (i.e. depleted/deleterious, orange) to neutral (white) to ≥ +3 (i.e. enriched, dark blue). ACE2 primary structure is on the vertical axis, amino acid substitutions are on the horizontal axis. *, stop codon.


Amino Acid Substitution
M L I V A S T N Q D E K R H W Y F P G C *
ACE2 residue position
+3 (enriched)
Log2 Enrichment Ratio
-3 (depleted)

99

Transcripts in the sorted populations were deep sequenced, and frequencies of variants were compared to the naive plasmid library to calculate the enrichment or depletion of all 2,340 coding mutations in the library (Fig. 2). This approach of tracking an *in vitro* selection or evolution by deep sequencing is known as deep mutagenesis (*35*). Enrichment ratios (Fig. 3A and 3B) and residue conservation scores (Fig. 3D and 3E) closely agree between two independent sort experiments, giving confidence in the data. For the most part, enrichment ratios (Fig. 3C) and conservation scores (Fig. 3F) in the nCoV-S-High sorts are anticorrelated with the nCoV-S-Low sorts, with the exception of nonsense mutations which were appropriately depleted from both gates. This indicates that most, but not all, nonsynonymous mutations in ACE2 did not eliminate surface expression. The library is biased towards solvent-exposed residues and has few substitutions of buried hydrophobics that might have bigger effects on plasma membrane trafficking (*33*).

