



HAL
open science

Molecular docking simulation reveals ACE2 polymorphisms that may increase the affinity of ACE2 with the SARS-CoV-2 Spike protein

Matteo Calcagnile, Patricia Forgez, Antonio Iannelli, Cecilia Bucci, Marco Alifano, Pietro Alifano

► To cite this version:

Matteo Calcagnile, Patricia Forgez, Antonio Iannelli, Cecilia Bucci, Marco Alifano, et al.. Molecular docking simulation reveals ACE2 polymorphisms that may increase the affinity of ACE2 with the SARS-CoV-2 Spike protein. *Biochimie*, 2021, 180, pp.143 - 148. <10.1016/j.biochi.2020.11.004>. <hal-03492758>

HAL Id: hal-03492758

<https://hal.science/hal-03492758v1>

Submitted on 21 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

1 **Manuscript type:** Short Communication

2

3 **Molecular docking simulation reveals ACE2 polymorphisms that may increase the affinity of**
4 **ACE2 with the SARS-CoV-2 Spike protein**

5

6 Matteo Calcagnile¹, Patricia Forgez², Antonio Iannelli^{3,4}, Cecilia Bucci^a, Marco Alifano^{5,6,*,#}, and
7 Pietro Alifano^{1,*,#}

8

9 ¹Department of Biological and Environmental Sciences and Technologies, University of Salento,
10 Lecce, Italy

11 ²INSERM UMR-S 1124 T3S, Eq 5 CELLULAR HOMEOSTASIS, CANCER and THERAPY,
12 University of Paris, Campus Saint Germain, Paris, France

13 ³Digestive Disease Department, Archet 2 Hospital, Nice University Hospital, University of Nice
14 Côte d'Azur, Nice, France

15 ⁴INSERM, U1065, Team 8 "Hepatic complications of obesity", University Nice Côte d'Azur,
16 France

17 ⁵Thoracic Surgery Department, Cochin Hospital, APHP Centre, University of Paris, France

18 ⁶INSERM U1138 Team «Cancer, Immune Control, and Escape», Cordeliers Research Center,
19 University of Paris, France

20

21 #These authors contributed equally to the work

22 *To whom correspondence should be addressed:

23 Pietro Alifano: Department of Biological and Environmental Sciences and Technologies, University
24 of Salento, Lecce, Italy; e-mail: pietro.alifano@unisalento.it

25 Marco Alifano: Thoracic Surgery Department, AP-HP, University of Paris, France; e-mail:
26 marco.alifano@aphp.fr

27

28 **Running title:** ACE2 polymorphism and interaction with SARS-CoV-2 Spike protein

29

30

31

32 **ABSTRACT**

33 There is increasing evidence that ACE2 gene polymorphism can modulate the interaction between
34 ACE2 and the SARS-CoV-2 spike protein affecting the viral entry into the host cell, and / or
35 contribute to lung and systemic damage in COVID-19. Here we used *in silico* molecular docking to
36 predict the effects of ACE2 missense variants on the interaction with the spike protein of SARS-
37 CoV-2. HDOCK and FireDock simulations identified 6 ACE2 missense variants (I21T, A25T,
38 K26R, E37K, T55A, E75G) with higher affinity for SARS-CoV-2 Spike protein receptor binding
39 domain (RBD) with respect to wild type ACE2, and 11 variants (I21V, E23K, K26E, T27A, E35K,
40 S43R, Y50F, N51D, N58H, K68E, M82I) with lower affinity. This result supports the hypothesis
41 that ACE2 genetic background may represent the first “genetic gateway” during the disease
42 progression.

43

44

45 **Keywords:** COVID-19; SARS-CoV-2 Spike protein; ACE2 polymorphism; *In silico* modeling;
46 SARS-CoV-2 Infectiousness; SARS-CoV-2 severity of infection

47

48

49 **1. Introduction**

50 Different phases can be distinguished during the progression of COVID-19 [1, 2]. During the first
51 phase, after the incubation period lasting 6 days in average, the onset of disease may be
52 characterized by influenza-like symptoms, from mild to moderate. The second phase, which is
53 known as the pulmonary phase and involves ~30% of all SARS-CoV-2 infected subjects, is
54 characterized by progressive respiratory involvement with onset of pneumonia-like symptoms. The
55 third phase, which develops in ~15% of all patients, is known as the pro-inflammatory phase, and is
56 characterized by severe interstitial pneumonia with focal and systemic iper-inflammation, which
57 may lead to acute respiratory distress syndrome, and systemic inflammatory response syndrome.
58 The fourth phase of COVID-19, which is known as the pro-thrombotic phase, develops in ~5% of
59 patients, and is characterized by the onset of microvascular and macrovascular thrombosis possibly
60 promoted by strong focal and/or systemic inflammation. During this phase patients require medical
61 treatment in intensive care units, and most of them do not survive.

62
63 SARS-CoV-2 infection susceptibility and severity seem to be influenced by environmental factors
64 (climate, pollution, cultural, social and economic inequalities, climate, health care system
65 organizations), co-morbidities (high blood pressure, cardiovascular disease, other heart and lung
66 conditions, diabetes, cancer, or compromised immune system), and inter-individual genetic
67 differences [3-6]. Inter-individual genetic differences may affect the **spatial transmission dynamics**
68 **of COVID-19, the susceptibility and severity of disease, and the inflammatory and immune**
69 **response**, and three “genetic gateways” have been proposed accounting for disease progression [7].

