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, and no approved therapeutics or vaccines are currently available. 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 at a cell surface. Mutations are found across the interface and also at buried sites where they are predicted to enhance folding and presentation of the interaction epitope. The N90-glycan on ACE2 hinders association. The mutational landscape offers a blueprint for engineering high affinity ACE2 receptors to meet this unprecedented challenge.

In December, 2018, 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 quarantined. 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*, *21-23*). In principle, the virus has limited potential to escape sACE2-mediated neutralization without simultaneously decreasing affinity for native ACE2

41 receptors, thereby attenuating virulence. Furthermore, fusion of sACE2 to the Fc region of
42 human immunoglobulin can provide an avidity boost while recruiting immune effector
43 functions and increasing serum stability, an especially desirable quality if intended for
44 prophylaxis (*23*, *24*), and sACE2 has proven safe in healthy human subjects (*25*) and
45 patients with lung disease (*26*). Recombinant sACE2 has now been rushed into a clinical
46 trial for COVID-19 in Guangdong province, China (Clinicaltrials.gov #NCT04287686).

47 Since human ACE2 has not evolved to recognize SARS-CoV-2 S, it was hypothesized that
48 mutations may be found that increase affinity for therapeutic and diagnostic applications.
49 The coding sequence of full length ACE2 with an N-terminal c-myc epitope tag was
50 diversified to create a library containing all possible single amino acid substitutions at 117
51 sites spanning the entire interface with S and lining the substrate-binding cavity. S binding
52 is independent of ACE2 catalytic activity (*23*) and occurs on the outer surface of ACE2 (*19*,
53 *20*), whereas angiotensin substrates bind within a deep cleft that houses the active site
54 (*27*). Substitutions within the substrate-binding cleft of ACE2 therefore act as controls that
55 are anticipated to have minimal impact on S interactions, yet may be useful for engineering
56 out substrate affinity to enhance *in vivo* safety. It is important to note though that
57 catalytically active protein may have desirable effects for replenishing lost ACE2 activity in
58 COVID-19 patients in respiratory distress (*28*, *29*).

59 The ACE2 library was transiently expressed in human Expi293F cells under conditions that
60 typically yield no more than one coding variant per cell, providing a tight link between
61 genotype and phenotype (*30*, *31*). Cells were then incubated with a subsaturating dilution
62 of medium containing the RBD of SARS-CoV-2 fused C-terminally to superfolder GFP
63 (sfGFP: (*32*)) (Fig. 1A). Levels of bound RBD-sfGFP correlate with surface expression levels
64 of myc-tagged ACE2 measured by dual color flow cytometry. Compared to cells expressing
65 wild type ACE2 (Fig. 1C), many variants in the ACE2 library fail to bind RBD, while there
66 appeared to be a smaller number of ACE2 variants with higher binding signals (Fig. 1D).
67 Cells expressing ACE2 variants with high or low binding to RBD were collected by
68 fluorescence-activated cell sorting (FACS), referred to as "nCoV-S-High" and "nCoV-S-Low"
69 sorted populations, respectively. During FACS, fluorescence signal for bound RBD-sfGFP
70 continuously declined, requiring the collection gates to be regularly updated to 'chase' the
71 relevant populations. This is consistent with RBD dissociating over hours during the
72 experiment. Reported affinities of RBD for ACE2 range from 1 to 15 nM (*8*, *10*).