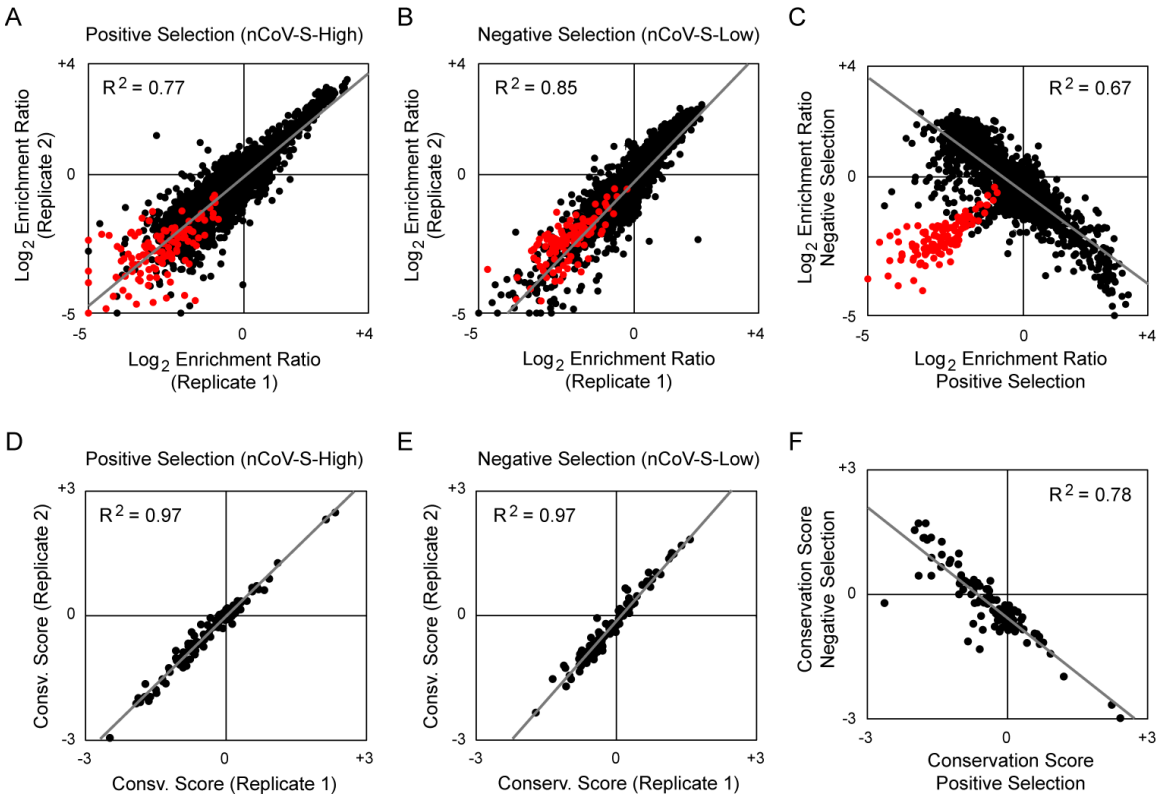


Figure 3. Data from independent replicates show close agreement.
(A-B) Log_2 enrichment ratios for ACE2 mutations in the nCoV-S-High (A) and nCoV-S-Low (B) sorts closely agree between two independent FACS experiments. Nonsynonymous mutations are black, nonsense mutations are red. Replicate 1 used a 1/40 dilution and replicate 2 used a 1/20 dilution of RBD-sfGFP-containing medium. R^2 values are for nonsynonymous mutations.
(C) Average log_2 enrichment ratios tend to be anticorrelated between the nCoV-S-High and nCoV-S-Low sorts. Nonsense mutations (red) and a small number of nonsynonymous mutations (black) are not expressed at the plasma membrane and are depleted from both sort populations (i.e. fall below the diagonal).
(D-F) Correlation plots of residue conservation scores from replicate nCoV-S-High (D) and nCoV-S-Low (E) sorts, and from the averaged data from both nCoV-S-High sorts compared to

both nCoV-S-Low sorts (F). Conservation scores are calculated from the mean of the $\log_2$ enrichment ratios for all amino acid substitutions at each residue position.

Mapping the experimental conservation scores from the nCoV-S-High sorts to the structure of RBD-bound ACE2 (*19*) shows that residues buried in the interface tend to be conserved, whereas residues at the interface periphery or in the substrate-binding cleft are mutationally tolerant (Fig. 4A). The region of ACE2 surrounding the C-terminal end of the ACE2 α1 helix and β3-β4 strands has a weak tolerance of polar residues, while amino acids at the N-terminal end of α1 and the C-terminal end of α2 prefer hydrophobics (Fig. 4B), likely in part to preserve hydrophobic packing between α1-α2. These discrete patches contact the globular RBD fold and a long protruding loop of the RBD, respectively.

Two ACE2 residues, N90 and T92 that together form a consensus N-glycosylation motif, are notable hot spots for enriched mutations (Fig. 2 and 4A). Indeed, all substitutions of N90 and T92, with the exception of T92S which maintains the N-glycan, are highly favorable for RBD binding, and the N90-glycan is thus predicted to partially hinder S/ACE2 interaction.