70
71 Specifically, **there is evidence that angiotensin-converting enzyme 2 (ACE2) is the** human cell
72 **receptor of SARS-CoV-2 [8-10], and it was speculated [5,7,11-17] that ACE2 gene polymorphism**
73 **may modulate the interaction between ACE2 and the Spike protein of SARS-CoV-2 during the**
74 **virus entry into the host cell. In particular, differential affinity of a number of ACE2 missense**
75 **variants for Spike protein was predicted using different computational approaches [12,18-21].**
76 **Moreover, since ACE2 regulates the renin-angiotensin-aldosterone system [22], ACE2 missense**
77 **variants or expression quantitative trait loci (eQTL) variants may contribute to pulmonary and**
78 **systemic injury by fostering vasoconstriction, inflammation, oxidation and fibrosis, thereby**
79 **affecting the clinical outcome [4,11,15, 23-25]. The possible association between specific ACE2**
80 **gene variants and COVID-19 susceptibility, severity, and clinical outcomes is supported by massive**
81 **genomic data from general population [26], while large-scale genome-wide association studies are**
82 **urgently needed to firmly establish the causal link [27].**

83

84 In this study we have used *in silico* molecular docking to analyze the possible effects of ACE2
85 single nucleotide polymorphisms (SNPs) leading to missense variants on the interaction between
86 ACE2 and SARS-CoV-2 Spike protein. Molecular docking was performed with HDOCK, a
87 powerful pipeline for integrated protein-protein docking, which is based on hybrid docking
88 algorithm of template-based modeling and *ab initio* free docking to optimize the adjustment of
89 ligand [28-31]. The HDOCK pipeline differs from other molecular docking platforms in its ability
90 to support amino acid sequences as inputs, and in its hybrid docking strategy in which experimental
91 information on the protein-protein binding site and small-angle X-ray scattering are incorporated
92 during the docking and post-docking processes [28].

93

94 With respect to the other pipelines that were previously used to model the interaction between
95 SARS-CoV-2 Spike protein and ACE2 missense variants [5,7,11,13-17], HDOCK has the
96 advantage of integrating two approaches with the same software, together with a remarkable
97 simplicity of use, and it is completely automated with consequent high reproducibility.

98

99

100 **2. Materials and methods**

101 *2.1. Databases*

102 3D structures of proteins were downloaded from Research Collaboratory for Structural
103 Bioinformatics Protein Data Bank (RCSB PDB) (<http://www.rcsb.org/>) [324]. We focused our
104 analysis on structures of SARS-COV-2 Spike Receptor Binding Domain (RBD)/ACE2 complexes
105 6M17 (10.2210/pdb6M17/pdb) [2233] 6LZG (10.2210/pdb6LZG/pdb) [2334], and 6M0J
106 (10.2210/pdb6M0J/pdb) [2435] models. The Single Nucleotide Polymorphism Database (dbSNP)
107 [25, 26 36,37] was used to identify the ACE2 SNPs leading to missense variants. Functional
108 information of ACE2 was acquired by UniProt database (Q9BYF1, ACE2_HUMAN) [2738].
109 ACE2 SNP frequencies were obtained by the database GnomAD-Exomes
110 (<https://gnomad.broadinstitute.org/>). ACE2 bat sequences were downloaded from NCBI database.
111 Multiple alignments of human and bat sequences were carried out by Clustal Omega [39].

112

113 *2.2. Molecular docking and statistical analysis*

114 HDOCK server (<http://hdock.phys.hust.edu.cn/>) was used to carry out molecular docking between
115 receptor binding domain (RBD) of SARS-CoV-2 Spike protein and ACE2 wild type or missense
116 variants from the dbSNP. We focused our analysis on two ACE2 N-terminal alpha helices that form

117 the major binding interface with SARS-CoV-2 Spike protein RBD based on X-ray crystallography
118 [33-35]. In our analysis we used, as a receptor, the amino acid sequence of ACE2 wild type or
119 missense variants, and, as a ligand, the SARS-CoV-2 Spike protein RBD models (6LZG chain B,
120 6M0J chain E or 6M17 chain E) downloaded from RCSB-PDB database. 6M17 was the most
121 complete structure because it contains the ACE2 collectrin-like domain [33]. Although this domain
122 is far from the binding interface, it could still affect the geometry of the protein. Since HDOCK
123 only provides score for *ab initio* free docking, to compare the complex scores obtained by *ab initio*
124 free docking and template-based modeling we used FireDock
125 (<http://bioinfo3d.cs.tau.ac.il/FireDock/>) [~~32,33~~ 40,41].

126
127 Results of HDOCK/FireDock simulations were confirmed by submitting HDOCK-generated ACE2/
128 Spike protein RBD complexes to PRODIGY [42]. Furthermore, wild type ACE2 and K26R ACE2
129 models were also built by using MODELLER 9.25 [43] via Chimera [44], and these models were
130 used as receptors in SwarmDock simulations [45]. HDOCK/FireDock pipeline was also used to
131 evaluate the impact of SARS-CoV-2 Spike protein RBD variants on binding to wild type or K26R
132 ACE2. QMEANDisCo (SwissDock) [46] and MolProbity [47] were used for bad bonds and angles
133 metrics. A detailed protocol of the computational workflow is depicted in Supplementary Figure S1.

134
135

136 3. Results

137 3.1. Molecular docking of SARS-CoV-2 Spike protein RBD and ACE2 wild type or missense
138 variants.