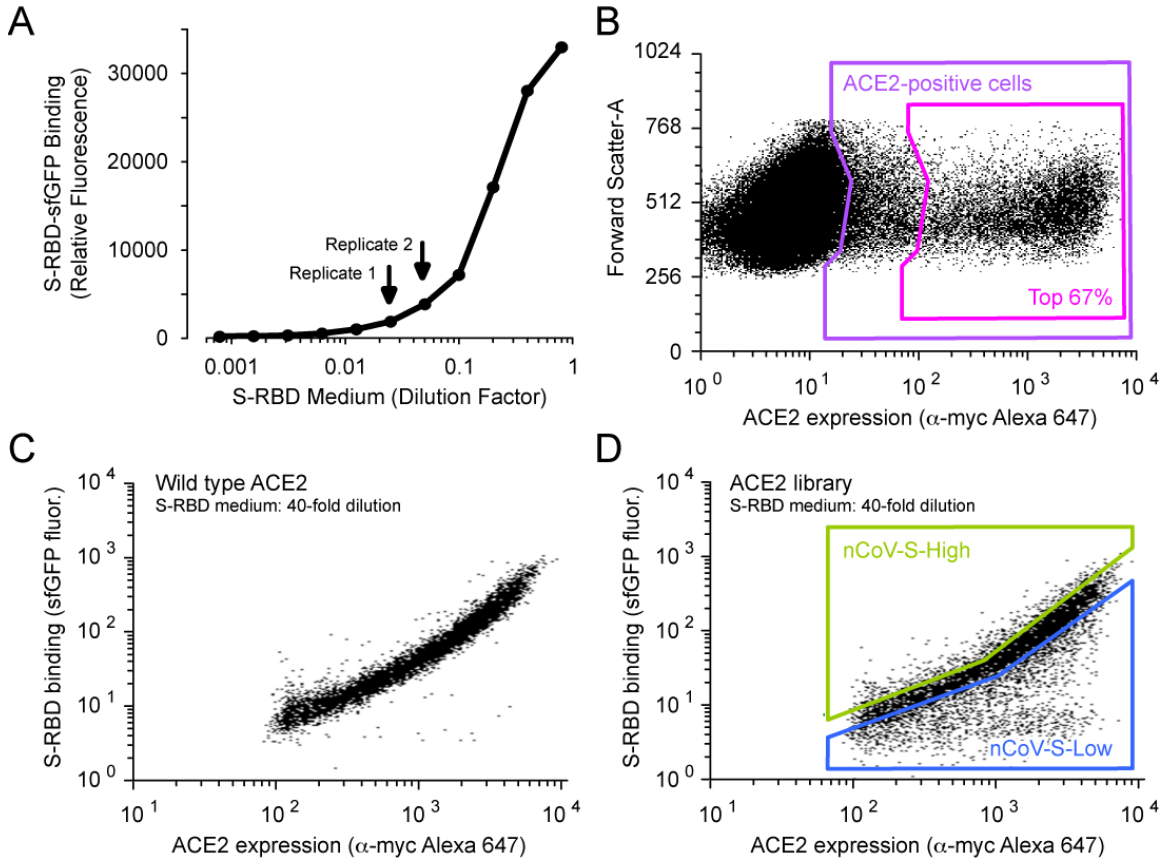


Figure 1. A selection strategy for ACE2 variants with high binding to the RBD of SARS-CoV-2 S.