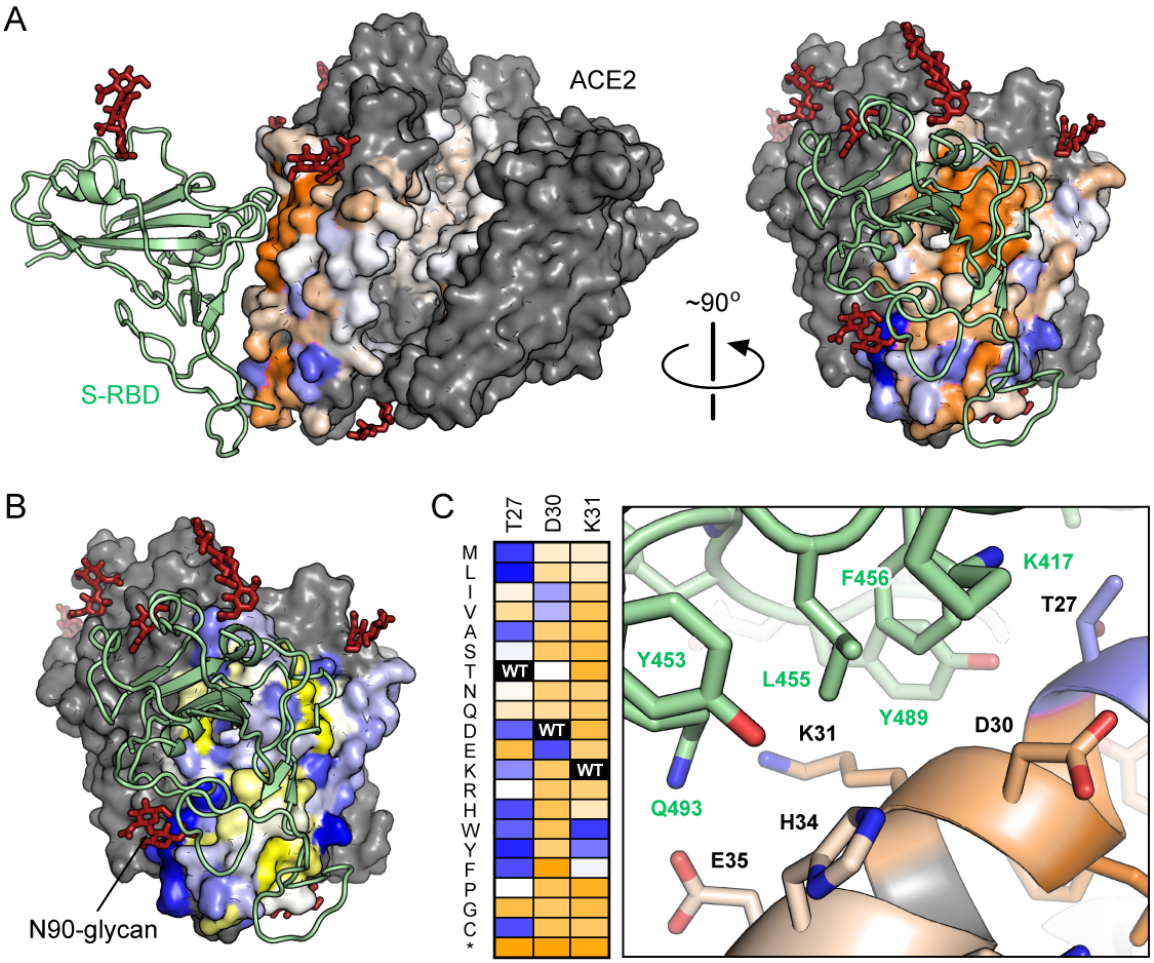


Figure 4. Sequence preferences of ACE2 residues for high binding to the RBD of SARS-CoV-2 S.

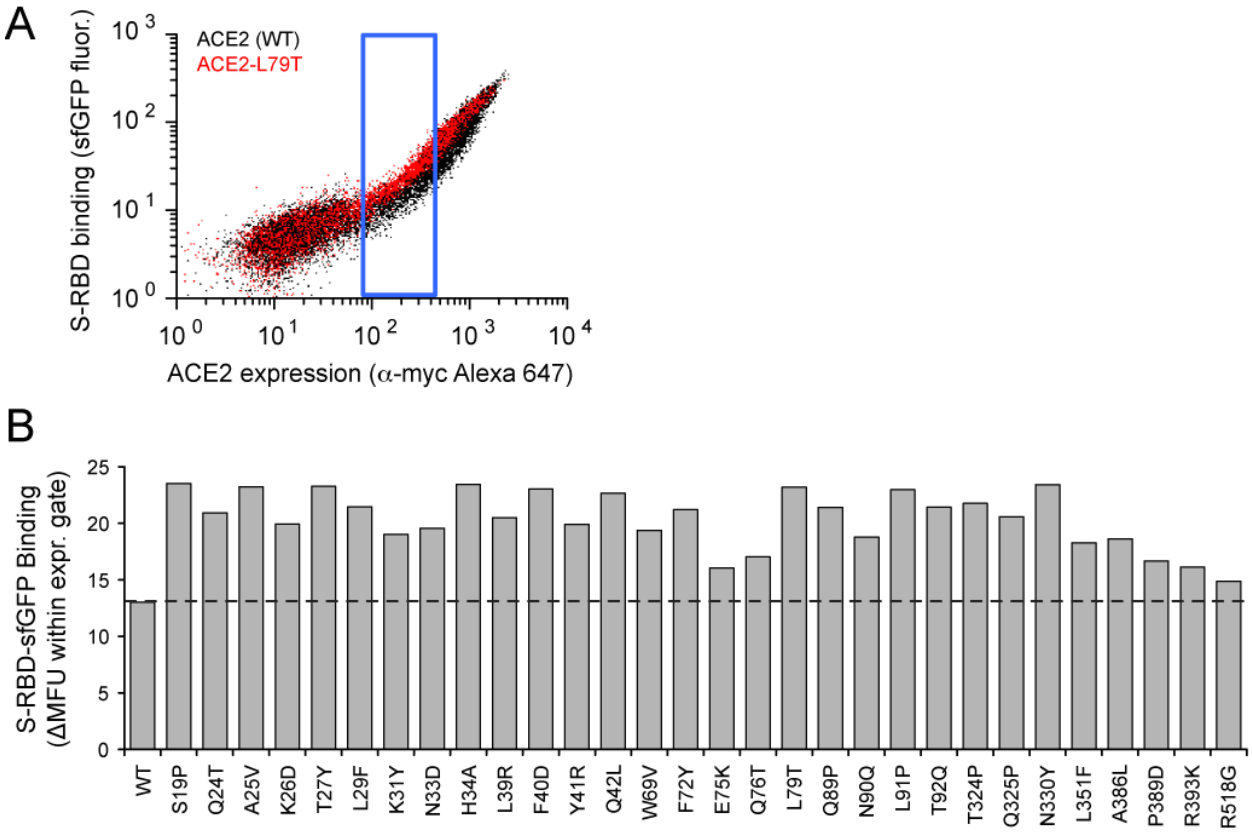
141 **(A)** Conservation scores from the nCoV-S-High sorts are mapped to the cryo-EM structure
142 (PDB 6M17) of RBD (pale green ribbon) bound ACE2 (surface). The view at left is looking
143 down the substrate-binding cavity, and only a single protease domain is shown for clarity.
144 Residues conserved for high RBD binding are orange; mutationally tolerant residues are pale
145 colors; residues that are hot spots for enriched mutations are blue; and residues maintained as
146 wild type in the ACE2 library are grey. Glycans are dark red sticks.
147 **(B)** Average hydrophobicity-weighted enrichment ratios are mapped to the RBD-bound ACE2
148 structure, with residues tolerant of polar substitutions in blue, while residues that prefer
149 hydrophobic amino acids are yellow.
150 **(C)** A magnified view of part of the ACE2 (colored by conservation score as in A) / RBD (pale
151 green) interface. Accompanying heatmap plots $\log_2$ enrichment ratios from the nCoV-S-High
152 sort for substitutions of ACE2-T27, D30 and K31 from ≤ -3 (depleted) in orange to ≥ +3
153 (enriched) in dark blue.