139 Two ACE2 N-terminal alpha helices form the major binding interface with SARS-CoV-2 Spike
140 protein RBD based on X-ray crystallography [~~22-24~~33-35]. In this region 25 SNPs causing leading
141 to ACE2 missense variants are listed in the dbSNP.

142

143 HDOCK and FireDock pipelines were used for molecular docking. For each ACE2 missense
144 variant, three docking simulations were carried, each with a different PDB model (6M17, 6LZG,
145 6M0J), and the results obtained with the two methods (template-based modeling and *ab initio* free
146 docking) were analyzed separately.

147

148 Before proceeding with the simulations, the quality of the models generated by HDOCK was
149 analyzed and compared with the quality of the corresponding models generated by MODELLER.
150 The analysis with QMEANDisCo demonstrated good quality of all models with global scores

151 similar to those of the control PDB models (6LZG, 6M0J and 6M17). Moreover, percentage of bad
152 angles according to MolProbity was slightly lower with HDOCK (wild type ACE2 = 0.8%; K26R
153 ACE2 = 0.77) compared to MODELLER (1.24% for both wild type and K26R ACE2)
154 (Supplementary Figure S2)

155

156 Overall, HDOCK/FireDock results with the different PDB models and methods were concordant in
157 92% of cases (Table 1; Supplementary Table S1; Fig. S34). For only two polymorphisms (N58K
158 and M62V), the template-based method produced results that differed from those produced by *ab*
159 *initio* docking.

160

161 We performed 156 docking simulations (i.e., 26 ACE2 SNPs, S19AQP ACE2, wild type ACE2, all
162 multiplied by three ligands and two methods). Global energy score (GES) average with all
163 simulations was -47.20 Kcal/mol (Fig. 1A, gray line), total standard deviation was 6.39 Kcal/mol
164 and confidence interval was ± 1.0035 Kcal/mol (Fig. 1A, dotted line). The highest GES was -37.94
165 Kcal/mol (M82I), while the lowest one was -56.24 (T55A). We used total GES average, GES value
166 with wild type ACE2 (Fig. 1A, red line), and confidence interval as a threshold to screen the SNPs,
167 considering as relevant only the SNPs that affected significantly the binding with SARS-CoV-2
168 Spike protein RBD. By using this approach, we found 6 out of the 25 ACE2 missense variants
169 (24%) (I21T, A25T, K26R, E37K, T55A, E75G) that showed higher affinity for SARS-CoV-2
170 Spike protein RBD with respect to wild type ACE2, and 11 variants (44%) (I21V, E23K, K26E,
171 T27A, E35K, S43R, Y50F, N51D, N58H, K68E, M82I) that exhibited lower affinity *in silico* (Fig.
172 1A)

173

174 3.2. Geographical distribution of ACE2 SNPs affecting binding to SARS-CoV-2 spike protein RBD

175 GnomAD-Exomes database was used to gain information about frequencies of the examined ACE2
176 SNPs worldwide (Table 1; Supplementary Table S2). K26R is the most diffused one with a global
177 frequency of 0.3971%. The large diffusion of this SNP is also confirmed by others database:
178 0.4579% in TOPMED; 0.595% in 4ALFA Project; 0.368% in ExAC; 0.315% in GnomAD; 0.511%
179 in GO-ESP; 0.21% in 1000G; 0.62% in TWINSUK; 0.93% in ALSPAC. In particular, the K26R
180 occurs with highest frequency in European (0.503%) and American (0.329%) populations with
181 maximum value in Ashkenazi Jewish (1.2%), while it is less common in both African (0.099%) and
182 Asian (0.079%) populations.

183

184 Cumulative frequency analysis of ACE SNPs demonstrated that ACE2 missense variants exhibiting
185 increased affinity for SARS-CoV-2 Spike protein were more common in European and American
186 populations (Fig. 1B), while those exhibiting reduced affinity were more common in African and
187 Asian populations (Fig. 1C). The frequencies of each ACE2 missense variants were plotted
188 individually in Supplementary Figure S4.

189

190 *3.3 Impact of SARS-CoV-2 Spike protein RBD variants on binding to ACE2*

191 A number of missense variants affecting the SARS-CoV-2 Spike protein have recently been
192 identified worldwide, and listed in a comprehensive database [48]. In particular, some of these
193 variants including N439K, L455F, F456L, A475V, Q493R, Q493L and N501Y, fall into the
194 interfaces of binding of RBD. The impact of these RBD variants on binding to wild type ACE2 or
195 K26R ACE2 was then evaluated. We focused on K26R missense variant because of its high
196 frequency in the general population.

197

198 Preliminary, PRODIGY was used to confirm the effect of ACE2 K26R missense on wild type Spike
199 protein RBD binding as predicted by FireDock with HDock complexes. Moreover, PRODIGY
200 calculated dissociation constants (K_d) that were $8.8E^{-10}$ for the ACE2 K26R and $4.610E^{-9}$ for wild
201 type ACE2. The effect of K26R missense was further confirmed using models that were generated
202 by MODELLER, and then submitted to SwarmDock obtaining an energy of -39.88 Kcal/mol for
203 wild type ACE2 and of -46.13 Kcal/mol for K26R ACE2.