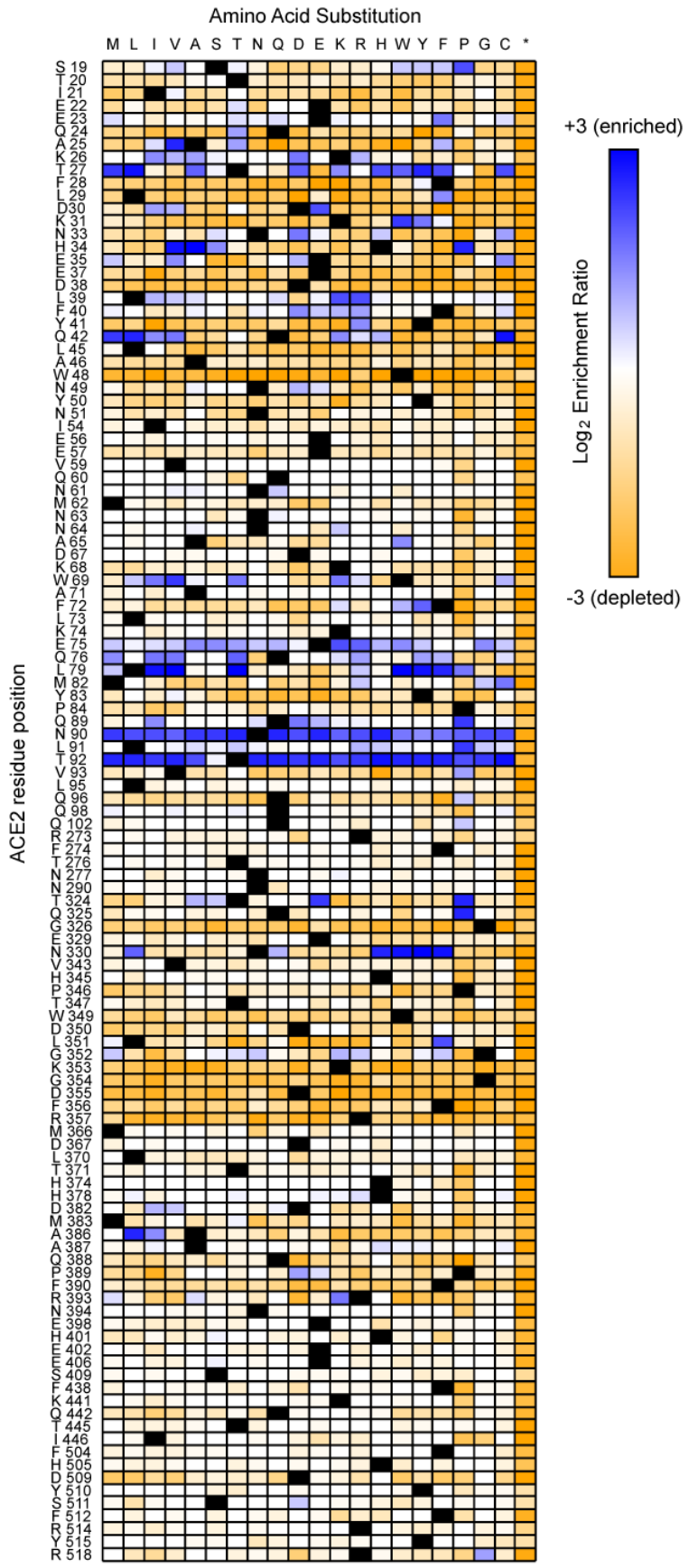
(A) Media from Expi293F cells secreting the SARS-CoV-2 RBD fused to sfGFP was collected and incubated at different dilutions with Expi293F cells expressing myc-tagged ACE2. Bound RBD-sfGFP was measured by flow cytometry. The dilutions of RBD-sfGFP-containing medium used for FACS selections are indicated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

(B-C) Expi293F cells were transiently 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
S 19
T 20
I 21
E 22
E 23
Q 24
A 25
K 26
T 27
F 28
L 29
D 30
K 31
N 33
H 34
E 35
E 37
D 38
L 39
F 40
Y 41
Q 42
L 45
A 46
W 48
N 49
Y 50
N 51
I 54
E 56
E 57
V 59
Q 60
N 61
M 62
N 63
N 64
A 65
D 67
K 68
W 69
A 71
F 72
L 73
K 74
E 75
Q 76
L 79
M 82
Y 83
P 84
Q 89
N 90
L 91
T 92
V 93
L 95
Q 96
Q 98
Q 102
R 273
F 274
T 276
N 277
N 290
T 324
Q 325
G 326
E 329
N 330
V 343
H 345
P 346
T 347
W 349
D 350
L 351
G 352
K 353
G 354
D 355
F 356
R 357
M 366
D 367
L 370
T 371
H 374
H 378
D 382
M 383
A 386
A 387
Q 388
P 389
F 390
R 393
N 394
E 398
H 401
E 402
E 406
S 409
F 438
K 441
Q 442
T 445
I 446
F 504
H 505
D 509
Y 510
S 511
F 512
R 514
Y 515
R 518
+3 (enriched)
Log2 Enrichment Ratio
-3 (depleted)

97

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 (*33*). 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 (*31*).

