



Mushtaq Hussain ORCID iD: 0000-0001-8862-179X

## **Structural Variations in Human ACE2 may Influence its Binding with SARS-CoV-2 Spike Protein**

Dr. Mushtaq Hussain<sup>1\*</sup>, Dr. Nusrat Jabeen<sup>2</sup>, Fozia Raza<sup>1</sup>, Sanya Shabbir<sup>1,2</sup>, Ayesha Ashraf Baig<sup>1</sup>, Anusha Amanullah<sup>1</sup> and Basma Aziz<sup>1</sup>

<sup>1</sup>Bioinformatics and Molecular Medicine Research Group, Dow College of Biotechnology, Dow Research Institute of Biotechnology and Biomedical Sciences, Dow University of Health Sciences, <sup>2</sup>Department of Microbiology, University of Karachi, Karachi-75280, Pakistan

### **\*Corresponding Author**

Dr. Mushtaq Hussain,  
Assistant Professor and Vice Principal,  
Dow College of Biotechnology  
Dow University of Health Sciences  
Karachi-Pakistan

Email: mushtaq.hussain@duhs.edu.pk

Cell Number: 0092-333-3468293

**Shortened Title:** ACE2 Variants and SARS-CoV-2 Interaction

### **Abstract**

The recent pandemic of COVID-19, caused by SARS-CoV-2, is unarguably the most fearsome compared to the earlier outbreaks caused by other coronaviruses, SARS-CoV and MERS-CoV. Human ACE2 is now established as a receptor for the SARS-CoV-2

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jmv.25832.

This article is protected by copyright. All rights reserved.

Accepted Article

spike protein. Where variations in the viral spike protein in turn lead to the cross species transmission of the virus, genetic variations in the host receptor ACE2, may also contribute to the susceptibility and/or resistance against the viral infection. This study aims to explore the binding of the proteins encoded by different human ACE2 allelic variants with SARS-CoV-2 spike protein. Briefly, coding variants of ACE2 corresponding to the reported binding sites for its attachment with coronavirus spike protein were selected and molecular models of these variants were constructed by homology modelling. The models were then superimposed over the native ACE2 and ACE2-spike protein complex, to observe structural changes in the ACE2 variants and their intermolecular interactions with SARS-CoV-2 spike protein respectively. Despite strong overall structural similarities, spatial orientation of the key interacting residues varies in the ACE2 variants compared to the wild type molecule. Most ACE2 variants showed similar binding affinity for SARS-CoV-2 spike protein as observed in the complex structure of wild type ACE2 and SARS-CoV-2 spike protein. However, ACE2 alleles, rs73635825 (S19P) and rs143936283 (E329G) showed noticeable variations in their intermolecular interactions with the viral spike protein. In summary, our data provide structural basis of potential resistance against SARS-CoV-2 infection driven by ACE2 allelic variants.

**Keywords:** SARS-CoV-2, ACE2, Spike protein, COVID-19

## **Introduction**

Evolution is a dynamic and continuous interplay, where pathogens and hosts are in continuous battle to overpower each other. Random mutations in pathogens then sculpted by natural selection empower them to attain increased virulence and mechanisms to evade host immune response. The recent pandemic of COVID-19 is yet another manifestation of the might of natural selection<sup>1</sup>. SARS-CoV-2 is the seventh strain of coronavirus and fourth of beta coronavirus, so far known to infect humans. Prior to SARS-CoV-2, two strains of coronavirus, namely SARS-CoV and MERS-CoV have caused outbreaks of relatively limited scale<sup>2</sup>. SARS infection epidemic, dubbed as Severe Acute Respiratory Syndrome (SARS), during 2002-2003, spread across 37 countries affecting nearly 8,500 individuals with 916 deaths<sup>3</sup>. In comparison, by January 2020,

MERS-CoV infected 2,519 individuals of 27 countries with 866 deaths<sup>4</sup>. The recent outbreak of COVID-19, now categorized as pandemic, is by all means considerably wide spread compared to earlier coronavirus outbreaks. By 29<sup>th</sup> March 2020, the infection has its presence around the globe with over half a million of world population being infected and has claimed over 30,000 lives and counting<sup>5</sup>.