154 Mining the data identifies many ACE2 mutations that are enriched for RBD binding. For
155 instance, there are 122 mutations to 35 positions in the library that have $\log_2$ enrichment
156 ratios >1.5 in the nCoV-S-High sort. At least a dozen ACE2 mutations at the structurally
157 characterized interface enhance RBD binding, and will be useful for engineering highly
158 specific and tight binders of SARS-CoV-2 S. The molecular basis for how some of these
159 mutations enhance RBD binding can be rationalized from the RBD-bound cryo-EM
160 structure (Fig. 4C): hydrophobic substitutions of ACE2-T27 increase hydrophobic packing
161 with aromatic residues of S, ACE2-D30E extends an acidic side chain to reach S-K417, and
162 aromatic substitutions of ACE2-K31 contribute to an interfacial cluster of aromatics.
163 However, engineered ACE2 receptors with mutations at the interface may present binding
164 epitopes that are sufficiently different from native ACE2 that virus escape mutants can
165 emerge, or they may be strain specific and lack breadth to bind S in future coronavirus
166 outbreaks.

167 Instead, attention was drawn to mutations in the second shell and farther from the
168 interface that do not directly contact S but instead have putative structural roles. For
169 example, proline substitutions were enriched at five library positions (S19, L91, T92, T324
170 and Q325) where they might entropically stabilize the first turns of helices. Proline was
171 also enriched at H34, where it may enforce the central bulge in α1. Multiple mutations
172 were also enriched at buried positions where they will change local packing (e.g. A25V,
173 L29F, W69V, F72Y and L351F). The selection of ACE2 variants for high binding signal
174 therefore not only reports on affinity, but also on presentation at the membrane of folded
175 structure recognized by SARS-CoV-2 S. The presence of enriched structural mutations in
176 the sequence landscape is especially notable considering the ACE2 library was biased
177 towards solvent-exposed positions.

178 Thirty single substitutions highly enriched in the nCoV-S-High sort were validated by
179 targeted mutagenesis (Fig. 5). Binding of RBD-sfGFP to full length ACE2 mutants increased
180 compared to wild type, yet improvements were small and were most apparent on cells
181 expressing low ACE2 levels (Fig. 5A). To rapidly assess mutations in a format more
182 relevant to therapeutic development, the soluble ACE2 protease domain was fused to
183 sfGFP. Expression levels of sACE2-sfGFP were qualitatively evaluated by fluorescence of
184 the transfected cultures (Fig. 6A), and binding of sACE2-sfGFP to full length S expressed at

185 the plasma membrane was measured by flow cytometry (Fig. 6B). A single substitution
186 (T92Q) that eliminates the N90 glycan gave a small increase in binding signal (Fig. 6B).
187 Focusing on the most highly enriched substitutions in the selection for S binding that were
188 also spatially segregated to minimize negative epistasis (*36*), combinations of mutations in
189 sACE2 gave large increases in S binding (Table 1 and Fig. 6B). While this assay only
190 provides relative differences, the combinatorial mutants have enhanced binding by at least
191 an order of magnitude, which would give K_Ds in the picomolar range based on the
192 published K_D of wild type ACE2 with S (*8*, *10*). Other combinations of mutations may have
193 even greater effects. While monomeric sACE2 has been extensively studied *in vivo*
194 (including in man) and is the most likely to advance first to clinical use, fusions of sACE2 to
195 Fc of IgG or IgA classes may provide additional benefits of avidity, recruitment of immune
196 effector functions, and secretion in to the respiratory mucosa. Plasmids for these fusions
197 are deposited with Addgene.