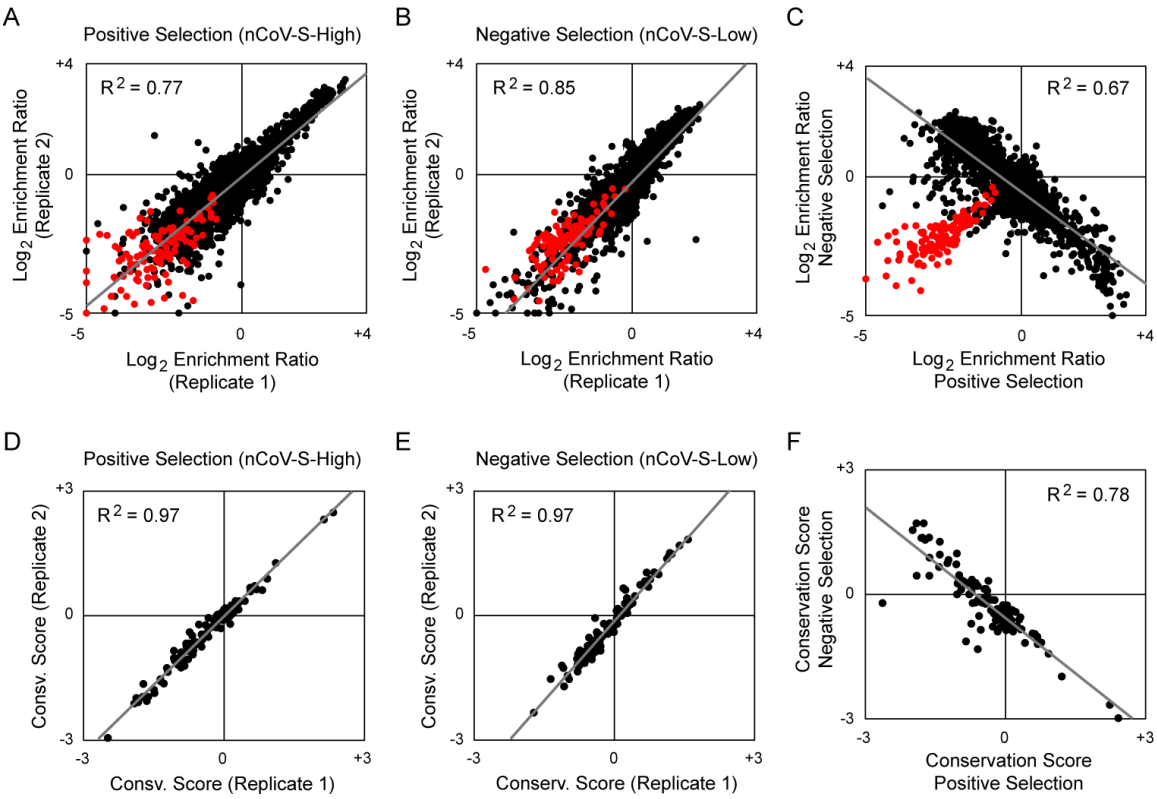


Figure 3. Data from independent replicates show close agreement.
(A-B) Log_2 enrichment ratios for ACE2 mutations in the nCoV-S-High (A) and nCoV-S-Low (B) sorts closely agree between two independent FACS experiments. Nonsynonymous mutations are black, nonsense mutations are red. Replicate 1 used a 1/40 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.

