Diversity of ACE2 and its interaction with SARS-CoV-2 receptor binding domain

Jessie Low-Gan[1,2], Ruiqi Huang[2], Gabrielle Warner[2], Abigail Kelley[2], Duncan McGregor[2], Vaughn Smider[2,3]

[1]San Diego Jewish Academy, San Diego, CA 92130

[2]Applied Biomedical Science Institute, San Diego, CA 92127

[3]The Scripps Research Institute, La Jolla, CA 92037

Correspondence: vaughn.smider@absinstitute.org

Abstract

COVID-19, the clinical syndrome caused by the SARS-CoV-2 virus, has rapidly spread globally causing millions of infections and hundreds of thousands of deaths. The potential animal reservoirs for SARS-CoV-2 are currently unknown, however sequence analysis has provided plausible potential candidate species. SARS-CoV-2 binds to the angiotensin I converting enzyme 2 (ACE2) to enable its entry into host cells and establish infection. We analyzed the binding surface of ACE2 from several important animal species to begin to understand the parameters for the ACE2 recognition by the SARS-CoV-2 spike protein receptor binding domain (RBD). We employed Shannon entropy analysis to determine the variability of ACE2 across its sequence and particularly in its RBD interacting region, and assessed differences between various species' ACE2 and human ACE2. As cattle are a known reservoir for coronaviruses with previous human zoonotic transfer, and has a relatively divergent ACE2 sequence, we compared the binding kinetics of bovine and human ACE2 to SARS-CoV-2 RBD. This revealed a nanomolar binding affinity for bovine ACE2 but an approximate ten-fold reduction of binding compared to human ACE2. Since cows have been experimentally infected by SARS-CoV-2, this lower affinity sets a threshold for sequences with lower homology to human ACE2 to be able to serve as a productive viral receptor for SARS-CoV-2.

Introduction

The novel coronavirus SARS-CoV-2 which causes COVID-19 is a zoonotic disease[1-3] that has thus far resulted in over one million deaths worldwide and over 42 million infections [4]. The virus crossed the species boundary, possibly from bats and potentially through an intermediate species, to humans and has spread through respiratory route across the globe in the past year. SARS-CoV-2 is a member of the betacoronavirus genera, of which other coronaviruses like SARS-CoV that caused severe acute respiratory syndrome in a pandemic in 2002-2004 (resulting in approximately 800 deaths in 37 countries), and MERS-CoV which caused Middle Eastern Respiratory Syndrome in 2012 (resulting in over 800 deaths in 27 countries). Both of these viruses originated in other species, SARS-CoV in bats and MERS-CoV in camels, and crossed the species barrier to humans. Other coronaviruses like HCoV-NL63, HCoV-229E, HCoV-HKU1, and HCoV-O43 are known to cause mild respiratory disease in humans and also crossed the species barrier[5-7]. Interestingly, the latter two viruses which cause more clinically mild disease appear to have derived from bovine coronavirus (BCoV)[8-10], which causes respiratory and intestinal disease in cattle [7,11-13]. Betacoronaviruses as a group have a wide host range including several agricultural and companion animal species[8]. Thus, coronavirus disease in humans appears to largely be through zoonotic transfer and has resulted in enormous worldwide morbidity and mortality as well as significant economic loss.

Given that coronaviruses appear to have crossed the species barrier several times in human history with devastating consequences, there is a need to understand the interactions between coronaviruses and their receptors in different animal species. This is particularly important since humans interact with companion animals as well as many agricultural species, often in close quarters where respiratory spread can easily occur. In addition to the impact of spread from species to species on human health, coronaviruses can cause devastating effects to animals resulting in morbidity, mortality, and major economic losses [5,7,11,14]. Indeed SARS-CoV-2 has been documented to infect dogs (*Canis lupus*), cats (*Felis catus*) [15], and mink (*Mustela lutrola*) [16] in nature, and ferrets (*Mustela putoris)* [15], hamsters (*Cricetulus griseus*) [17] and cows (*Bos taurus*) [18] have been experimentally infected. Human to tiger transfer occurred at the Bronx Zoo [19,20]. A mink farm in the Netherlands was ravaged by SARS-CoV-2 infection and farms in the U.S. and Spain were also infected [16,21]. Additionally, SARS-CoV-2 is now studied *in vivo* in hamsters (*Cricetulus griseus*) and ferrets (*Mustela putoris*) as animal models to understand viral pathology and evaluate therapeutics. The need to understand animal susceptibility to coronavirus infection is therefore important to public health, the economy, as well as to establish well understood animal models for therapeutic and vaccine development.

65 SARS-CoV-2 utilizes its trimeric spike protein to bind to the angiotensin I converting enzyme 2 (ACE2)
66 on target pneumocytes or other host cells [22-28]. This interaction is high affinity, and results in viral
67 membrane fusion to the host cell and initiates the infectious process. SARS-CoV-1 also utilizes ACE2
68 as a receptor, however their spike proteins bind with lower affinity (31 nM K_D) than SARS-CoV-2 (4.2
69 nM K_D)[23]. Crystal structure and electron microscopy analysis of SARS-CoV-2 with ACE2 has
70 revealed the interacting amino acid residues of the spike receptor binding domain (RBD) and the
71 human ACE2 surface[23,25,29,30]. With sequences available for many companion and agriculturally
72 important species, the ability to assess potential spike RBD binding is an important first step towards
73 prediction of infection of these alternative hosts. Here we analyze the conservation and diversity of the
74 ACE2 protein in multiple important species, with particular emphasis on the region that interacts with
75 SARS-CoV-2 spike RBD. We confirm that an ACE2 from *Bos taurus*, which is somewhat divergent
76 from human ACE2 in its binding region, interacts with SARS-CoV-2 spike RBD with high affinity,
77 suggesting that multiple mammalian species may be susceptible to infection with this coronavirus.

78

79 **Results**

80 In order to assess the potential of SARS-CoV-2 to interact with ACE2 of important companion and
81 agricultural species, we assembled the ACE2 sequences of several species including (Table 1 and
82 Supplemental Table 1). The human ACE2/RBD cocrystal structure (PDB: 6M17)[27] was utilized to
83 visualize interacting residues between the SARS-CoV-2 spike RBD and human ACE2 (Figure 1). In
84 ACE2, 8 residues on two helices make direct contact with spike RBD and an additional 17 residues are
85 within 5 angstroms of the RBD. These 25 residues are color coded in Figure 1 as purple (contact) and
86 cyan (nearby), respectively. We focused our evaluation on these interacting residues in the following
87 analyses.

88 In order to determine the overall differences between ACE2 of different species, we employed protein
89 sequence alignment as well as variability analysis (Supplemental Figures 1 and 2). All vertebrate
90 ACE2 sequences showed significant homology to human ACE2, with cynomologous and rhesus
91 monkeys being 95% and 94% identical to human (Table 1). Other mammals were between 80-87%
92 identical to human. Horseshoe bats, a potential reservoir for SARS-CoV-2 [28,31] was only 81% and
93 76% identical through the entire sequence and interacting residues, respectively. Surprisingly, the
94 percent identity across the entire ACE2 sequence did not fully correlate with percent identity of the
95 interacting residues. For example, cows show lower homology throughout the entire ACE2 sequence
96 at only 79% compared to the other species, but is 84% identical within the identified interacting

97 residues. In contrast, dogs are 83.3% identical across the entire sequence, but only 76% identical in
98 the interacting residues. A similar lower identity in binding site residues is also seen for rabbits (Table
99 1). The relative contribution of the interacting residues versus residues outside the RBD binding site is
100 currently not known although it is expected that the interacting residues are far more important to
101 infection relative to the residues outside of the RBD binding site.