SARS-CoV and SARS-CoV-2 spike proteins interact with human Angiotensin-Converting Enzyme 2 (ACE2) as their receptor, whereas MERS-CoV spike protein attaches with DiPeptidyl Peptidase 4 (DPP4)<sup>6,7</sup>. Recently, several structures of SARS-CoV-2 spike protein complexed with human ACE2 have been resolved (PDBids: 6LZG; 6VW1; 6M17), highlighting the critical residues involved in the intermolecular interactions between the viral spike protein and its receptor<sup>7</sup>. Since the virus has been transmitted in humans from bats and/or other intermediate hosts, considerable work has been conducted to explore the sequence and structural variations in the spike proteins and animal host ACE2 receptors<sup>8,9,10</sup>. It has been shown that variations in both viral spike protein and host ACE2 sequences may act as a barrier for viral infection across species<sup>2,8,9</sup>. This in turn raises two interesting possibilities: first, do the natural genetic polymorphism in human *ACE2* gene and/or protein influences their attachment with SARS-CoV-2 spike protein? Second, are these genetic variants of *ACE2* benign or could be considered as evolutionary trade-off? where a variation in a trait may bring both advantage and disadvantage to the fitness of an organism. This study aims to explore both these possibilities by a battery of sequence and structure analysis tools. The findings provide interesting insights to the potential relationship between natural genetic variants in human ACE2 and susceptibility and/or resistance against COVID-19 infection.

## Methodology

### Data mining for structure and genetic polymorphism

Protein sequences of SARS-CoV-2 spike protein (YP\_009724390.1) and human ACE2 (Q9BYF1) were retrieved from NCBI and UniProt, respectively. Structures of both the proteins were identified by PDB-BLAST and retrieved from RCSB protein data bank<sup>11</sup>. Data for genetic variants and allele frequencies were obtained from Ensembl Genome

Browser<sup>12</sup> and gnomAD<sup>13</sup>. Appropriate filters were employed to limit the data for only coding region variants of *ACE2* gene. Coding variants of *ACE2* corresponding to the critical binding sites between *ACE2*-SARS-CoV<sup>14</sup>, *ACE2*-SARS-CoV-2 (PDBid;6LZG), and those reported to show impaired binding with coronavirus spike protein in the *in vitro* mutation analysis<sup>15</sup> were selected for further investigations (Supplementary Table 1).

### **Impact analysis of allelic variants**

Effect of amino acid substitution on the protein (*ACE2*) stability was predicted at 25°C and 37°C using I-Mutant2.0<sup>16</sup>. The program predicts the change in the free energy ( $\Delta\Delta G$ ) in the sequence variants compared to the wildtype. Increase or decrease in the protein stability is indicated by negative or positive  $\Delta\Delta G$  values, respectively. Functional impact of all selected allelic variants of *ACE2* were predicted using Sorting Intolerant From Tolerant (SIFT)<sup>17</sup>, PolyPhen-2<sup>18</sup>, Combined Annotation-Dependent Depletion (CADD)<sup>19</sup> and Rare Exome Variant Ensemble Learner (REVEL)<sup>20</sup>.

### **Protein molecular modelling**

Total 17 *ACE2* coding region allelic variants were identified that correspond to the critical interaction points between *ACE2* and receptor binding domain of coronavirus spike protein. Structural models of these variants were generated using atomic coordinates of PDBid; 2AJFA (wild type *ACE2*) as template by Modeller9.16<sup>21</sup>. The models were optimized for Gibb's free energy and loop conformation. Each model was then individually superimposed over template and Root Mean Square Deviation (RMSD) in the C $\alpha$  back bone was estimated in Å by SWISS-PdBViewer 4.1.0<sup>22</sup>. In addition, variations in the spatial orientation of the key residues and corresponding change in the intramolecular hydrogen bonds were also noted by DS visualizer 2016.

### **Structure analysis of *ACE2* variants and SARS-CoV-2 spike protein complexes**

Very recently, several co-crystal structures of human *ACE2* complexed with SARS-CoV-2 spike protein, (PDBids: 6LZG; 6VW1; 6M17), have been resolved via X-ray diffraction and cryo-electron microscopy. In this study, 6LZG was chosen because of the non-chimeric nature of the spike protein and resolution at lower wave length (2.5Å) for

generation of docking poses by superimposition. Docking poses were developed by Carbon  $\alpha$  backbone superimposition of each ACE2 variant model over PDBid; 6LZG using SWISS-PdBViewer 4.1.0<sup>22</sup>. Intermolecular hydrogen bonds, electrostatic and hydrophobic interactions between SARS-CoV-2 spike protein and human ACE2 variants were monitored using DS visualizer 2016. PRODIGY web server was used to predict binding affinity and different types of intermolecular interactions: charged-charged, charged-polar, charged-apolar, polar-polar, polar-apolar and apolar-apolar of the complexes<sup>23</sup>.