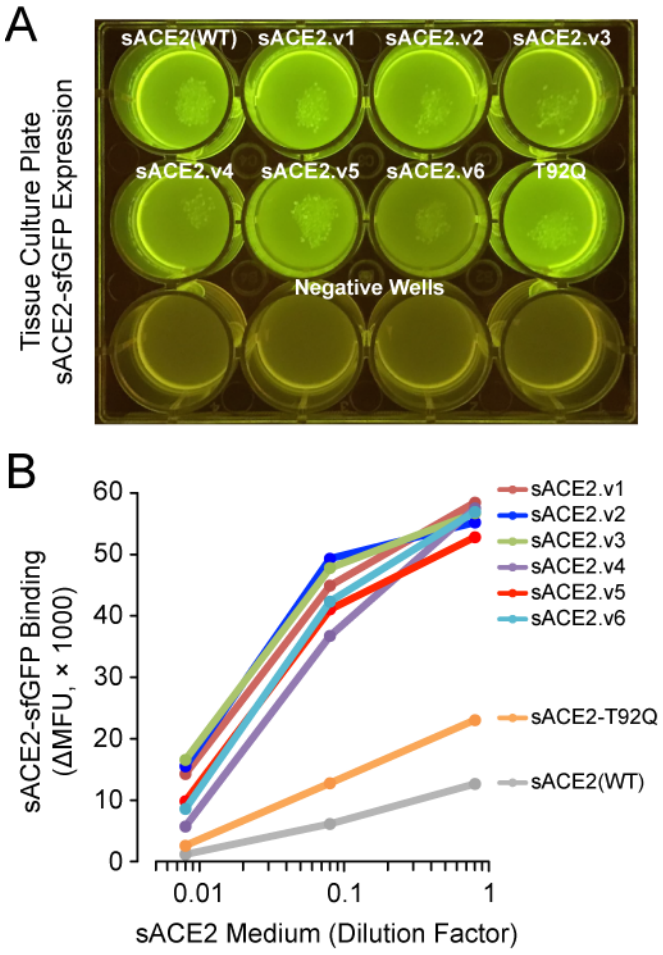
198

199 **Figure 5. Single amino acid substitutions in ACE2 predicted from the deep mutational**
200 **scan to increase RBD binding have small effects.**
201 **(A)** Expi293F cells expressing full length ACE2 were stained with RBD-sfGFP-containing
202 medium and analyzed by flow cytometry. Data are compared between wild type ACE2 (black)
203 and a single mutant (L79T, red). Increased RBD binding is most discernable in cells expressing
204 low levels of ACE2 (blue gate).
205 **(B)** RBD-sfGFP binding was measured for 30 single amino acid substitutions in ACE2. Data are
206 GFP mean fluorescence in the low expression gate (blue box in panel A) with background
207 fluorescence subtracted.

208

Table 1. Combinatorial mutants of sACE2.	
Variant	**Mutations**
sACE2.v1	H34A, T92Q, Q325P, A386L
sACE2.v2	T27Y, L79T, N330Y, A386L
sACE2.v3	A25V, T27Y, T92Q, Q325P, A386L
sACE2.v4	H34A, L79T, N330Y, A386L
sACE2.v5	A25V, T92Q, A386L
sACE2.v6	T27Y, Q42L, L79T, T92Q, Q325P, N330Y, A386L

209



210

211 **Figure 6. Engineered sACE2 with enhanced binding to S.**
212 **(A)** Expression of sACE2-sfGFP mutants was qualitatively evaluated by fluorescence of the
213 transfected cell cultures.
214 **(B)** Cells expressing full length S were stained with dilutions of sACE2-sfGFP-containing media
215 and binding was analyzed by flow cytometry.

216 While deep mutagenesis of viral proteins in replicating viruses has been extensively
217 pursued to understand escape mechanisms from drugs and antibodies, the work here

218 shows how deep mutagenesis can be directly applicable to therapeutic design when the
219 selection method is decoupled from virus replication and focused on host factors.

220 **METHODS**

221 **Plasmids.** The mature polypeptide (a.a. 19-805) of human ACE2 (GenBank NM_021804.1)
222 was cloned in to the NheI-XhoI sites of pCEP4 (Invitrogen) with a N-terminal HA leader
223 (MKTIIALSYIFCLVFA), myc-tag, and linker (GSPGGA). Soluble sACE2 fused to superfolder
224 GFP (*34*) was constructed by genetically joining the protease domain (a.a. 1-615) of ACE2
225 to sfGFP (GenBank ASL68970) via a gly/ser-rich linker (GSGGSGSGG), and pasting between
226 the NheI-XhoI sites of pcDNA3.1(+) (Invitrogen). A synthetic human codon-optimized gene
227 fragment (Integrated DNA Technologies) for the RBD (a.a. 333-529) of SARS-CoV-2 S
228 (GenBank YP_009724390.1) was N-terminally fused to a HA leader and C-terminally fused
229 to superfolder GFP (*34*) and ligated in to the NheI-XhoI sites of pcDNA3.1(+). Human
230 codon-optimized full length S (a.a. 16-1273) was subcloned from pUC57-2019-nCoV-
231 S(Human) (Molecular Cloud) with a N-terminal HA leader (MKTIIALSYIFCLVFA), myc-tag,
232 and linker (GSPGGA).