204

205 HDock/FireDock analysis was performed using the wild type or K26R ACE2 receptor and
206 missense variants of the Spike protein RBD listed above. Results demonstrated that 5 of the 7 RBD
207 mutations increased binding affinity for wild type ACE2 (Fig1D), while 5 of the 7 RBD mutations
208 decreased the binding affinity for K26R ACE2 compared to wild type RBD (Fig1E).

209

210

211 **4. Discussion and conclusions**

212 The present study supports the hypothesis that ACE2 gene polymorphism may contribute to the
213 genetic susceptibility to COVID-19 affecting the SARS-CoV-2 entry into the host cells, thus
214 representing the first “genetic gateways” during disease progression [7]. Our results broaden the list
215 of ACE2 missense variants that can affect the interaction with the SARS-CoV-2 spike protein [7, ~~9-~~
216 ~~12~~, 168, 189, 1921]. Specifically, we focused our attention on ACE2 SNPs affecting two N-terminal
217 alpha helices that form the major binding interface with SARS-CoV-2 Spike protein [22-24 33-35].

218

219 We did not include in our analysis the S19P variant because it falls into the cleavage site of ACE2
220 precursor, and it may affect the N-terminal sequence of the mature protein. Besides, there is
221 evidence that the S19P may reduce the affinity for SARS-CoV-2 Spike protein [189]. Results about
222 K26R that is expected to increase the affinity for Spike protein is noteworthy because this variant is
223 relatively frequent in European people with a frequency about 0.5%, which would correspond to a
224 potential target population of 2,230,000 people at the European Union level [912]. In this study we
225 confirmed the results of K26R by using HDOCK that also allowed us to identify additional
226 missense variants (I21T, A25T, E37K, T55A, E75G) with higher affinity for SARS-CoV-2 Spike
227 protein, and 11 variants (I21V, E23K, K26E, T27A, E35K, S43R, Y50F, N51D, N58H, K68E,
228 M82I) with lower affinity.

229

230 It is worth of noticing the K26R variant of ACE2 was identified in a COVID-19 patient but not in
231 control subjects in Italy, in a recent genome-wide association study enrolling a cohort of 131
232 patients and 258 controls [27], further reinforcing the hypothesis that this missense variant may be
233 associated with clinical susceptibility to disease.

234

235 Beside, it may be also relevant to note is rather common in different families of bats including
236 Vespertilionidae and Phyllostomidae (Supplementary Fig. S5). Specifically, Phyllostomidae are
237 diffused in South America (*Desmodus rotundus* XP_024425698.1, *Phyllostomus discolor*
238 XP_028378317.1), while Phyllostomidae are very common in China (*Pipistrellus abramus*
239 ACT66266.1) and Indochina (*Kerivoula pellucida* QJF77795.1), and the presence of *P. abramus*
240 was confirmed in the Wuhan area [49].

241

242 It is conceivable that the polymorphisms responsible for a higher affinity may be responsible for a
243 greater severity of the disease in humans, especially when very high affinity receptors are
244 overexpressed due to the environmental and pharmacological factors. Of course, underlying
245 diseases would contribute to an even more severe course of the disease, with an intense viral
246 replication capable of infecting in turn a large number of persons, including some individuals with
247 similar ACE2 polymorphisms, and so on. Another aspect to consider is the co-evolution of Spike
248 protein. Indeed, missense mutations in the Spike RBD may have conflicting effects on binding
249 affinity for wild type and K26R ACE2 (Fig.1D and Fig.1E).

250

251 Polymorphisms in genes coding for proteases from the respiratory tract belonging to the
252 transmembrane protease/serine subfamily (TMPRSS) may also contribute to inter-individual
253 differences in susceptibility and severity of disease [26, 50]. Indeed, there is evidence that TMPRSS
254 proteolytic activity induces SARS-CoV Spike protein fusogenic activity, and, notably, SARS-CoV-
255 2 cell entry is dependent on TMPRSS2, and blocked by protease inhibitors [51].

256

257 Obviously, the impact of these polymorphisms on severity of outcome should be weighted by
258 appropriate demographic and clinical factors. If this difference were confirmed, this would pave the
259 way for the identification, on a population scale, of healthy individuals whose molecular
260 phenotypes would be responsible for disease that is more serious. Apart from the usual social
261 distancing measures, targeted drug prevention strategies could be evaluated. It could be logical to
262 assess pharmacological prophylactic interventions, as proposed in categories of healthy people at
263 particular risk of exposure such as care-givers. The serine protease inhibitor camostat mesylate,
264 approved in Japan to treat unrelated diseases, has been shown to block TMPRSS2 activity [52,53],
265 and is thus an interesting candidate. Conversely, the identification of broader categories of people
266 with lower risk of developing severe disease, could allow a safer exit from the lock-down phases,
267 while facilitating the establishment of a faster herd immunity, and waiting reliable serological tests
268 and effective vaccine.

269

270

271 **Conflict of interest**

272 The authors declare no conflict of interest.

273

274 **Authors' contribution**

275 M.C. contributed to experimental set-up, pipeline development, *in silico* analysis; P.F., A.I., and
276 C.B. contributed to study designing and data providing; M.A. and P.A. contributed to coordination,
277 conception, designing and writing. M.A. and P.A. contributed equally to the work. All authors
278 critically revised draft versions of the manuscript and approved the final version.

279

280 **Acknowledgements**

281 We wish to thank prof. Diane Damotte (University of Paris) for advice and critical reading of the
282 manuscript.

283

284

285

286 **References**

287

288 [1] W.G. Dos Santos, Natural history of COVID-19 and current knowledge on treatment therapeutic
289 options, *Biomed. Pharmacother.* 129 (2020) 110493. <https://doi.org/10.1016/j.biopha.2020.110493>.