102 For variability analysis, the structural importance of protein regions, and even individual amino acid
103 residues, can be compared across multiple species. Such diversity analyses initially identified the
104 complementary determining regions within antibodies as important interacting domains with antigen by
105 Kabat and Wu[32,33], and more recently Shannon entropy evaluation has been employed to identify
106 conserved and diverse domains of multiple proteins through multiple sequence analysis[34,35]. First,
107 we aligned the ACE2 sequences and determined their percent identities (Table 1 and Supplemental
108 Table 1). Then, we calculated Shannon entropy (SE) across the ACE2 sequences. Of note, ACE2 is
109 remarkably conserved across its sequence, with few residues exhibiting high variability (Figure 2).

110 Within the interacting residues, 21 of 25 are highly conserved, with SE values below 2. Significantly, no
111 residues had values above 3. The more variable residues are red and conserved residues blue in
112 Figure 2. Five residues, N330, G352, D355, R357, and R393 are completely conserved, with SE
113 values of zero (Figure 2 and Supplemental Table 1). Four of these are nearby residues with RBD, with
114 only R357 being a contact residue. The complete conservation of these residues suggests that they
115 play an important role in the protease function or structural integrity of ACE2. Residues with values
116 between 0 and 1 are K353, which is a contact residue, and L45, Y83, T324, which are nearby residues
117 (Supplemental Table 1). Amino acids with SE values between 1-2 are K26, T27, D30, K31, E35,
118 D38, Q42, M82, and G354. The most diverse residues, with SE values over 2, are Q24, H34, L79, and
119 Q81. Of these, Q24 and H34 are contact residues, with H34 found in a central location in the ACE2-
120 RBD interface (Figure 1, B), and having by far the highest SE value at 2.88 . Others have analyzed the
121 evolution of ACE2 residues and have found positions 24 and 34 to be undergoing positive evolutionary
122 selection pressure [36], and suggested that these positions could play a role in predicting infectivity by
123 SARS-CoV-2 [36].

124 Since the interacting residues are likely most important for viral interaction with ACE2 on the host cell,
125 we evaluated the residues that differed between the various species' ACE2 and human ACE2 in this
126 region (Figure 3 and Supplemental Figure 3). As mentioned, horseshoe bat (*Rhinolophus sinicus*)
127 shares only 19/25 interacting residues (76%) with human ACE2, and only 5/8 contact residues.
128 Specifically, D30E, H34T, Y41H, and M82N (contact residues), and T27M, E35K (nearby residues) are
129 mutated in horseshoe bat ACE2 relative to human ACE2 (Figure 3, upper left), suggesting potentially

130 lower affinity for spike RBD. Pangolin, a possible intermediate host of SARS-CoV-2, has only one
131 contact residue (M82N) and three nearby residues (D38E, L79I, and G354H) altered, and both D38E
132 and L79I are conservative changes (Figure 3, upper right). Felines, which have had documented
133 natural infection [15,20], have only two active site residues mutated (D38E and M82T), and only M82T
134 is a contact residue (Figure 3, middle left). Dogs, which also have been infected naturally, have similar
135 binding site residues as cat but with the notable exception of H34Y (Figure 3, middle), a residue
136 reported to be important in binding. Cow has three interaction site residues mutated, of which M82T is
137 a contact residue, and pig has 5 mutations where Q24L, D30E, H34L, M82T are contact residues
138 (Figure 3 bottom).

139 Cows serve as a reservoir for bovine coronavirus (BCoV) a respiratory infection of cattle that is a
140 betacoronavirus [11,12] distantly related to SARS-CoV-2[7]. Of considerable note, two BCoV-related
141 coronaviruses, HCoV-OC43 and HCoV-HKU1, have crossed the species barrier, with OC43 likely from
142 cows to humans to cause “common cold” respiratory disease in humans[37]. Whereas BCoV, HCoV-
143 OC43 and HCoV-HKU1 utilize 9-*O*-acetylated sialoglycans as cellular receptors[38], and SARS-CoV-2
144 utilizes ACE2, cows are a potential important species to evaluate for possible SARS-CoV-2 infection as
145 they are a known coronavirus reservoir and they could theoretically serve as host for SARS-CoV-2 and
146 other coronaviruses like BCoV, which could potentially provide a host for coronavirus recombination,
147 selection, and evolution. From a biochemical standpoint, cows have a more distantly related ACE2
148 protein compared to many other vertebrates (78.8% compared to most other species which are over
149 80%), however their binding site residues are more conserved (84%)(Table 1). Therefore, it would be
150 useful to know whether the lower homology across the entire ACE2 sequence prohibits productive
151 ACE2/RBD interaction. To address this question, we expressed human and bovine ACE2 as antibody
152 Fc fusion proteins and compared their interaction with SARS-CoV-2 by enzyme linked immunosorbent
153 assay (ELISA) and further quantified their binding kinetics by surface plasmon resonance analysis
154 (Figure 4). By ELISA, bovine ACE2 had an approximately ten-fold worse binding EC_{50} (0.129 nM for
155 human ACE2 compared to 1.299 nM for bovine ACE2)(Figure 4, A). This lower apparent affinity for
156 bovine ACE2 was confirmed by surface plasmon resonance analysis which showed a K_D for bovine
157 ACE2 of 36.25 nM versus 7.5 nM for human ACE2 (Figure 4, B and Supplementary Table 1). This
158 dissociation constant difference relates primarily to a faster off-rate for bovine ACE2 compared to
159 human ACE2 (Supplementary Figure 5). Of note, despite this lower K_D for bovine ACE2, this affinity is
160 very similar to the K_D for human ACE2 for the RBD of SARS-CoV-1. Interestingly, for both human and
161 bovine ACE2, the SPR data fit more consistently with a two-site model for interaction, suggesting that
162 the RBD may be multimerizing to produce avidity effects on the chip surface. A potential second site
163 would have ten-fold lower K_D values (Supplemental Figure 5), which interestingly, are more in line with

164 the EC_{50} values of the ELISA (Figure 4, A). Such interactions would be important to explore to
165 understand the details of interaction between coronavirus spike RBD and ACE2 on the cell surface.
166 These interactions would be important to inhibit by therapeutic agents, for example by monoclonal
167 antibodies targeting the virus. The avidity interaction may decrease the K_D (increase the affinity) of
168 bovine ACE2 with spike RBD from 36 nM to 2.5 nM, and 7.5 nM to 0.4 nM for human ACE2, a
169 substantial enhancement of the interaction. Regardless of the mechanisms of interaction, it is clear that
170 bovine ACE2 still has high affinity towards SARS-CoV-2 RBD, albeit with five to ten fold worse binding
171 than human ACE2, and can still apparently mediate infection of bovine cells[18].

172

173 **Discussion**

174 The recent pandemic has spread rapidly across the globe through human populations, but additionally
175 has also infected several animals. While zoonotic in origin, it is still unclear which species provided the
176 reservoir for transfer to humans. Coronaviruses as a group have a very wide range of host species,
177 and have jumped the species barrier multiple times to humans[39,40]. While bats appear to be a host
178 species for coronaviruses related to SARS-CoV and SARS-CoV-2, it is possible that an as yet
179 unidentified species serves as the reservoir for this virus. Additionally, it is clear that SARS-CoV-2 can
180 naturally infect other species, such as cats, dogs, and mink have been observed naturally. In order to
181 understand the host range of SARS-CoV-2 as well as identify potential reservoirs for the virus it is
182 critically important to understand (i) details about the identity and binding properties between the virus
183 and its host cell receptor and possible co-receptors, (ii) other biological requirements needed for the
184 virus to replicate and transmit, for example host cell enzymes needed for viral processing, replication,
185 and assembly. For SARS-CoV-2, ACE2 appears to be the major receptor required for cell entry, so
186 understanding its interaction with ACE2 from humans as well as other species is important to enable
187 predictive methods for viral host range. Here we analyze ACE2 diversity across several species, and
188 specifically evaluate binding of SARS-CoV-2 RBD to bovine ACE2, finding key interacting residues to
189 be highly conserved.