233 **Tissue Culture.** Expi293F cells (ThermoFisher) were cultured in Expi293 Expression
234 Medium (ThermoFisher) at 125 rpm, 8 % CO_2, 37 °C. For production of RBD-sfGFP, cells
235 were prepared to 2×10^6 / ml. Per ml of culture, 500 ng of pcDNA3-RBD-sfGFP and 3 μg of
236 polyethylenimine (MW 25,000; Polysciences) were mixed in 100 μl of OptiMEM (Gibco),
237 incubated for 20 minutes at room temperature, and added to cells. Transfection Enhancers
238 (ThermoFisher) were added 19 h post-transfection, and cells were cultured for 110 h. Cells
239 were removed by centrifugation at 800 × g for 5 minutes and medium was stored at -20 °C.
240 After thawing and immediately prior to use, remaining cell debris and precipitates were
241 removed by centrifugation at 20,000 × g for 5 minutes. Soluble ACE2-sfGFP protein was
242 produced by the same protocol with the following modifications: Transfection Enhancers
243 were added 22-$^1/_2$ h post-transfection, and medium supernatant was harvested after 60 h.

244 **Deep mutagenesis.** 117 residues within the protease domain of ACE2 were diversified by
245 overlap extension PCR (*37*) using primers with degenerate NNK codons. The plasmid
246 library was transfected in to Expi293F cells using Expifectamine under conditions
247 previously shown to typically give no more than a single coding variant per cell (*32*, *33*); 1
248 ng coding plasmid was diluted with 1,500 ng pCEP4-ΔCMV carrier plasmid per ml of cell
249 culture at 2×10^6 / ml, and the medium was replaced 2 h post-transfection. The cells were
250 collected after 24 h, washed with ice-cold PBS supplemented with 0.2 % bovine serum
251 albumin (PBS-BSA), and incubated for 30 minutes on ice with a 1/20 (replicate 1) or 1/40
252 (replicate 2) dilution of medium containing RBD-sfGFP into PBS-BSA. Cells were co-stained
253 with anti-myc Alexa 647 (clone 9B11, 1/250 dilution; Cell Signaling Technology). Cells
254 were washed twice with PBS-BSA, and sorted on a BD FACS Aria II at the Roy J. Carver
255 Biotechnology Center. The main cell population was gated by forward/side scattering to
256 remove debris and doublets, and DAPI was added to the sample to exclude dead cells. Of
257 the myc-positive (Alexa 647) population, the top 67% were gated (Fig. 1B). Of these, the 15
258 % of cells with the highest and 20% of cells with the lowest GFP fluorescence were
259 collected (Fig. 1D) in tubes coated overnight with fetal bovine serum and containing

260 Expi293 Expression Medium. Total RNA was extracted from the collected cells using a
261 GeneJET RNA purification kit (Thermo Scientific), and cDNA was reverse transcribed with
262 high fidelity Accuscript (Agilent) primed with gene-specific oligonucleotides. Diversified
263 regions of ACE2 were PCR amplified as 5 fragments. Flanking sequences on the primers
264 added adapters to the ends of the products for annealing to Illumina sequencing primers,
265 unique barcoding, and for binding the flow cell. Amplicons were sequenced on an Illumina
266 NovaSeq 6000 using a 2×250 nt paired end protocol. Data were analyzed using Enrich (*38*),
267 and commands are provided in the GEO deposit. Briefly, the frequencies of ACE2 variants in
268 the transcripts of the sorted populations were compared to their frequencies in the naive
269 plasmid library to calculate an enrichment ratio.