290

291 [2] G. Lippi, F. Sanchis-Gomar, B.M. Henry, COVID-19: unravelling the clinical progression of
292 nature's virtually perfect biological weapon, *Ann. Transl. Med.* 8 (2020) 693.
293 <https://dx.doi.org/10.21037/atm-20-3989>.

294

295 [3] Centers for Disease control and Prevention (CDC), Assessing risk factors for severe COVID-19
296 illness. [https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/assessing-](https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/assessing-risk-factors.html)
297 [risk-factors.html](https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/assessing-risk-factors.html) , 2020 (accessed 27 August 2020)

298

299 [4] M. Alifano, P. Alifano, P. Forgez, A. Iannelli, Renin-angiotensin system at the heart of COVID-
300 19 pandemic, *Biochimie* 174 (2020) 30-33. <https://doi.org/10.1016/j.biochi.2020.04.008>

301

302 [5] M. Alifano, G. Attanasi, F. Iannelli, F. Cherikh, A. Iannelli, COVID-19 pandemic: a European
303 perspective on health economic policies, *JBEP COVID-19 Special Issue 4* (2020) 35-43.

304

305 [6] A. Iannelli, G. Favre, S. Frey, V. Esnault, J. Gugenheim, S. Bouam, L. Schiavo, A. Tran, M.
306 Alifano M, Obesity and COVID-19: ACE 2, the Missing Tile, *Obes. Surg.* 25 (2020) 1–3.
307 <https://dx.doi.org/10.1007/s11695-020-04734-7>

308

309 [7] M. Debnath, M. Banerjee, M. Berk, Genetic gateways to COVID-19 infection: Implications for
310 risk, severity, and outcomes, *FASEB J.* Epub ahead of print. <https://doi.org/10.1096/fj.202001115R>

311

312 [8] M. Letko, A. Marzi, V. Munster. Functional assessment of cell entry and receptor usage for
313 SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol.* 5(4) (2020) 562-569.
314 [doi:10.1038/s41564-020-0688-y](https://doi.org/10.1038/s41564-020-0688-y)

315

316 [9] A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Velesler. Structure, Function,
317 and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell.* 181(2) (2020) 281-292.e6. [doi:](https://doi.org/10.1016/j.cell.2020.02.058)
318 [10.1016/j.cell.2020.02.058](https://doi.org/10.1016/j.cell.2020.02.058)

319

320 [10] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.L. Hsieh, O. Abiona, B.S. Graham, J.S.
321 McLellan. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science.*
322 [367\(6483\) \(2020\)1260-1263. doi: 10.1126/science.abb2507](https://doi.org/10.1126/science.abb2507)

323

324 [811] M. Bosso, T.A. Thanaraj, M. Abu-Farha, M. Alanbaei, J. Abubaker, F. Al-Mulla, The Two
325 Faces of ACE2: The Role of ACE2 Receptor and Its Polymorphisms in Hypertension and COVID-
326 19, *Mol. Ther. Methods Clin. Dev.* 18 (2020) 321-327.
327 <https://dx.doi.org/10.1016/j.omtm.2020.06.017>

328

329 [912] M. Calcagnile, P. Forgez, A. Iannelli, C. Bucci, M. Alifano, P. Alifano. ACE2
330 polymorphisms and individual susceptibility to SARS-CoV-2 infection: insights from an *in silico*
331 study. *bioRxiv* 2020.04.23.057042. <https://doi.org/10.1101/2020.04.23.057042>.

332

333 [103] Y. Cao, L. Li, Z. Feng, S. Wan, P. Huang, X. Sun, F. Wen, X. Huang, G. Ning, W. Wang,
334 Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2
335 in different populations, *Cell Discov.* 6 (2020) 11. <https://doi.org/10.1038/s41421-020-0147-1>

336

337 [114] B. Darbani, The Expression and Polymorphism of Entry Machinery for COVID-19 in
338 Human: Juxtaposing Population Groups, Gender, and Different Tissues, *Int. J. Environ. Res. Public*
339 *Health* 17 (2020) 3433. <https://doi.org/10.3390/ijerph17103433>
340

341 [125] C.A. Devaux, J.M. Rolain, D. Raoult, ACE2 receptor polymorphism: Susceptibility to SARS-
342 CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome, *J. Microbiol. Immunol.*
343 *Infect.* 53 (2020) 425-435. <https://dx.doi.org/10.1016%2Fj.jmii.2020.04.015>
344

345 [146] K.J. Godri Pollitt, J. Peccia, A.I. Ko, N. Kaminski, C.S. Dela Cruz, D.W. Nebert,
346 J.K.V.Reichardt, D.C. Thompson, V. Vasiliou, COVID-19 vulnerability: the potential impact of
347 genetic susceptibility and airborne transmission, *Hum. Genomics* 4 (2020) 17.
348 <https://doi.org/10.1186/s40246-020-00267-3>
349

350 [157] G. Lippi, C.J. Lavie, B.M. Henry, F. Sanchis-Gomar, Do genetic polymorphisms in
351 angiotensin converting enzyme 2 (ACE2) gene play a role in coronavirus disease 2019 (COVID-
352 19)? *Clin. Chem. Lab. Med.* Epub ahead of print. <https://doi.org/10.1515/cclm-2020-0727>
353