190 Several studies have used sequence homology and/or structural modeling to attempt to predict SARS-
191 CoV-2 RBD binding to various species' ACE2 in order to predict the virus host range[15,36,39-43]. In
192 an effort to predict species permissive to infection, Damas et.al. developed a five-tiered scoring scheme
193 based on percent identity of 410 vertebrate species, as well as specific structural features of SARS-
194 CoV-1 or SARS-CoV-2 interactions with ACE2[36]. They also focused on 25 amino acid residues at
195 the RBD binding interface, however their 25 residues differed from ours in that they included S16, N53,

196 N90, and N322, with the asparagines included as potential glycosylation sites that may impact RBD
197 binding. However, our approach was agnostic in choosing residues that were either (i) known contact
198 residues with the RBD, or (ii) within 5 angstroms of contact residues. Residues included in our analysis
199 which were not included in Damas et.al. were K26, Q81, and N352. Damas et. al. also note that the
200 host range of SARS-CoV-2 might be quite broad and suggest new species that should be evaluated for
201 animal models of virus infection, which notably include cows, which scored in their "medium" category
202 for predictive binding to RBD. Like our assessment of homology, they find that bats score very low in
203 predicted ACE2 binding. Interestingly, using sequence evolution and selection analysis, they identify
204 Q24 and H34 as positions undergoing positive selection and evolution. We identified these as amongst
205 the most variable residues by Shannon entropy analysis, and these also appear to be important
206 residues at the ACE2/RBD interface. As in our analysis, pangolins scored low in potential RBD binding
207 based on interacting residue homology, suggesting that SARS-CoV-2 may bind other pangolin
208 receptors, or have other mechanisms to interact with ACE2. As pangolins are thought to be a possible
209 intermediate host for SARS-CoV-2, much more biochemical and infectivity data with this controversial
210 species should be obtained.

211 In a different approach, Lam et.al. used structural modeling to predict the change in free energy, $\Delta\Delta G$,
212 for 215 ACE2 sequences derived from different species[44]. They correlated the $\Delta\Delta G$ with published
213 infectability information to provide a framework to predict which species may be susceptible to SARS-
214 CoV-2 infection. Like other studies, their work suggests a broad range of mammal susceptibility, with
215 the exception of non-placental mammals. They similarly find that horseshoe bats have higher $\Delta\Delta G$
216 values (*i.e.* lower affinity), calling into question their susceptibility and potential as a reservoir for SARS-
217 CoV-2.

218 There have been several studies that have measured the K_D between SARS-CoV-2 spike (or RBD) and
219 ACE2, with values ranging from 1.2 to 44 nM [23,25,29,30]. In these studies, differences in the
220 experimental conditions, such as whether ACE2 or RBD was immobilized, and technique used such as
221 biolayer interferometry versus surface plasmon resonance, could account for differences in kinetic
222 values. In our study, we immobilized ACE2 and measured RBD interaction as the analyte using
223 surface plasmon resonance, which is similar to Lan et. al. who reported a value of K_D = 4.6 nM, which is
224 close to our value of 7.5 nM. However, Lan et.al. as well as all of the other studies utilized a 1:1 model
225 for binding, whereas we found evidence for a two site complexation on the surface, and applied a two
226 site model that gave K_{D1} and K_{D2} values of 7.5 nM and 0.4 nM, respectively. Of considerable note,
227 Forssen et.al. reanalyzed the binding data from Lan et.al. and Tian et.al. and found evidence of
228 multimeric interactions, similar to our study[45]. Since the RBD exists in close proximity to two other

229 RBD subunits in the spike protein trimer, it is plausible that cooperative interactions between RBDs
230 exist during ACE2 interaction events. Such events could considerably enhance the interactions
231 providing avidity to enable infection. In this regard, ACE2 variants from species with mutations in the
232 RBD interface may still be permissive to infection because of the cooperative and enhanced avidity of
233 binding events.

234 One of the major challenges in developing predictive methods for viral infectivity is the limited amount
235 of both infection and RBD-ACE2 biochemical interaction data for most species. For example, only few
236 species have had documented infection in the real world, or even in well-controlled experimental
237 systems[15,36,40,44]. In this regard, detailed biochemical analysis of multiple species' receptors would
238 provide valuable information that could be used in predictive modeling studies. With the decreased
239 costs of synthetic DNA and high-throughput screening approaches, such analyses could potentially be
240 accomplished rapidly. As a first step in this process, here we evaluate cow ACE2 for binding SARS-
241 CoV-2 RBD by both ELISA and surface plasmon resonance, a species with documented experimental
242 infection by SARS-CoV-2 and of significant agricultural importance that serves as a reservoir for other
243 betacoronaviruses, including at least one, HCoV-O43, that has jumped the species barrier to become
244 an endemic respiratory virus in humans. The affinity of bovine ACE2 is nearly ten-fold worse than
245 human ACE2 in binding the coronavirus RBD, however binding is still in the mid-nanomolar range. Of
246 potential importance is the possible cooperativity in binding that may occur between ACE2 and RBD,
247 which we observed in surface plasmon resonance analysis which suggested a two-site model for
248 binding. Interestingly, this data was consistent with ELISA binding data, suggesting K_D in the
249 subnanomolar range for human, and low nanomolar range for bovine ACE2. Further study into various
250 species' viral receptors, including the recently discovered neuropilin-2 co-receptor for SARS-CoV-2[46],
251 should enable much more accurate and rapid development of predictive algorithms based on viral
252 receptor sequence and structural modelling.

253

254 **Methods**

255 *Sequence and structure analysis*

256 ACE2 sequences from 19 species were obtained from the NCBI protein sequence database. A number
257 of multiple sequence alignments were performed using Clustal Omega. Specific multiple sequence
258 alignments for (i) companion animals (*Canis lupus familiaris, Felis catus, Mus musculus, Mesocricetus*
259 *auratus, Oryctolagus cuniculus*), (ii) agriculturally important animals (*Bos taurus, Capra hircus, Equus*
260 *asinus, Equus caballus, Gallus gallus, Ovis aries, and Sus scrofa*), and (iii) species relevant to animal

261 models or potential reservoirs of interest to vaccine testing research or suspected reservoirs (*Macaca*
262 *mulatta, Macaca fascicularis, Felis catus, Manis javanica, Cricetulus griseus, Canis lupus familiaris,*
263 *Mustela putoris, Sus scrofa, Rhinolophus sinicus, Bos taurus*) were investigated. In all sequence
264 alignments the *Homo sapiens* ACE2 sequence was included as a reference.

265 Further analysis on the multiple sequence alignments were carried out with a variety of software.
266 Phylogenetic trees and pairwise percent identities were obtained with Clustal Omega [47] and
267 visualized in UGENE (http://ugene.net). The open-source EMBL Boxshade Server was used to produce
268 the boxshade alignments (https://embnet.vital-it.ch/software/BOX_form.html). The Protein Variability
269 Server (http://imed.med.ucm.es/PVS/) was employed to calculate Shannon Entropy values for each
270 residue position.

271 For structural analysis, the ACE2 and SARS-CoV-2 Receptor Binding Domain (RBD) complex structure
272 files (PDB) were obtained from the RCSB PDB (PDB ID: 6M17)[27]. Eight contact residues of ACE2
273 were identified based on the co-crystal structure [27], and for our interaction residue analysis we
274 additionally added 17 ACE2 residues (nearby residues) that were within at least 5 angstroms of the
275 RBD region and classified the full list as “interaction residues”. There were a total of 25 interaction
276 residues, with 8 known RBD contact residues and 17 nearby residues. Using a *de novo* python script,
277 these residues were extracted into a separate “sequence” and an additional, specific multiple sequence
278 alignment was constructed.

279 The interaction residue multiple sequence alignments were investigated to give more focused insight
280 into which species would be at risk of infection, with the assumption that residues directly interacting
281 with the SARS-CoV-2 spike protein would be essential for infection. Structural analyses were
282 performed in Visual Molecular Dynamics (VMD, https://www.ks.uiuc.edu/Research/vmd/).