270 **Flow Cytometry Analysis of ACE2-S Binding.** Expi293F cells were transfected with
271 pcDNA3-myc-ACE2 or pcDNA3-myc-S plasmids (500 ng DNA per ml of culture at 2×10^6 /
272 ml) using Expifectamine (ThermoFisher). Cells were analyzed by flow cytometry 24 h
273 post-transfection. To analyze binding of RBD-sfGFP to full length myc-ACE2, cells were
274 washed with ice-cold PBS-BSA, and incubated for 30 minutes on ice with a 1/30 dilution of
275 medium containing RBD-sfGFP and a 1/240 dilution of anti-myc Alexa 647 (clone 9B11,
276 Cell Signaling Technology). Cells were washed twice with PBS-BSA and analyzed on a BD
277 LSR II. To analyze binding of sACE2-sfGFP to full length myc-S, cells were washed with
278 PBS-BSA, and incubated for 30 minutes on ice with a serial dilution of medium containing
279 sACE2-sfGFP and a 1/240 dilution of anti-myc Alexa 647 (clone 9B11, Cell Signaling
280 Technology). Cells were washed twice with PBS-BSA and analyzed on a BD Accuri C6. Data
281 were processed with FCS Express (De Novo Software)

282 **Reagent and data availability.** Plasmids are deposited with Addgene under IDs 141183-5
283 and 145145-78. Raw and processed deep sequencing data are deposited in NCBI's Gene
284 Expression Omnibus (GEO) with series accession no. GSE147194.

285 **ACKNOWLEDGEMENTS.** Staff at the UIUC Roy J. Carver Biotechnology Center assisted with
286 FACS and Illumina sequencing. Hannah Choi assisted with plasmid preparation. The
287 development of deep mutagenesis to study virus-receptor interactions was supported by
288 NIH award R01AI129719.

289 **CONFLICT OF INTEREST STATEMENT.** E.P. is the inventor on a provisional patent filing
290 by the University of Illinois covering aspects of this work. E.P. is a cofounder of Orthogonal
291 Biologics, Inc.

292 **REFERENCES**

293 1. N. Zhu *et al.*, A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N.*
294 *Engl. J. Med.* **382**, 727–733 (2020).

295 2. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable
296 bat origin. *Nature*. **579**, 270–273 (2020).

297 3. J. S. M. Peiris *et al.*, Coronavirus as a possible cause of severe acute respiratory
298 syndrome. *Lancet*. **361**, 1319–1325 (2003).

299 4. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses,
300 The species Severe acute respiratory syndrome-related coronavirus: classifying
301 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*. **4**, 3 (2020).

302 5. A. Patel, D. B. Jernigan, 2019-nCoV CDC Response Team, Initial Public Health
303 Response and Interim Clinical Guidance for the 2019 Novel Coronavirus Outbreak -
304 United States, December 31, 2019-February 4, 2020. *MMWR Morb. Mortal. Wkly. Rep.*
305 **69**, 140–146 (2020).

306 6. W. Wang, J. Tang, F. Wei, Updated understanding of the outbreak of 2019 novel
307 coronavirus (2019-nCoV) in Wuhan, China. *J. Med. Virol.* **92**, 441–447 (2020).

308 7. C. Huang *et al.*, Clinical features of patients infected with 2019 novel coronavirus in
309 Wuhan, China. *Lancet*. **395**, 497–506 (2020).

310 8. A. C. Walls *et al.*, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike
311 Glycoprotein. *Cell* (2020), doi:10.1016/j.cell.2020.02.058.

312 9. Y. Wan, J. Shang, R. Graham, R. S. Baric, F. Li, Receptor recognition by novel
313 coronavirus from Wuhan: An analysis based on decade-long structural studies of
314 SARS. *J. Virol.* (2020), doi:10.1128/JVI.00127-20.

315 10. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion
316 conformation. *Science*, eabb2507 (2020).

317 11. M. Hoffmann *et al.*, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is
318 Blocked by a Clinically Proven Protease Inhibitor. *Cell* (2020),
319 doi:10.1016/j.cell.2020.02.052.

320 12. W. Li *et al.*, Angiotensin-converting enzyme 2 is a functional receptor for the SARS
321 coronavirus. *Nature*. **426**, 450–454 (2003).

322 13. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor
323 usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol*. **11**, 1860
324 (2020).

325 14. M. A. Tortorici, D. Veesler, Structural insights into coronavirus entry. *Adv. Virus Res.*
326 **105**, 93–116 (2019).

327 15. S. K. Wong, W. Li, M. J. Moore, H. Choe, M. Farzan, A 193-amino acid fragment of the
328 SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J. Biol.*
329 *Chem.* **279**, 3197–3201 (2004).

330 16. I. G. Madu, S. L. Roth, S. Belouzard, G. R. Whittaker, Characterization of a highly
331 conserved domain within the severe acute respiratory syndrome coronavirus spike
332 protein S2 domain with characteristics of a viral fusion peptide. *J. Virol.* **83**, 7411–
333 7421 (2009).

334 17. A. C. Walls *et al.*, Tectonic conformational changes of a coronavirus spike
335 glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11157–
336 11162 (2017).