## Results

SARS-CoV-2 spike glycoprotein is a 1273 amino acid long structural protein located on the outer envelope of the virus. The protein has two main functional subunits: a long N-terminal S1 subunit and relatively short C-terminal S2 subunit. A 200 amino acid long Receptor Binding Domain (RBD) is stationed within S1 subunit of the spike protein that is mainly involved in its interaction with ACE2 receptor<sup>6,7,10</sup>. Human ACE2 is an 805 amino acid long protein with two functional domains: N-terminal peptidase M2 domain and C-terminal collectrin domain. Partial structure of the ACE2 has been resolved (PDBid; 1R42) comprising peptidase M2 domain. Structurally, the peptidase domain has two catalytic sub domains with an active site sandwiched in between the two subdomains<sup>24</sup>. Co-crystal structures of SARS-CoV-2 spike protein complexed with human ACE2 (PDBid; 6LZG), SARS-CoV with human ACE2<sup>14</sup> and *in vitro* mutation analysis of human ACE2<sup>15</sup> have identified the critical residues that underpin the interaction between the RBD of viral spike protein and ACE2 peptidase domain. In this study, amongst 345 and 242 natural ACE2 coding variants identified from Ensembl Genome Browser and gnomAD, respectively, 17 were found at the positions that had shown to be important for the binding of ACE2 with the viral spike protein (Supplementary table 1). Allele frequencies of these variants range from 5.45E-6 (rs1299103394) to 3.88E-3 (rs4646116). To predict the potential pathological consequences rendered by ACE2 alleles, changes in the free energy of ACE2 variants were estimated. The results showed that amino acid substitutions in only three alleles, rs73635825 (S19P), rs1299103394 (K26E) and rs766996587 (M82I) may adversely

affect the stability of the encoded protein compared to the wild type (Table 1). However, to date no pathogenic consequences have been reported in relation to these variants in humans. Moreover, except for rs73635825 (S19P), which was predicted to be damaging by only PolyPhen-2, different prediction tools designed to assess the functional impact of the mutations and/or SNPs did not reveal any ensuing pathogenicity rendered by these variants (Table 1). Conversely, rs961360700 (D355N) and rs762890235 (P389H) were predicted to be detrimental by SIFT and Polyphen-2, whereas rs1396769231 (M383T) was predicted to be damaging by SIFT, Polyphen-2 and REVEL prediction tools. Nevertheless, no report of any functional impact was found in the literature for these allelic variants of ACE2 as well.

In order to explore the structural changes in the protein encoded by different alleles of ACE2, molecular models of all the selected protein variants were developed and superimposed over the structurally resolved template of wild type ACE2. Structurally, all ACE2 variants bear the characteristic two subdomains of peptidase M2 domain with active site cleft present in between. The overall protein architecture of ACE2 allelic variants is largely conserved with RMSD of C $\alpha$  back bone varies from 0.17Å to 0.58Å, compared to the wild type (Table 1; Figure 1A,B). Like SARS-CoV, RBD of SARS-CoV-2 spike protein comprises two subdomains: core and extended loop. Co-crystal structures of SARS-CoV and SARS-CoV-2 spike proteins complexed with ACE2 have demonstrated that the extended loop of RBD directly interacts with loops flanked by  $\alpha$ 2 and  $\alpha$ 3 helices and a  $\beta$  hairpin loop between  $\beta$ 3 and  $\beta$ 4 sheets of ACE2<sup>7,14</sup>. The additional residues of ACE2 reported to affect its binding with the viral spike protein by mutation analysis<sup>15</sup> lies at C-terminal to these regions. Though the C $\alpha$  back bone of ACE2 allelic variants was found highly conserved, the spatial orientation of substituting residues varied noticeably compared to the wild type (Figure 1C,D). Therefore it is possible that these variations in the orientation of amino acid side chain may in turn alter the intramolecular interactions in different allelic variants. This consequently may affect the orientation of other critical amino acids of ACE2, otherwise not substituted in the respective variants, but required for the interaction with SARS-CoV-2 spike protein.