283 *Protein purification*

284 Human and bovine ACE2 proteins were produced as fusion proteins to human IgG1 Fc according to
285 our published methods for monoclonal antibody purification[2-4]. Briefly, 30 M HEK293 Freestyle cells
286 were transfected with 293fectin combined with 30 µg of pFuse-based vectors containing the human or
287 bovine ACE2 fused to the human Fc region. Cells were shaken at 37 °C for 4 days with 8% CO_2. The
288 media was clarified by centrifugation at 4000 RPM for 5 minutes followed by filtration through a 0.22 µm
289 filter. The media was concentrated and buffer-exchanged into PBS using Amicon Ultra Centrifugal Filter
290 unit (MWCO = 10,000) (MilliporeSigma) at 4 °C. The concentrated media was then loaded onto a
291 protein A-sepharose column (Cytiva) pre-equilibrated with 20 mM sodium phosphate, pH 7.0, followed
292 by washing of the column with 10 column volumes of the same buffer and eluted twice with 1 column

293 volume of 0.1 M glycine-HCl, pH 2.7 into fractions containing 0.1 column volume of 1M Tris, pH 8.
294 Purified proteins were buffer-exchanged into PBS using Amicon Ultra Centrifugal Filter unit (MWCO =
295 10,000) (MilliporeSigma), quantified using 280 nM absorbance on a Nanodrop spectrophotometer
296 (Thermo Fisher Scientific), and resolved on an SDS-PAGE stained with InstantBlue Coomassie Protein
297 Stain (Abcam).

298 The SARS-CoV-2 RBD in plasmid NR-52309 (BEI Resources) was transfected and harvested as
299 described above, but purified using TALON cobalt metal affinity resin (Takara Bio) following the
300 manufacturer's protocol, except that 50 mM, 100 mM, 200 mM and 300 mM imidazole gradient elution
301 fractions (1 column volume of each) were collected. Each elution fraction was resolved on an SDS-
302 PAGE gel stained with InstantBlue Coomassie Protein Stain (Abcam), and fractions containing a single
303 RBD band were pooled, buffer-exchanged into PBS and quantified as described above. Purified
304 protein was resolved on SDS-PAGE with Coomassie (Gel-code blue) staining.

305 *Enzyme linked immunosorbent assay (ELISA)*

306 ELISA plates (Corning 3690) were coated with 100 ng soluble SARS-CoV-2 RBD at 4 ºC in 1 × PBS
307 (pH 7.4) overnight. Plates were washed three times with Tris buffered saline pH 7.4 containing 0.1%
308 Tween-20 (TBST) and blocked with 2% milk (Marvel, dried skim milk dissolved in TBST) at room
309 temperature for 1 h. Serially diluted human or bovine ACE2-Fc were added to the wells in 2%
310 milk/TBST, plates were incubated at room temperature for 1 h, washed four times with TBST and then
311 goat anti-human Fc-HRP (Jackson ImmunoResearch #109-035-098, diluted 1: 5000 in 2% milk/TBST)
312 was added to the wells. Plates were incubated at room temperature for 30 minutes and washed five
313 times with TBST, then developed by adding 50 µl TMB substrate solution (Thermo Scientific) per well
314 and incubated at room temperature for 3 minutes. The HRP-TMB reaction was stopped by adding 50 µl
315 1.0 N sulfuric acid per well. The optical density at 450 nm was read on a microplate reader
316 (SPECTRAMAX M2, Molecular Devices). Antigen-binding curves and EC_{50} values were generated and
317 calculated using four parameter logistic regression in GraphPad Prism 8 (GraphPad Software, Inc., San
318 Diego, CA).

319 *Surface plasmon resonance*

320 Surface plasmon resonance of immobilized ACE2 proteins to RBD as the analyte was performed by
321 Biosensor Tools (https://www.biosensortools.com/ , Salt Lake City, Utah). ACE2-Fc ligands were
322 captured onto a Protein A surface at four different surface densities. Data were collected in Hepes
323 Buffered Saline (HBS) containing 0.1 mg/ml BSA at 25 degrees C. The RBD analyte (36 kDa) was
324 tested in a 3-fold titration series up to 1 uM and 100 nM for both ACE2 protein surfaces at the 4

325 different densities. The data for each data set were fit to a two independent site model. The K_D values
326 indicated in Figure 4 are the average of four experiments at different Rmax values for 100 nM analyte.

327

328

329

330

331 **Table 1. Various species ACE2 percent identity to human ACE2.** The percent identity was
332 calculated across all amino acids of ACE2 (Full length ACE2) or only the interacting residues as shown
333 in Figure 1.

		% Sequence Identity to Human ACE2	
Species		**Full length ACE2**	**Interacting Residues**
Human	*Homo sapiens*	100	100
Cynomolgus monkey	*Macaca fascicularis*	95.03	100
Rhesus monkey	*Macaca mulatta*	94.78	100
Horse	*Equus caball*	86.71	72
Donkey	*Equus asinus*	85.57	72
Cat	*Felis catus*	85.22	80
Rabbit	*Oryctolagus cuniculus*	84.84	76
Pangolin	*Manis javanica*	84.72	68
Golden hamster	*Mesocricetus auratus*	84.47	88
Chinese hamster	*Cricetulus griseus*	84.35	88
Dog	*Canis lupus*	83.33	76
Ferret	*Mustela putorius*	82.48	68
Mouse	*Mus musculus*	82.11	68
Goat	*Capra hircus*	81.72	84
Sheep	*Ovis aries*	81.69	84
Pig	*Sus scrofa*	81.37	76
Bat	*Rhinolophus sinicus*	80.75	72
Cow	*Bos taurus*	78.8	84
Chicken	*Gallus gallus*	66.25	52

334

335

336

337

338

339

340 **Figure Legends**

341 **Figure 1. ACE2 residues interacting with SARS-CoV-2 spike RBD. (A).** The cocrystal structure for
342 the spike RBD (salmon color) and ACE2 (grey, PDB: 6M17) was used to visualize residues that contact
343 the RBD (shaded purple, Q24, D30, H34, Y41, Q42, M82, K353, R357) or are within 5 A of the RBD
344 (shaded cyan, K26, T27, F28, K31, K35, E37, D38, L45, L79, Q81, Y83, T324, N330, G352, G354, ,
345 D355, R393). Throughout the manuscript we refer to these as "contact residues" (purple) and "nearby
346 residues" (cyan), and together as "interacting residues". **(B)** Top down view of ACE2 in space filling
347 mode, with residues color-coded as in (A). G352 is not visible in this orientation. **(C)** Boxshade
348 alignment of only the RBD-interacting (contact residues) and nearby amino acid residues from (B) for
349 multiple species. Contact residues are indicated with a cyan bar on top of the sequence, and nearby
350 residues with a purple bar.

351

352 **Figure 2. Shannon entropy analysis of ACE2 shows low variability. (A)** Shannon entropy plot of
353 the entire ACE2 sequence. The residue positions, and schematic of functional domains are illustrated
354 below the plot. Few residues have values above 2.5, indicating low variability **(B)** (Left) Shannon
355 entropy values projected onto the human ACE2 structure as a heat map from blue (low) to red (high).
356 (Right) View of the RBD-interacting surface as in Figure 1, with contact and nearby residues labeled.
357 Blue residues are very highly conserved.

358

359 **Figure 3. Variable residues at the ACE2-RBD interface of individual species.** With human ACE2
360 as a reference, the variant interacting residues for each species are colored red, and conserved
361 residues colored cyan. For the remainder of the ACE2 protein, conserved residues are white and
362 variable residues light red. Certain residues like M82, Q24, D30, and H34 are often mutated relative to
363 human. D30 is conservatively changed to glutamate, however more non-conservative changes can be
364 seen for H34, for example. More species have been analyzed and are shown in Supplemental Figure
365 3.