354 [168] H. Hadi-Alijanvand, M. Rouhani, Studying the Effects of ACE2 Mutations on the Stability,
355 Dynamics, and Dissociation Process of SARS-CoV-2 S1/hACE2 Complexes, *J. Proteome Res.*
356 Epub ahead of print. <https://doi.org/10.1021/acs.jproteome.0c00348>
357

358 [189] M. Hussain, N. Jabeen, F. Raza, S. Shabbir, A.A. Baig, A. Amanullah, B. Aziz, Structural
359 variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein, *J. Med.*
360 *Virol.* Epub ahead of print. <https://doi.org/10.1002/jmv.25832>
361

362 [20] E.W. Stawiski, D. Diwanji, K. Suryamohan, R. Gupta, F.A. Fellouse, J.F. Sathirapongsasuti, J.
363 Liu, Y.P. Jiang, A. Ratan, Mis M. Human ACE2 receptor polymorphisms predict SARS-CoV-2
364 susceptibility. *bioRxiv.* 2020 2020.04.07.024752.
365

366 [4921] H. Othman, Z. Bouslama, J.T. Brandenburg, J. da Rocha, Y. Hamdi, K. Ghedira, N. Srairi-
367 Abid, S. Hazelhurst, Interaction of the spike protein RBD from SARS-CoV-2 with ACE2:
368 Similarity with SARS-CoV, hot-spot analysis and effect of the receptor polymorphism, *Biochem.*
369 *Biophys. Res. Commun.* 527 (2020) 702-708. <https://doi.org/10.1016/j.bbrc.2020.05.028>
370

371 [22] Y.Y. Zheng, Y.T. Ma, J.Y. Zhang, X. Xie. COVID-19 and the cardiovascular system. *Nat Rev*
372 *Cardiol.* 17(5) (2020)259-260. doi: 10.1038/s41569-020-0360-5
373

374 [423] D. Gemmati, B. Bramanti, M.L. Serino, P. Secchiero, G. Zauli, V. Tisato, COVID-19 and
375 Individual Genetic Susceptibility/Receptivity: Role of ACE1/ACE2 Genes, Immunity,
376 Inflammation and Coagulation. Might the Double X-chromosome in Females Be Protective against
377 SARS-CoV-2 Compared to the Single X-Chromosome in Males? *Int. J. Mol. Sci.* 21 (2020) 3474.
378 <https://doi.org/10.3390/ijms21103474>
379

380 [24] C.P. Nelson, I.E. Sama, V. Codd, T.R. Webb, S. Ye, C.C. Lang, A.A. Voors, L.L. Ng, N.J.
381 Samani. Genetic Associations With Plasma Angiotensin Converting Enzyme 2 Concentration:
382 Potential Relevance to COVID-19 Risk. *Circulation.* 15142(11) (2020) 1117-1119. doi:
383

384 [205] C. Strafella, V. Caputo, A. Termine, S. Barati, S. Gambardella, P. Borgiani, C. Caltagirone,
385 G. Novelli, E. Giardina, R. Cascella, Analysis of ACE2 Genetic Variability among Populations
386 Highlights a Possible Link with COVID-19-Related Neurological Complications, *Genes (Basel)* 11
387 (2020) 741. <https://doi.org/10.3390/genes11070741>

388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438

[4726] Y. Hou, J. Zhao, W. Martin, A. Kallianpur, M.K. Chung, L. Jehi, N. Sharifi, S. Erzurum, C. Eng, F. Cheng, New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis, *BMC Med.* 18 (2020) 216. <https://doi.org/10.1186/s12916-020-01673-z>

[27] E. Benetti, R. Tita, O. Spiga, A. Ciolfi, G. Birolo, A. Bruselles, G. Doddato, A. Giliberti, C. Marconi, F. Musacchia, T. Pippucci, A. Torella, A. Trezza, F. Valentino, M. Baldassarri, A. Brusco, R. Asselta, M. Bruttini, S. Furini, M. Seri, V. Nigro, G. Matullo, M. Tartaglia, F. Mari, GEN-COVID Multicenter Study, A. Renieri, A.M. Pinto. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *Eur J Hum Genet.* 17 (2020) 1–13. doi: 10.1038/s41431-020-0691-z

[28] Y. Yan, H. Tao, J. He, S.Y. Huang, The HDock server for integrated protein-protein docking, *Nat. Protoc.* 15 (2020) 1829-1852. <https://doi.org/10.1038/s41596-020-0312-x>

[29] Y. Yan, Z. Wen, X. Wang, S.Y. Huang, Addressing recent docking challenges: A hybrid strategy to integrate template-based and free protein-protein docking, *Proteins* 85 (2017) 497-512. <https://doi.org/10.1002/prot.25234>

[30] Y. Yan, D. Zhang, P. Zhou, B. Li, S.Y. Huang, HDock: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy, *Nucleic Acids Res.* 45 (2017) W365-W373. <https://doi.org/10.1093/nar/gkx407>

[31] S.Y. Huang, X. Zou X, An iterative knowledge-based scoring function for protein-protein recognition, *Proteins* 72 (2008) 557-579. <https://doi.org/10.1002/prot.21949>

[324] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, Shindyalov I.N., P.E. Bourne, The Protein Data Bank, *Nucleic Acids Res.* 28 (2000) 235-242. <https://doi.org/10.1107/s0907444902003451>