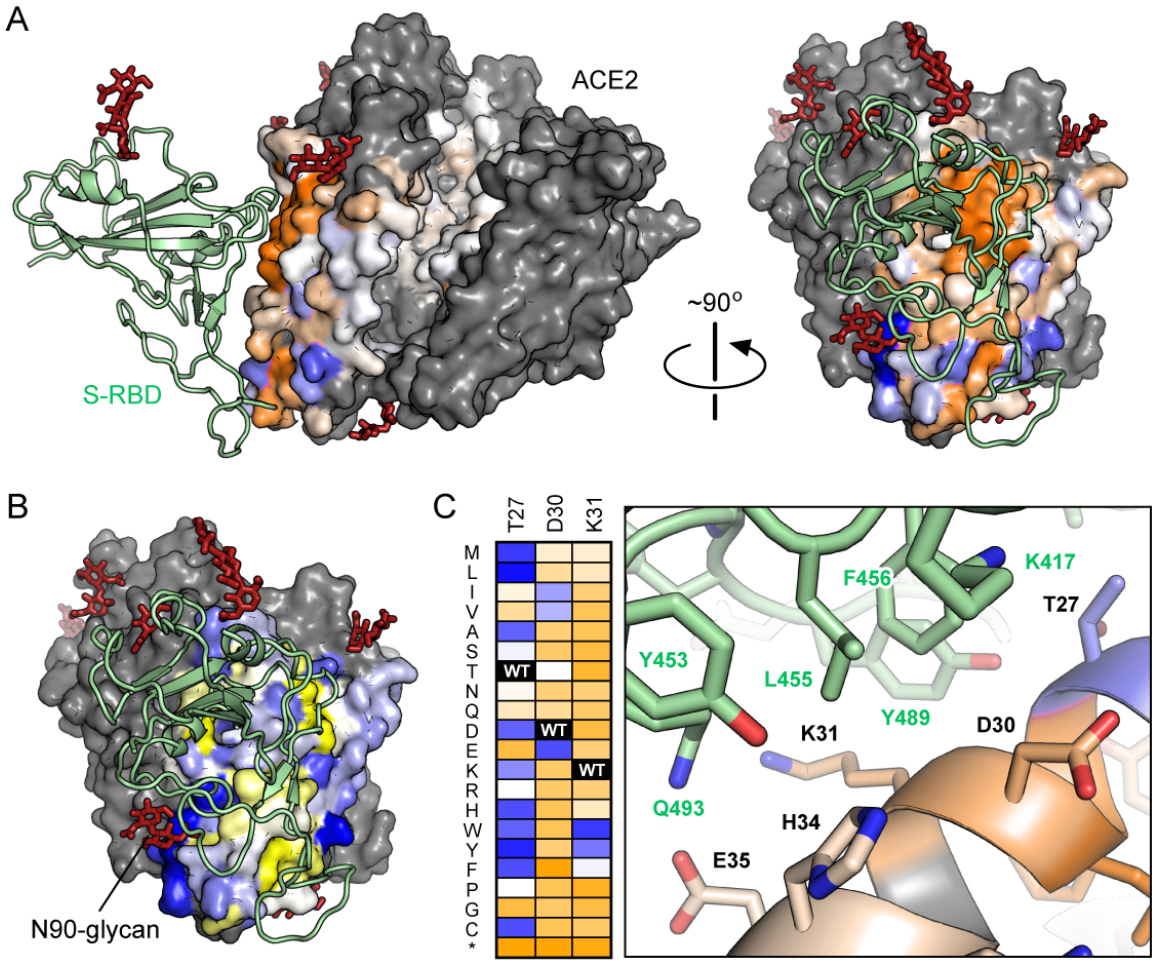


Figure 4. Sequence preferences of ACE2 residues for high binding to the RBD of SARS-CoV-2 S.

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

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

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

Deep mutational scans in human cells have errors (*34*), and it is unclear how large an effect an enriched mutation in a selection will have when introduced in a purified protein. Mutations of interest for ACE2 engineering will need careful assessment by targeted mutagenesis, as well as considerations on how best to combine mutations for production of conformationally-stable, high affinity sACE2. Other considerations will be whether to fuse sACE2 to Fc of IgG1 or IgA1 to evoke specialized immune effector functions, or to fuse with albumin to boost serum stability without risking an excessive inflammatory response. These are unknowns.

183 While deep mutagenesis of viral proteins in replicating viruses has been extensively
184 pursued to understand escape mechanisms from drugs and antibodies, the work here
185 shows how deep mutagenesis can be directly applicable to therapeutic design when the
186 selection method is decoupled from virus replication and focused on host factors.