To compare the binding of different ACE2 allelic variants with SARS-CoV-2 spike protein, docking poses were generated by superimposition of structures of ACE2 variants over ACE2-SARS-CoV-2 complex structure (PDBid; 6LZG). All ACE2 variants were found to bind with the viral spike protein with topology nearly identical to the resolved ACE2-SARS-CoV-2 complex structure (PDBid; 6LZG) (Figure 2A). This is not surprising, since no major structural change was observed in the protein backbone of different ACE2 variants (Figure 1A, B). By and large, nature and numbers of intermolecular contacts between SARS-CoV-2 spike protein and humans ACE2 variants were found comparable (Figure 2B). However, some important variations were observed, for example, compared to total of 70 interactions found in SARS-CoV-2 spike protein and wild type ACE2 complex, 60 and 81 intermolecular interactions were observed between SARS-CoV-2 spike protein and rs143936283 (E329G) and rs961360700 (D355N) variants of ACE2, respectively. Amongst the six compared types of interactions, three most strongest, charged-charged, charged-polar and charged-apolar were found nearly identical between rs961360700 (D355N) and SARS-CoV-2 spike protein compared to the wild type ACE2 and SARS-CoV-2 spike protein complex. In comparison, ACE2 variant, rs143936283 (E329G) showed less charged-charged, charged-polar, charged-apolar, polar-apolar and apolar-apolar interactions with SARS-CoV-2 spike protein. This implies that rs143936283 (E329G) in humans may confer some level of resistance against the attachment of SARS-CoV-2 to the receptor molecule. However, accounting the dynamic nature of protein-protein interaction, it is difficult to draw a fair conclusion from simply these observations. Therefore, inter residual interaction maps were developed for complexes between each ACE2 allelic variant and SARS-CoV-2 spike protein.

Recent co-crystal structure of SARS-CoV-2 spike protein with ACE2 (PDBid; 6LZG) highlighted the interaction points of both the proteins (Figure 2C). Many of these interactions were also reported in previously resolved structure of SARS-CoV spike protein-ACE2 complex (PDBid; 2AJF)<sup>14</sup>. A representation for the comparison of these interactions for different ACE2 alleles is shown in Figure 2D. Briefly, the most conserved intermolecular interactions (hydrogen bonds) between different ACE2 variants and SARS-CoV-2 spike protein were found between Y41, H34, Y83 and K353 of ACE2

and T500, Y453 (except M82I), N487 (except T27A and D427Y) and G502 (except S19P and E329G) of SARS-CoV-2 spike protein, respectively. Similarly, hydrophobic interactions between K31 and K353 of ACE2 and Y489 and Y505 of SARS-CoV-2 spike protein, respectively, were also found conserved in complexes of all ACE2 variants with the viral spike protein. This implies that none of these ACE2 allelic variants offer complete resistance against the attachment of SARS-CoV-2. However, some important variations are noteworthy to mention here. For example, Q42 of wild type ACE2 interacts with G446 and Q449 of the SARS-CoV-2 spike protein, but all ACE2 alleles-spike protein complexes did not show any such interaction. The Q42 of ACE2 also interacts with Y449 of the spike protein and this interaction is also absent in ACE2 alleles corresponding to T27A, E35D, E37K, M82I, E329G and M383T amino acid substitutions. Similar observation was made for the interaction between E35 of ACE2 and Q493 of the SARS-CoV-2 spike protein (Figure 2D; Supplementary Figure). This may suggest that ability of E35 and Q42 to interact with SARS-CoV-2 spike protein could be affected by its spatial positioning in ACE2 protein which may get affected by the change in the intramolecular interaction and/or electrostatic potential due to the substitution in the flanking residues with respect to three dimensional conformation of ACE2. Conversely, K353 of ACE2 establishes hydrogen bond interactions with G496 and G502 of SARS-CoV-2 spike protein. Both these interaction were found absent in only S19P (rs73635825) variant of ACE2. Moreover, the binding affinity of the S19P ACE2 variant for SARS-CoV-2 spike protein was also predicted to be lowest (-10.3) amongst all the alleles including the wild type ACE2. Similarly, E329G ACE2 variant which was predicted to have second lowest binding affinity (-10.8) for SARS-CoV-2 spike protein also lacked K353-G502 interaction in the respective complex. It is important to note here that K353 is one of the key residues of ACE2 that not only showed interaction with the spike proteins of both SARS-CoV<sup>14</sup> and SARS-CoV-2, but its substitution in *in vitro* mutation analysis has shown to completely abolish the interaction between ACE2 and coronavirus spike protein<sup>15</sup>. Therefore it is likely that rs73635825 (S19P) and rs143936283 (E329G) alleles of ACE2 may offer some level of resistance against the SARS-CoV-2 attachment with the human ACE2 receptor.