366

367 **Figure 4. SARS-CoV-2 spike RBD binds to bovine ACE2 with lower affinity than human ACE2.**
368 **(A)** ELISA of human (red) and bovine (blue) ACE2-Fc binding to immobilized SARS-CoV-2 spike RBD.
369 **(B)** Surface plasmon resonance analysis of RBD as an analyte on immobilized Human ACE-2 Fc (top)
370 or bovine ACE2-Fc (bottom). The data was fit to a two site model (Supplemental Figure 5). The K_D for
371 site one is indicated, and site two is in parentheses.

372

373 **Supplemental Table 1. Shannon Entropy for RBD interacting residues of ACE2.**

374 **Supplemental Figure 1. Boxshade amino acid sequence alignment of multiple ACE2 species.**

375 **Supplemental Figure 2. Phylogenetic tree of ACE2 from multiple species.**

376 **Supplemental Figure 3. Structures showing variations of the ACE2-RBD interface for four**
377 **species. This analysis was identical to Figure 3.**

378 **Supplemental Figure 4. Pairwise ACE2 percent identities for multiple vertebrate species.**

Supplemental Figure 5. Surface plasmon resonance data for human and bovine ACE2-Fc binding SARS-CoV-2 RBD.

Acknowledgements

The following reagent was produced under HHSN272201400008C and obtained through BEI Resources, NIAID, NIH: Vector pCAGGS Containing the SARS-Related Coronavirus 2, Wuhan-Hu-1 Spike Glycoprotein Receptor Binding Domain (RBD), NR-52309. We thank David Myszka of Biosensor Tools for surface plasmon resonance data analysis.

Author Contributions

JG and VS designed experiments and analyzed data. JG assembled sequences and performed bioinformatic analysis and generated figures. RH designed SARS-CoV-2 RBD expression protocols and purified the RBD protein. GW and AK purified ACE2-Fc proteins and performed ELISA assays and analyzed SPR results. DM and GW designed ACE2 expression vectors. VS and JG wrote the manuscript. VS supervised the project.

Funding

This project was supported in part by National Institutes of Health grant R01 HD088400 to V.S.

References

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, et al.: **Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China**. *Lancet* 2020, **395**:497-506.

2. Wang C, Horby PW, Hayden FG, Gao GF: **A novel coronavirus outbreak of global health concern**. *Lancet* 2020, **395**:470-473.

3. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, et al.: **A Novel Coronavirus from Patients with Pneumonia in China, 2019**. *N Engl J Med* 2020, **382**:727-733.

4. **Johns Hopkins Coronavirus Resource Center**. Edited by; 2020.

5. Hulswit RJG, de Haan CAM, Bosch BJ: **Coronavirus Spike Protein and Tropism Changes**. *Advances in virus research* 2016, **96**:29-57.

6. Li F: **Receptor recognition and cross-species infections of SARS coronavirus**. *Antiviral Res* 2013, **100**:246-254.

7. Saif LJ, Jung K: **Comparative Pathogenesis of Bovine and Porcine Respiratory Coronaviruses in the Animal Host Species and SARS-CoV-2 in Humans**. *J Clin Microbiol* 2020, **58**.

8. Lang Y, Li W, Li Z, Koerhuis D, van den Burg ACS, Rozemuller E, Bosch BJ, van Kuppeveld FJM, Boons GJ, Huizinga EG, et al.: **Coronavirus hemagglutinin-esterase and spike proteins coevolve for functional balance and optimal virion avidity**. *Proc Natl Acad Sci U S A* 2020, **117**:25759-25770.

9. Vijgen L, Keyaerts E, Lemey P, Maes P, Van Reeth K, Nauwynck H, Pensaert M, Van Ranst M: **Evolutionary history of the closely related group 2 coronaviruses: porcine hemagglutinating encephalomyelitis virus, bovine coronavirus, and human coronavirus OC43**. *J Virol* 2006, **80**:7270-7274.

10. Vijgen L, Keyaerts E, Moes E, Thoelen I, Wollants E, Lemey P, Vandamme AM, Van Ranst M: **Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event**. *J Virol* 2005, **79**:1595-1604.

11. Amer HM: **Bovine-like coronaviruses in domestic and wild ruminants**. *Anim Health Res Rev* 2018, **19**:113-124.

12. Hodnik JJ, Jezek J, Staric J: **Coronaviruses in cattle**. *Trop Anim Health Prod* 2020.

13. Saif LJ: **Bovine respiratory coronavirus**. *Vet Clin North Am Food Anim Pract* 2010, **26**:349-364.

14. Hofmann H, Simmons G, Rennekamp AJ, Chaipan C, Gramberg T, Heck E, Geier M, Wegele A, Marzi A, Bates P, et al.: **Highly conserved regions within the spike proteins of human coronaviruses 229E and NL63 determine recognition of their respective cellular receptors**. *J Virol* 2006, **80**:8639-8652.

15. Shi J, Wen Z, Zhong G, Yang H, Wang C, Huang B, Liu R, He X, Shuai L, Sun Z, et al.: **Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2**. *Science* 2020, **368**:1016-1020.

16. Oreshkova N, Molenaar RJ, Vreman S, Harders F, Oude Munnink BB, Hakze-van der Honing RW, Gerhards N, Tolsma P, Bouwstra R, Sikkema RS, et al.: **SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020**. *Euro Surveill* 2020, **25**.

440 17. Rogers TF, Zhao F, Huang D, Beutler N, Burns A, He WT, Limbo O, Smith C, Song G, Woehl J, et
441 al.: **Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease**
442 **in a small animal model**. *Science* 2020.

443 18. Ulrich L, Wernike K, Hoffmann D, Mettenleiter TC, Beer M: **Experimental Infection of Cattle with**
444 **SARS-CoV-2**. *Emerg Infect Dis* 2020, **26**.

445 19. Ji W, Wang W, Zhao X, Zai J, Li X: **Cross-species transmission of the newly identified**
446 **coronavirus 2019-nCoV**. *J Med Virol* 2020, **92**:433-440.

447 20. Wang L, Mitchell PK, Calle PP, Bartlett SL, McAloose D, Killian ML, Yuan F, Fang Y, Goodman LB,
448 Fredrickson R, et al.: **Complete Genome Sequence of SARS-CoV-2 in a Tiger from a U.S.**
449 **Zoological Collection**. *Microbiol Resour Announc* 2020, **9**.

450 21. Enserink M: **SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020**.
451 *Science* 2020.

452 22. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, Schiergens TS,
453 Herrler G, Wu NH, Nitsche A, et al.: **SARS-CoV-2 Cell Entry Depends on ACE2 and**
454 **TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor**. *Cell* 2020.

455 23. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, et al.: **Structure of**
456 **the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor**. *Nature*
457 2020.

458 24. Letko M, Marzi A, Munster V: **Functional assessment of cell entry and receptor usage for**
459 **SARS-CoV-2 and other lineage B betacoronaviruses**. *Nat Microbiol* 2020, **5**:562-569.

460 25. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D: **Structure, Function, and**
461 **Antigenicity of the SARS-CoV-2 Spike Glycoprotein**. *Cell* 2020.

462 26. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, Lu G, Qiao C, Hu Y, Yuen KY, et al.: **Structural**
463 **and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2**. *Cell* 2020.

464 27. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q: **Structural basis for the recognition of SARS-CoV-**
465 **2 by full-length human ACE2**. *Science* 2020, **367**:1444-1448.

466 28. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, et al.: **A**
467 **pneumonia outbreak associated with a new coronavirus of probable bat origin**. *Nature*
468 2020, **579**:270-273.

469 29. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F: **Structural basis of**
470 **receptor recognition by SARS-CoV-2**. *Nature* 2020, **581**:221-224.

471 30. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS:
472 **Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation**. *Science* 2020,
473 **367**:1260-1263.

474 31. Zhou J, Li C, Liu X, Chiu MC, Zhao X, Wang D, Wei Y, Lee A, Zhang AJ, Chu H, et al.: **Infection of**
475 **bat and human intestinal organoids by SARS-CoV-2**. *Nat Med* 2020, **26**:1077-1083.