187 **METHODS**

188 **Plasmids.** The mature polypeptide (a.a. 19-805) of human ACE2 (GenBank NM_021804.1)
189 was cloned in to the NheI-XhoI sites of pCEP4 (Invitrogen) with a N-terminal HA leader
190 (MKTIIALSYIFCLVFA), myc-tag, and linker (GSPGGA). A synthetic human codon-optimized
191 gene fragment (Integrated DNA Technologies) for the RBD (a.a. 333-529) of SARS-CoV-2 S
192 (GenBank YP_009724390.1) was N-terminally fused to a HA leader and C-terminally fused
193 to superfolder GFP (*32*) and ligated in to the NheI-XhoI sites of pcDNA3.1(+) (Invitrogen).

194 **Tissue Culture.** Expi293F cells (ThermoFisher) were cultured in Expi293 Expression
195 Medium (ThermoFisher) at 125 rpm, 8 % CO_2, 37 °C. For production of RBD-sfGFP, cells
196 were prepared to 2×10^6 / ml. Per ml of culture, 500 ng of pcDNA3-RBD-sfGFP and 3 μg of
197 polyethylenimine (MW 25,000; Polysciences) were mixed in 100 μl of OptiMEM (Gibco),
198 incubated for 20 minutes at room temperature, and added to cells. Transfection Enhancers
199 (Thermo Fisher) were added 19 h post-transfection, and cells were cultured for 110 h. Cells
200 were removed by centrifugation at 800 × g for 5 minutes and medium was stored at -20 °C.
201 After thawing and immediately prior to use, remaining cell debris and precipitates were
202 removed by centrifugation at 20,000 × g for 5 minutes.

203 **Deep mutagenesis.** 117 residues within the protease domain of ACE2 were diversified by
204 overlap extension PCR (*35*) using primers with degenerate NNK codons. The plasmid
205 library was transfected in to Expi293F cells using Expifectamine under conditions
206 previously shown to typically give no more than a single coding variant per cell (*30*, *31*); 1
207 ng coding plasmid was diluted with 1,500 ng pCEP4-ΔCMV carrier plasmid per ml of cell
208 culture at 2×10^6 / ml, and the medium was replaced 2 h post-transfection. The cells were
209 collected after 24 h, washed with ice-cold PBS-BSA, and incubated for 30 minutes on ice
210 with a 1/20 (replicate 1) or 1/40 (replicate 2) dilution of medium containing RBD-sfGFP
211 into PBS supplemented with 0.2 % bovine serum albumin (PBS-BSA). Cells were co-stained
212 with anti-myc Alexa 647 (clone 9B11, 1/250 dilution; Cell Signaling Technology). Cells
213 were washed twice with PBS-BSA, and sorted on a BD FACS Aria II at the Roy J. Carver
214 Biotechnology Center. The main cell population was gated by forward/side scattering to
215 remove debris and doublets, and DAPI was added to the sample to exclude dead cells. Of
216 the myc-positive (Alexa 647) population, the top 67% were gated (Fig. 1B). Of these, the 15
217 % of cells with the highest and 20% of cells with the lowest GFP fluorescence were
218 collected (Fig. 1D) in tubes coated overnight with fetal bovine serum and containing
219 Expi293 Expression Medium. Total RNA was extracted from the collected cells using a
220 GeneJET RNA purification kit (Thermo Scientific), and cDNA was reverse transcribed with
221 high fidelity Accuscript (Agilent) primed with gene-specific oligonucleotides. Diversified
222 regions of ACE2 were PCR amplified as 5 fragments. Flanking sequences on the primers
223 added adapters to the ends of the products for annealing to Illumina sequencing primers,
224 unique barcoding, and for binding the flow cell. Amplicons were sequenced on an Illumina

225 NovaSeq 6000 using a 2×250 nt paired end protocol. Data were analyzed using Enrich (*36*),
226 and commands are provided in the GEO deposit. Briefly, the frequencies of ACE2 variants in
227 the transcripts of the sorted populations were compared to their frequencies in the naive
228 plasmid library to calculate an enrichment ratio.