337 18. J. K. Millet, G. R. Whittaker, Host cell entry of Middle East respiratory syndrome
338 coronavirus after two-step, furin-mediated activation of the spike protein. *Proc. Natl.*
339 *Acad. Sci. U.S.A.* **111**, 15214–15219 (2014).

340 19. R. Yan *et al.*, Structural basis for the recognition of the SARS-CoV-2 by full-length
341 human ACE2. *Science*, eabb2762 (2020).

342 20. F. Li, W. Li, M. Farzan, S. C. Harrison, Structure of SARS coronavirus spike receptor-
343 binding domain complexed with receptor. *Science*. **309**, 1864–1868 (2005).

344 21. H. Hofmann *et al.*, Susceptibility to SARS coronavirus S protein-driven infection
345 correlates with expression of angiotensin converting enzyme 2 and infection can be
346 blocked by soluble receptor. *Biochem. Biophys. Res. Commun.* **319**, 1216–1221
347 (2004).

348 22. C. Lei *et al.*, Potent neutralization of 2019 novel coronavirus by recombinant ACE2-Ig.
349 *bioRxiv*, 2020.02.01.929976 (2020).

350 23. M. J. Moore *et al.*, Retroviruses pseudotyped with the severe acute respiratory
351 syndrome coronavirus spike protein efficiently infect cells expressing angiotensin-
352 converting enzyme 2. *J. Virol.* **78**, 10628–10635 (2004).

353 24. V. Monteil *et al.*, Inhibition of SARS-CoV-2 infections in engineered human tissues
354 using clinical-grade soluble human ACE2. *Cell*. **DOI: 10.1016/j.cell.2020.04.004**, 1–
355 28 (2020).

356 25. P. Liu *et al.*, Novel ACE2-Fc chimeric fusion provides long-lasting hypertension
357 control and organ protection in mouse models of systemic renin angiotensin system
358 activation. *Kidney Int.* **94**, 114–125 (2018).

359 26. M. Haschke *et al.*, Pharmacokinetics and pharmacodynamics of recombinant human
360 angiotensin-converting enzyme 2 in healthy human subjects. *Clin Pharmacokinet*. **52**,
361 783–792 (2013).

362 27. A. Khan *et al.*, A pilot clinical trial of recombinant human angiotensin-converting
363 enzyme 2 in acute respiratory distress syndrome. *Crit Care*. **21**, 234 (2017).

364 28. G. Zhang, S. Pomplun, A. R. Loftis, A. Loas, B. L. Pentelute, The first-in-class peptide
365 binder to the SARS-CoV-2 spike protein. *bioRxiv*, 2020.03.19.999318 (2020).

366 29. P. Towler *et al.*, ACE2 X-ray structures reveal a large hinge-bending motion
367 important for inhibitor binding and catalysis. *J. Biol. Chem.* **279**, 17996–18007
368 (2004).

369 30. R. L. Kruse, Therapeutic strategies in an outbreak scenario to treat the novel
370 coronavirus originating in Wuhan, China. *F1000Res*. **9**, 72 (2020).

371 31. H. Zhang, J. M. Penninger, Y. Li, N. Zhong, A. S. Slutsky, Angiotensin-converting
372 enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential
373 therapeutic target. *Intensive Care Med*. **309**, 1864 (2020).

374 32. J. D. Heredia *et al.*, Mapping Interaction Sites on Human Chemokine Receptors by
375 Deep Mutational Scanning. *J. Immunol.* **200**, ji1800343–3839 (2018).

376 33. J. Park *et al.*, Structural architecture of a dimeric class C GPCR based on co-trafficking
377 of sweet taste receptor subunits. *Journal of Biological Chemistry*. **294**, 4759–4774
378 (2019).

379 34. J.-D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger, G. S. Waldo, Engineering and
380 characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79–
381 88 (2006).

382 35. D. M. Fowler, S. Fields, Deep mutational scanning: a new style of protein science. *Nat.*
383 *Methods*. **11**, 801–807 (2014).

384 36. J. D. Heredia, J. Park, H. Choi, K. S. Gill, E. Procko, Conformational Engineering of HIV-
385 1 Env Based on Mutational Tolerance in the CD4 and PG16 Bound States. *J. Virol.* **93**,
386 e00219–19 (2019).

387 37. E. Procko *et al.*, Computational design of a protein-based enzyme inhibitor. *J. Mol.*
388 *Biol.* **425**, 3563–3575 (2013).

389 38. D. M. Fowler, C. L. Araya, W. Gerard, S. Fields, Enrich: software for analysis of protein
390 function by enrichment and depletion of variants. *Bioinformatics*. **27**, 3430–3431
391 (2011).

392