476 32. Kabat EA, Wu TT: **Identical V region amino acid sequences and segments of sequences in**
477 **antibodies of different specificities. Relative contributions of VH and VL genes,**
478 **minigenes, and complementarity-determining regions to binding of antibody-combining**
479 **sites**. *J Immunol* 1991, **147**:1709-1719.

480 33. Wu TT, Kabat EA: **An analysis of the sequences of the variable regions of Bence Jones**
481 **proteins and myeloma light chains and their implications for antibody complementarity**.
482 *J. Exp. Med.* 1970, **132**:211-250.

483 34. Deiss TC, Vadnais M, Wang F, Chen PL, Torkamani A, Mwangi W, Lefranc M-P, Criscitiello MF,
484 Smider VV: **Immunogenetic factors driving formation of ultralong VH CDR3 in Bos taurus**
485 **antibodies**. *Cellular and Molecular Immunology* 2017, **14**:1-12.

486 35. Strait BJ, Dewey TG: **The Shannon information entropy of protein sequences**. *Biophys J* 1996,
487 **71**:148-155.

488 36. Damas J, Hughes GM, Keough KC, Painter CA, Persky NS, Corbo M, Hiller M, Koepfli KP,
489 Pfenning AR, Zhao H, et al.: **Broad host range of SARS-CoV-2 predicted by comparative**
490 **and structural analysis of ACE2 in vertebrates**. *Proc Natl Acad Sci U S A* 2020, **117**:22311-
491 22322.

492 37. Cui J, Li F, Shi ZL: **Origin and evolution of pathogenic coronaviruses**. *Nat Rev Microbiol* 2019,
493 **17**:181-192.

494 38. Hulswit RJG, Lang Y, Bakkers MJG, Li W, Li Z, Schouten A, Ophorst B, van Kuppeveld FJM, Boons
495 GJ, Bosch BJ, et al.: **Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic**
496 **acids via a conserved receptor-binding site in spike protein domain A**. *Proc Natl Acad Sci*
497 *U S A* 2019, **116**:2681-2690.

498 39. Mathavarajah S, Dellaire G: **Lions, tigers and kittens too: ACE2 and susceptibility to COVID-**
499 **19**. *Evol Med Public Health* 2020, **2020**:109-113.

500 40. Salajegheh Tazerji S, Magalhaes Duarte P, Rahimi P, Shahabinejad F, Dhakal S, Singh Malik Y,
501 Shehata AA, Lama J, Klein J, Safdar M, et al.: **Transmission of severe acute respiratory**
502 **syndrome coronavirus 2 (SARS-CoV-2) to animals: an updated review**. *J Transl Med* 2020,
503 **18**:358.

504 41. Hayashi T, Abiko K, Mandai M, Yaegashi N, Konishi I: **Highly conserved binding region of ACE2**
505 **as a receptor for SARS-CoV-2 between humans and mammals**. *Vet Q* 2020, **40**:243-249.

506 42. Lam TT, Jia N, Zhang YW, Shum MH, Jiang JF, Zhu HC, Tong YG, Shi YX, Ni XB, Liao YS, et al.:
507 **Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins**. *Nature* 2020,
508 **583**:282-285.

509 43. Qiu Y, Zhao YB, Wang Q, Li JY, Zhou ZJ, Liao CH, Ge XY: **Predicting the angiotensin**
510 **converting enzyme 2 (ACE2) utilizing capability as the receptor of SARS-CoV-2**. *Microbes*
511 *Infect* 2020, **22**:221-225.

512 44. Lam SD, Bordin N, Waman VP, Scholes HM, Ashford P, Sen N, van Dorp L, Rauer C, Dawson NL,
513 Pang CSM, et al.: **SARS-CoV-2 spike protein predicted to form complexes with host**
514 **receptor protein orthologues from a broad range of mammals**. *Sci Rep* 2020, **10**:16471.

515 45. Forssen P, Samuelsson J, Lacki K, Fornstedt T: **Advanced Analysis of Biosensor Data for**
516 **SARS-CoV-2 RBD and ACE2 Interactions**. *Anal Chem* 2020, **92**:11520-11524.

517 46. Daly JL, Simonetti B, Klein K, Chen K-E, Williamson MK, Antón-Plágaro C, Shoemark DK, Simón-
518 Gracia L, Bauer M, Hollandi R, et al.: **Neuropilin-1 is a host factor for SARS-CoV-2 infection**.
519 *Science* 2020:eabd3072.

520 47. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC,
521 Finn RD, et al.: **The EMBL-EBI search and sequence analysis tools APIs in 2019**. *Nucleic*
522 *Acids Res* 2019, **47**:W636-W641.