229 **Reagent and data availability.** Plasmids are deposited with Addgene under IDs 141183-5.
230 Raw and processed deep sequencing data are deposited in NCBI's Gene Expression
231 Omnibus (GEO). At this time, a series accession number has not been assigned.

232 **ACKNOWLEDGEMENTS.** Staff at the UIUC Roy J. Carver Biotechnology Center assisted with
233 FACS and Illumina sequencing. The development of deep mutagenesis to study virus-
234 receptor interactions was supported by NIH award R01AI129719.

235 **CONFLICT OF INTEREST STATEMENT.** E.P. is the inventor on a provisional patent filing
236 by the University of Illinois covering aspects of this work.

237 **REFERENCES**

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

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

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

244 4. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses,
245 The species Severe acute respiratory syndrome-related coronavirus: classifying
246 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*. **4**, 3 (2020).

247 5. A. Patel, D. B. Jernigan, 2019-nCoV CDC Response Team, Initial Public Health
248 Response and Interim Clinical Guidance for the 2019 Novel Coronavirus Outbreak -
249 United States, December 31, 2019-February 4, 2020. *MMWR Morb. Mortal. Wkly. Rep.*
250 **69**, 140–146 (2020).

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

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

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

257 9. Y. Wan, J. Shang, R. Graham, R. S. Baric, F. Li, Receptor recognition by novel
258 coronavirus from Wuhan: An analysis based on decade-long structural studies of

259 SARS. *J. Virol.* (2020), doi:10.1128/JVI.00127-20.

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

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

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

267 13. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor
268 usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol*. **11**, 1860
269 (2020).

270 14. M. A. Tortorici, D. Veesler, Structural insights into coronavirus entry. *Adv. Virus Res.*
271 **105**, 93–116 (2019).

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

275 16. I. G. Madu, S. L. Roth, S. Belouzard, G. R. Whittaker, Characterization of a highly
276 conserved domain within the severe acute respiratory syndrome coronavirus spike
277 protein S2 domain with characteristics of a viral fusion peptide. *J. Virol.* **83**, 7411–
278 7421 (2009).

279 17. A. C. Walls *et al.*, Tectonic conformational changes of a coronavirus spike
280 glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11157–
281 11162 (2017).

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

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

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

289 21. H. Hofmann *et al.*, Susceptibility to SARS coronavirus S protein-driven infection
290 correlates with expression of angiotensin converting enzyme 2 and infection can be
291 blocked by soluble receptor. *Biochem. Biophys. Res. Commun.* **319**, 1216–1221
292 (2004).

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

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

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

301 25. M. Haschke *et al.*, Pharmacokinetics and pharmacodynamics of recombinant human
302 angiotensin-converting enzyme 2 in healthy human subjects. *Clin Pharmacokinet*. **52**,
303 783–792 (2013).

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

306 27. P. Towler *et al.*, ACE2 X-ray structures reveal a large hinge-bending motion
307 important for inhibitor binding and catalysis. *J. Biol. Chem.* **279**, 17996–18007
308 (2004).

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

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

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

316 31. J. Park *et al.*, Structural architecture of a dimeric class C GPCR based on co-trafficking
317 of sweet taste receptor subunits. *Journal of Biological Chemistry*. **294**, 4759–4774
318 (2019).

319 32. J.-D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger, G. S. Waldo, Engineering and
320 characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79–
321 88 (2006).

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

324 34. J. D. Heredia, J. Park, H. Choi, K. S. Gill, E. Procko, Conformational Engineering of HIV-
325 1 Env Based on Mutational Tolerance in the CD4 and PG16 Bound States. *J. Virol.* **93**,
326 e00219–19 (2019).

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

329 36. D. M. Fowler, C. L. Araya, W. Gerard, S. Fields, Enrich: software for analysis of protein
330 function by enrichment and depletion of variants. *Bioinformatics*. **27**, 3430–3431
331 (2011).

332