[2233] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2, *Science* 367 (2020) 1444-1448. <https://doi.org/10.1126/science.abb2762>

[2334] Q. Wang, Y. Zhang, L. Wu, S. Niu, C. Song, Z. Zhang, G. Lu, C. Qiao, Y. Hu, K.Y. Yuen, Q. Wang, H. Zhou, J. Yan, J. Qi, Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2, *Cell* 181 (2020) 894-904. <https://doi.org/10.1016/j.cell.2020.03.045>

[2435] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, *Nature* 581 (2020) 215-220. <https://doi.org/10.1038/s41586-020-2180-5>

[2536] S.T. Sherry, M.H. Ward, M. Kholodov, J. Baker, L. Phan, E.M. Smigielski, K. Sirotkin, dbSNP: the NCBI database of genetic variation, *Nucleic Acids Res.* 29 (2001) 308-311. <https://doi.org/10.1093/nar/29.1.308>

[2637] M. Bhagwat, Searching NCBI's dbSNP database, *Curr. Protoc. Bioinformatics* 2010 Dec; Chapter 1:Unit 1.19. <https://dx.doi.org/10.1002/2F0471250953.bi0119s32>

439 [27 38] UniProt Consortium T, UniProt: the universal protein knowledgebase, *Nucleic Acids Res.*
440 46 (2018) 2699. <https://doi.org/10.1093/nar/gkw1099>
441
442 [39] F. Sievers, D.G. Higgins (2014). Clustal Omega, accurate alignment of very large numbers of
443 sequences. In *Multiple sequence alignment methods* (pp. 105-116). Humana Press, Totowa, NJ.
444
445 [3240] N. Andrusier, R. Nussinov, H.J. Wolfson, FireDock: fast interaction refinement in molecular
446 docking, *Proteins* 69 (2007) 139-159. <https://doi.org/10.1002/prot.21495>
447
448 [3341] E. Mashiach, D. Schneidman-Duhovny, N. Andrusier, R. Nussinov, H.J. Wolfson,
449 FireDock: a web server for fast interaction refinement in molecular docking, *Nucleic Acids Res* 36
450 (2008) W229-W232. <https://doi.org/10.1093/nar/gkn186>
451
452 [42] L.C. Xue, J.P. Rodrigues, P.L. Kastritis, A.M. Bonvin, A. Vangone, A. PRODIGY: a web
453 server for predicting the binding affinity of protein–protein complexes. *Bioinformatics*, 32(23)
454 (2016) 3676-3678. doi: 10.1093/bioinformatics/btw514
455
456 [43] B. Webb, A. Sali. Comparative protein structure modeling using MODELLER. *Current*
457 *protocols in bioinformatics*. 54.1 (2016) 5-6. doi: 10.1002/cpps.20
458
459 [44] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E.
460 Ferrin. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of*
461 *computational chemistry*. 25(13) (2004) 1605-1612. doi: 10.1002/jcc.20084.
462
463 [45] M. Torchala, I.H. Moal, R.A. Chaleil, J. Fernandez-Recio, P.A. Bates. SwarmDock: a server
464 for flexible protein–protein docking. *Bioinformatics*, 29(6) (2013) 807-809.
465 doi:10.1093/bioinformatics/btt038
466
467 [46] G. Studer, C. Rempfer, A. M. Waterhouse, R. Gumienny, J. Haas, T. Schwede, T.
468 QMEANDisCo—distance constraints applied on model quality estimation. *Bioinformatics*, 36(6),
469 (2020) 1765-1771. doi: 10.1093/bioinformatics/btz828
470
471 [47] I. W. Davis, A. Leaver-Fay, V.B. Chen, J.N. Block, G.J. Kapral, X. Wang, L.W. Murray
472 W.B. Arendall, J. Snoeyink, J.S. Richardson, D.C. Richardson. MolProbity: all-atom contacts and
473 structure validation for proteins and nucleic acids. *Nucleic acids research*. (2007) 35(suppl_2),
474 W375-W383. doi:10.1093/nar/gkm216
475
476 [48] M. Shaminur Rahman, M. Rafiul Islam, M. Nazmul Hoque, A.S.M. Rubayet UI Alam, m.
477 Akther, J. Akter Puspo, S. Akter, A. Anwar, M. Sultana, M. Anwar Hossain. Comprehensive
478 annotations of the mutational spectra of SARS-CoV-2 spike protein: a fast and accurate pipeline.
479 *Transbound Emerg.* (2020) . doi: 10.1111/tbed.13834
480
481 [49] C.P. Tong. Distribution and preference of landscape features and foraging sites of
482 insectivorous bats in Hong Kong urban parks. *HKU Theses Online (HKUTO)* (2016).
483
484 [50] G. Vargas-Alarcón, R. Posadas-Sánchez, J. Ramírez-Bello. Variability in genes related to
485 SARS-CoV-2 entry into host cells (ACE2, TMPRSS2, TMPRSS11A, ELANE, and CTSL) and its
486 potential use in association studies. *Life Sci.* 260 (2020)118313. doi: 10.1016/j.lfs.2020.118313
487
488