523

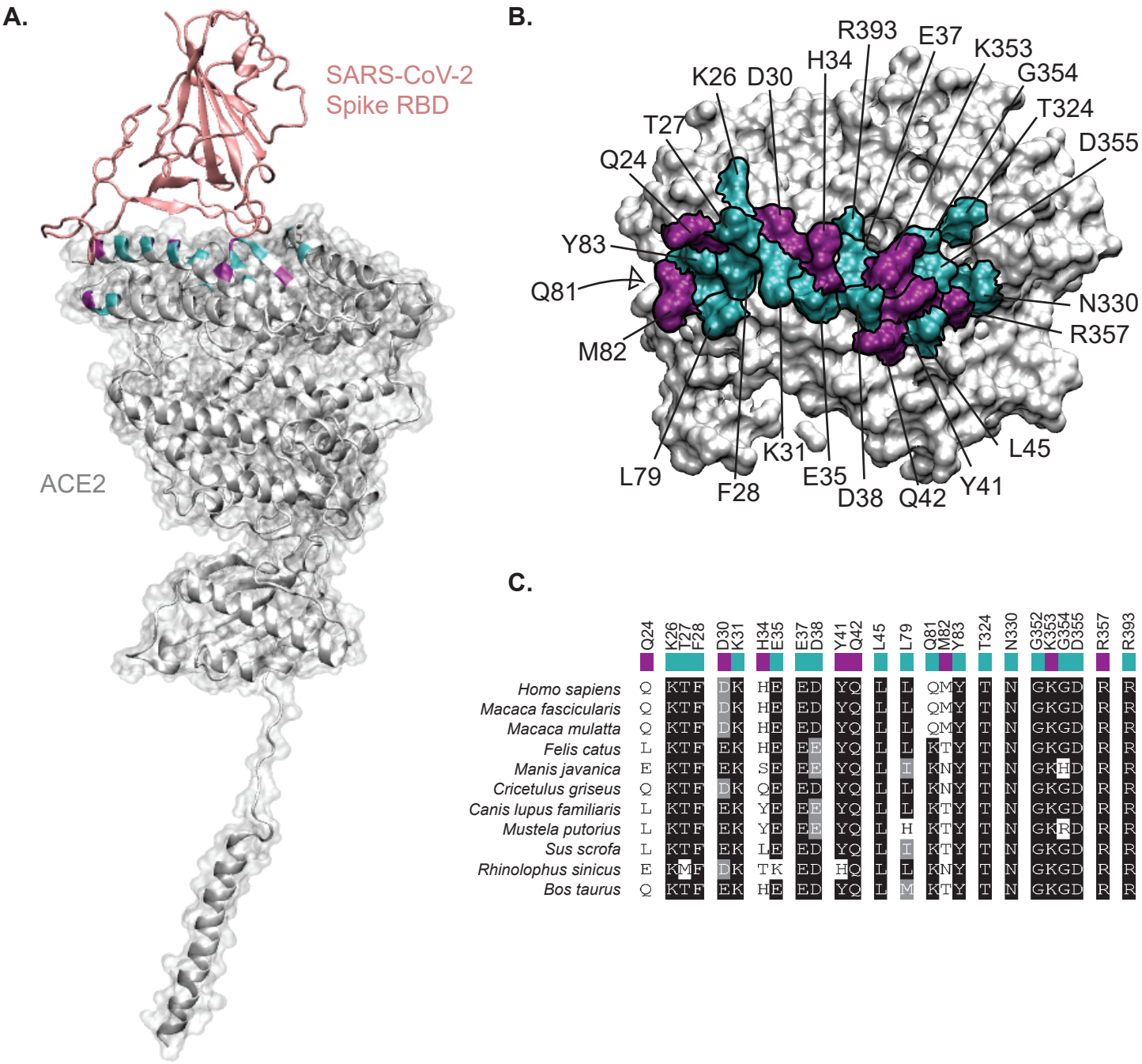

A.
SARS-CoV-2
Spike RBD
ACE2
B.
R393
E37
K353
H34
G354
K26
D30
T324
T27
D355
Q24
Y83
Q81
N330
R357
M82
L45
K31
L79
E35
Y41
F28
D38
Q42
C.
Q24
K26 T27 F28
D30 K31
H34 E35
E37 D38
Y41 Q42
L45
L79
Q81 M82 Y83
T324
N330
G352 K353 G354 D355
R357
R393
Homo sapiens Q KTF DK HE ED YQ L L QMY T N GKGD R R
Macaca fascicularis Q KTF DK HE ED YQ L L QMY T N GKGD R R
Macaca mulatta Q KTF DK HE ED YQ L L QMY T N GKGD R R
Felis catus L KTF EK HE EE YQ L L KTY T N GKGD R R
Manis javanica E KTF EK SE EE YQ L I KNY T N GKHD R R
Cricetulus griseus Q KTF DK QE ED YQ L L KNY T N GKGD R R
Canis lupus familiaris L KTF EK YE EE YQ L L KTY T N GKGD R R
Mustela putorius L KTF EK YE EE YQ L H KTY T N GKRD R R
Sus scrofa L KTF EK LE ED YQ L I KTY T N GKGD R R
Rhinolophus sinicus E KMF DK TK ED HQ L L KNY T N GKGD R R
Bos taurus Q KTF EK HE ED YQ L M KTY T N GKGD R R

Figure 1

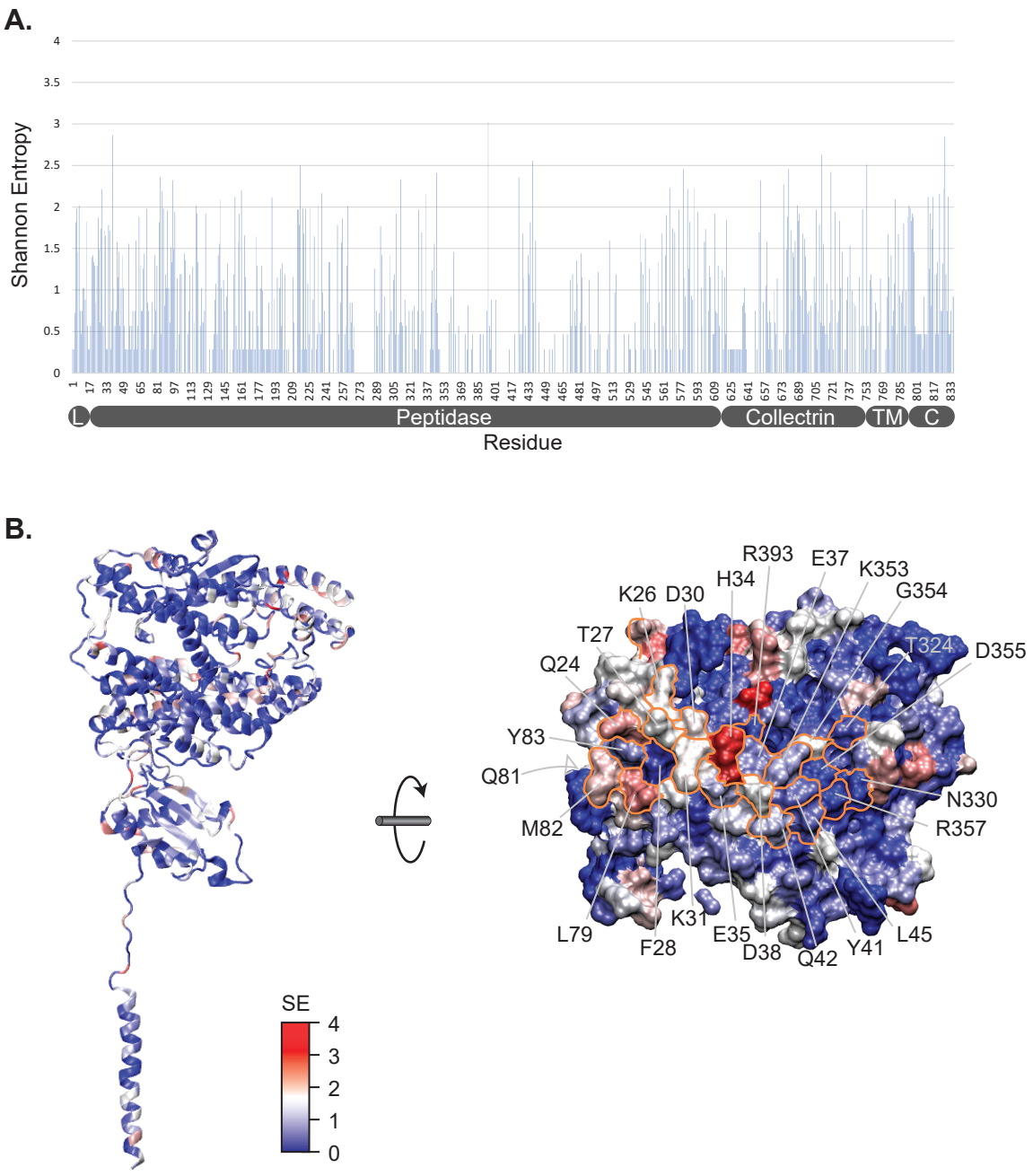
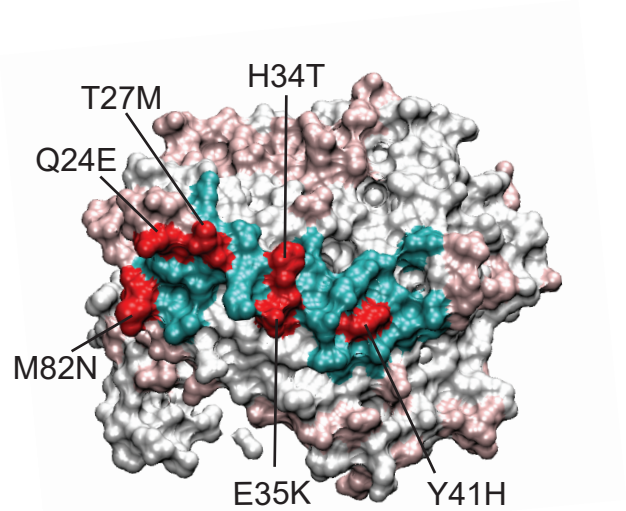
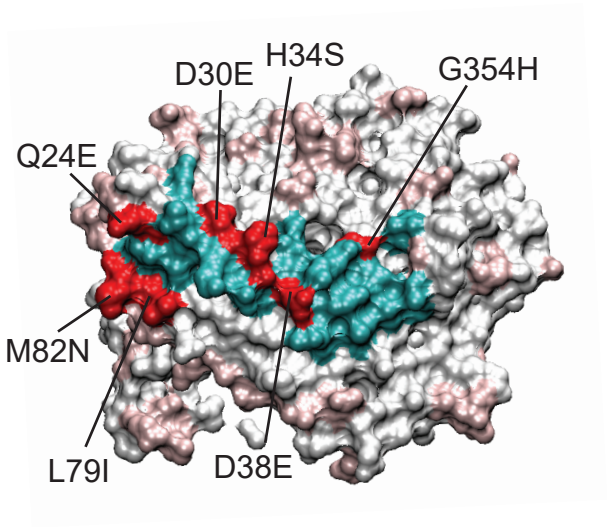