489 [51] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T.S.
490 Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Müller, C. Drosten, S. Pöhlmann. SARS-CoV-2
491 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease
492 Inhibitor. *Cell* 181 (2020) 271-280.e8. doi: 10.1016/j.cell.2020.02.052.
493
494 [52] M. Kawase, K. Shirato, L. van der Hoek, F. Taguchi, S. Matsuyama. Simultaneous treatment
495 of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe
496 acute respiratory syndrome coronavirus entry. *J. Virol.* 86 (2012) 6537-6545.
497 doi:10.1128/JVI.00094-12
498
499 [53] Y. Zhou, P. Vedantham, K. Lu, J. Agudelo, R. Jr Carrion, J.W. Nunneley, D. Barnard, S.
500 Pöhlmann, J.H. McKerrow, A.R. Renslo, G. Simmons. Protease inhibitors targeting coronavirus
501 and filovirus entry. *Antiviral Res.* 116 (2015) 76-84. doi: 10.1016/j.antiviral.2015.01.011
502
503
504
505
506
507
508

509 **Legends to Figure**

510

511 **Figure 1. Molecular docking simulations and frequencies of ACE2 SNPs that either enhance**
512 **or the affinity with SARS-Cov-2 Spike protein. A)** Global Energy Score (GES) (Kcal/mol) of the
513 interaction between wild type ACE2 or ACE2 missense variants and SARS-Cov-2 Spike protein.
514 Molecular docking simulation were carried out by using HDOCK, which is based on two methods:
515 template-based modeling and *ab initio* free docking. GES here shown are an average of the GES
516 with the two methods. As reference values are shown: the average GES obtained with all
517 simulations (solid line), the confidence interval (dotted line), and the GES obtained with wild type
518 ACE (red line). **B-C)** Frequencies of ACE2 SNPs that either enhance (I21T, K26R, E37K, T55A)
519 (**B**), or reduce (I21V, E23K, K26E, T27A, E35K, S43R, Y50F, N51D, N58H, K68E, M82I) (**C**) the
520 affinity with SARS-Cov-2 Spike protein, based on GES.

521

522

523

524 **Table**

525

526 **Table 1.** ACE2 SNPs analyzed in this study, their frequencies, and Global Energy Score (GES,
 527 Kcal/mol) of the interaction between wild type ACE2 or ACE2 missense variants and SARS-Cov-2
 528 Spike protein.

ACE2 wild type or missense variant	dbSNP ID	Frequency	Template-based modeling GES (Kcal/mol) ^a		Ab initio docking GES (Kcal/mol) ^a		Total GES (Kcal/mol) ^b	
		GnomAD Exomes	Average	SD	Average	SD	Average	SD
Wild type	-	-	-48.15	1.83	-50.45	0	-49.3	1.71
I21T	rs1244687367	0.000005	-53.83	4.35	-53.96	3.93	-53.9	3.71
I21V	rs778030746	0.000011	-42.76	0.43	-41.17	0.65	-41.97	1
E23K	rs756231991	0.000005	-37.66	5.38	-45.68	0.43	-41.67	5.56
A25T	rs1434130600	-	-55.89	0.54	-54.43	0.62	-55.16	0.95
K26R	rs4646116	0.003971	-55.2	0.36	-54.51	1.43	-54.86	1.01
K26E	rs1299103394	0.000005	-40.94	2.48	-40.24	0.69	-40.59	1.67
T27A	rs781255386	0.000011	-47.51	1.01	-40.99	2.07	-44.25	3.86
E35D	rs778500138	-	-47.87	1.24	-44.42	3.63	-46.14	3.08
E35K	rs1348114695	0.000016	-36.68	4.85	-42.19	1.03	-39.43	4.35
E37K	rs146676783	0.000033	-54.65	2.38	-53.92	1.33	-54.29	1.77
F40L	rs924799658	0.000016	-52.76	0.44	-50.75	3.34	-51.76	2.4
S43R	rs1447927937	0.000005	-42.82	0.99	-45.28	0.39	-44.05	1.5
Y50F	rs1192192618	0.000005	-45.33	0.32	-44.12	0.61	-44.72	0.79
N51D	rs760159085	0.000005	-39.55	0.72	-38.45	2.67	-39	1.85
T55A	rs775273812	0.000006	-57.8	4.06	-54.67	2.7	-56.24	3.53
N58K	rs771621249	0.000011	-45.68	3.24	-54	0.84	-49.84	5.03
N58H	rs1222417695	0.000011	-46.26	0.06	-40.53	2.16	-43.39	3.42
Q60R	rs759162332	0.000011	-51.12	4.87	-54.81	0.12	-52.96	3.68
M62V	rs1325542104	0.000006	-53.77	0.03	-47.37	3.07	-50.57	4.01
N64K	rs1199100713	0.000005	-45.51	3.81	-48.97	1.46	-47.24	3.2
K68E	rs755691167	0.000011	-39.34	0.63	-37.75	0.45	-38.54	1
F72V	rs1256007252	0.000005	-46.91	1.51	-50.35	0.88	-48.63	2.18
E75G	rs867318181	-	-52.47	1.29	-54.27	0.44	-53.37	1.31
S77F	rs1234981462	N.D.	-44.85	4.78	-49.97	0.24	-47.41	4.13
M82I	rs766996587	0.000011	-33.49	2.7	-42.39	1.05	-37.94	5.21

529 ^aValues are averages of GES obtained with the three different 3D PDB SRARS-CoV-2 Spike
 530 protein RDB structures (6LZG chain B, 6M0J chain E or 6M17 chain E), which were used as
 531 ligands. SD, standard deviation. ^bValues are averages of GES obtained with the two methods
 532 (template-based modeling or *ab initio* docking) and the three different 3D PDB structures.

533

534