A.
Shannon Entropy
Residue
L
Peptidase
Collectrin
TM
C
B.
SE
4
3
2
1
0
R393
E37
K353
H34
K26
D30
G354
T27
T324
D355
Q24
Y83
Q81
N330
M82
R357
K31
L79
E35
L45
F28
D38
Y41
Q42

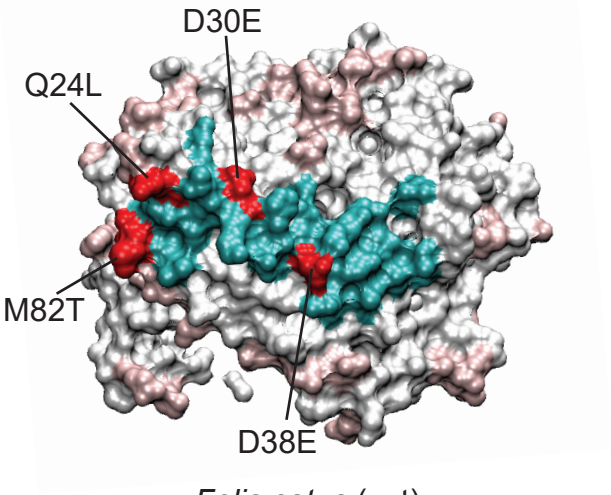
Figure 2



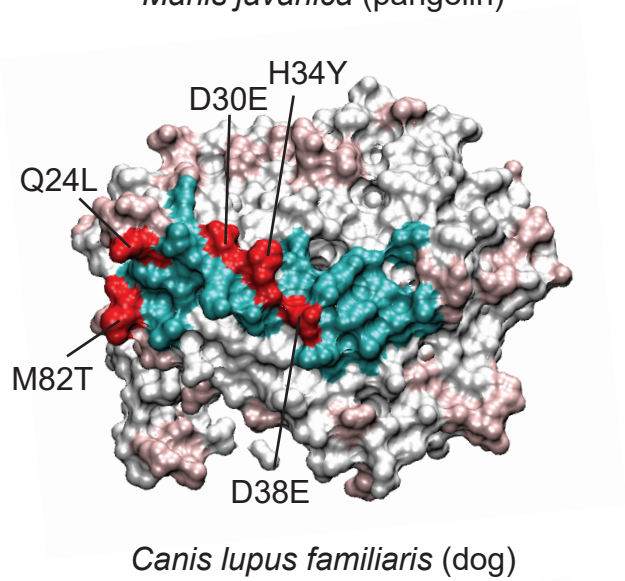
Rhinolophus sinicus (horseshoe bat)



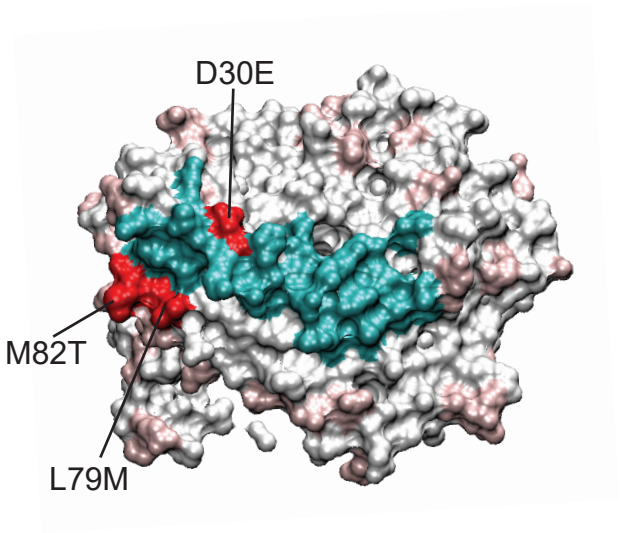
Manis javanica (pangolin)



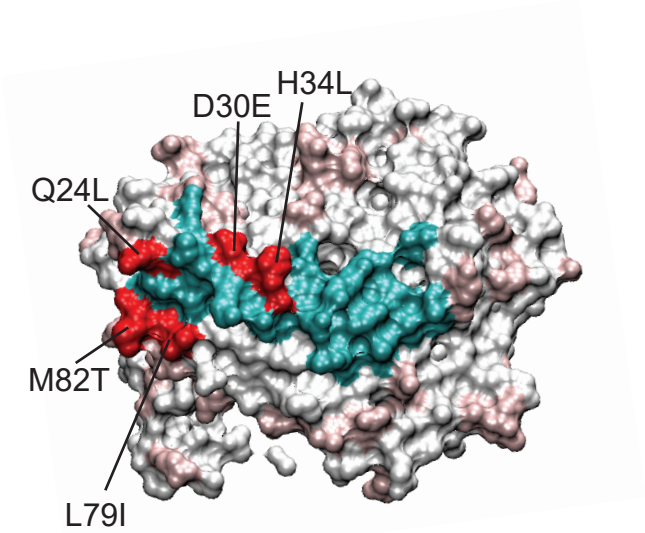
Felis catus (cat)



Canis lupus familiaris (dog)



Bos taurus (cow)



Sus scrofa (pig)

Figure 3

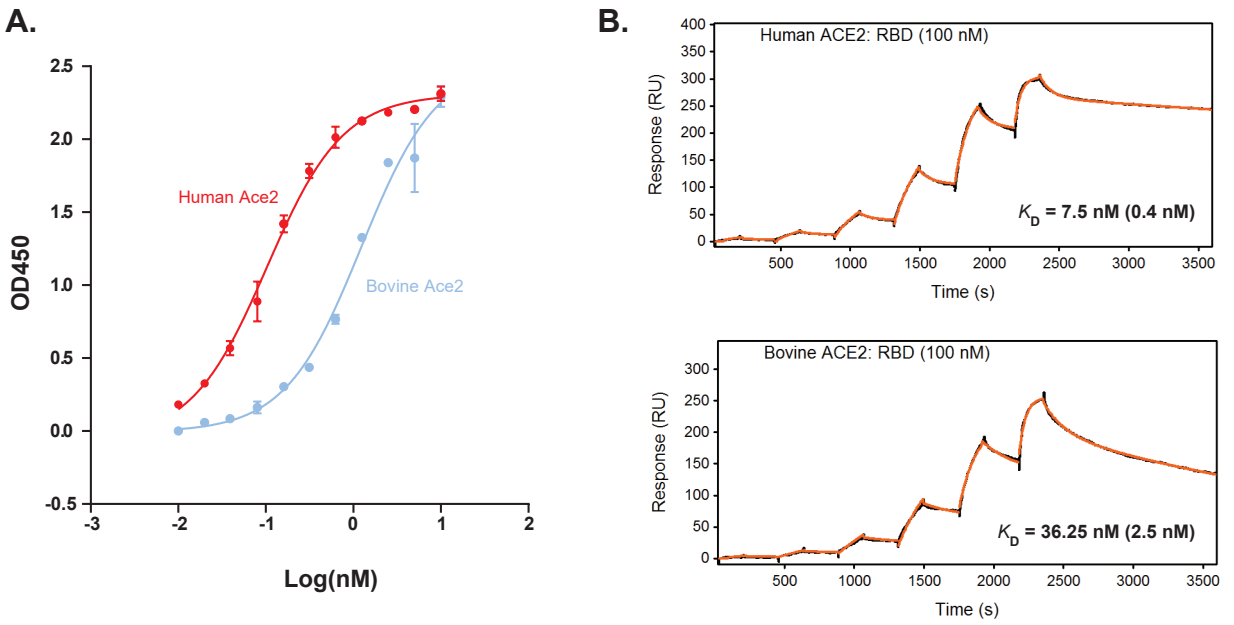

A.
OD450
Log(nM)
Human Ace2
Bovine Ace2
B.
Human ACE2: RBD (100 nM)
Response (RU)
Time (s)
K_D = 7.5 nM (0.4 nM)
Bovine ACE2: RBD (100 nM)
Response (RU)
Time (s)
K_D = 36.25 nM (2.5 nM)

Figure 4