## Discussion

*ACE2* gene encodes a zinc metalloprotease which acts as a receptor molecule for three strains of coronavirus, SARS-CoV, NL63 and recently SARS-CoV-2. Several studies have explained the variations in the coronavirus spike protein and host ACE2 receptor in order to predict and/or prove the origin and potential of cross transmission of the virus<sup>1,7,8,9</sup>. However, less attention has been given to explore the effect of natural genetic and expressional variants of human ACE2 for the potential susceptibility and/or resistance against the coronavirus infection. *In vitro* assays have shown that expression of ACE2 is positively correlated with the SARS-CoV infection<sup>25</sup>. Similarly, genetic variants in ACE2 have been proposed to affect interaction of the receptor with the viral spike protein<sup>15</sup>. Conversely, a recent preliminary study predicts no such link between the human ACE2 variants and coronavirus infection<sup>26</sup>. The present study focuses to explore the variation in the binding of protein encoded by ACE2 coding variants with SARS-CoV-2 spike protein.

Many coding variants of ACE2 in humans have been associated with cardiovascular disorders, hypertension and diabetes<sup>27,28</sup>. Although some variants of ACE2, assessed in the present investigation, were predicted to have some functional impact but to date no empirical data showing their link with any disease or genetic disorder in humans. Moreover, compared to the wild type, no significant change has been observed in the overall structural conformation of the ACE2 protein variants and the substituted amino acids are located on one of the subdomain of peptidase M2, distantly to the catalytic sites of ACE2<sup>24</sup>. Thereby it is plausible to suggest that these variants of ACE2 may be under neutral selection. Protein molecular modeling and related bioinformatics tool could provide significant insights to predict potential variations in the protein structures and complexes<sup>29</sup>. In the present study a comparative modeling and molecular superimposition revealed important variations in the intermolecular interaction between ACE2 alleles and SARS-CoV-2 spike protein. It is interesting to note that for two ACE2 alleles, rs73635825 (S19P) and rs143936283 (E329G), low binding affinity and lack of some of the key residues in the complex formation with SARS-CoV-2 spike protein could be suggestive of intrinsic resistance to some scale against the SARS-CoV-2 infection.

Therefore, it is conceivable that under new selection pressure as offered by SARS-CoV-2, these alleles may undergo positive selection. However, since protein-ligand and protein-protein interactions are dynamic<sup>30</sup>, therefore implementation of molecular dynamic simulation could further validate the absence of some of the critical interactions between these ACE2 alleles and SARS-CoV-2 spike protein. Consistently, empirical means of deducing binding affinity of intermolecular complexes could also provide important insights in this regard<sup>31</sup>.

Where mutations in viruses enable them to cross species barrier and/or become more virulent, natural genetic variants in host receptors may concomitantly offer susceptibility/resistance against evolving pathogenic viruses. For example, genetic variations in CD4 receptor (C868T) and CCR5-Δ32 in humans confer susceptibility and resistance, respectively, against certain HIV strains<sup>32,33</sup>. Present investigation highlights candidate alleles of ACE2 that may lead to the variation in the susceptibility and/or resistance against COVID-19. Moreover, it provides a template for similar investigations in relation to the recently reported molecule, TMPRSS2, required for priming the spike protein for cellular entry of the virus<sup>34</sup>. Finally, clinical manifestations and recovery rate of COVID-19 varies significantly between different age groups, nationalities and race<sup>4,5,35</sup>. It is possible that if not all, in some individuals positive prognosis of the COVID-19 may be due to the existence of ACE2 variants like rs73635825 (S19P) and rs143936283 (E329G). Therefore, the findings of this investigation provide clues to screen frequencies of the candidate alleles in different populations to predict the prognosis of COVID-19.

#### References:

1. Chan HL, Tsui SK, Sung JJ. Coronavirus in severe acute respiratory syndrome (SARS). *Trends Mol Med*. 2003;9(8):323-325.
2. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol*. 2019;17(3):181-192.
3. Chan-Yeung M, Xu RH. SARS: epidemiology. *Respirology*. 2003;8(Suppl):S9-14.

4. <http://www.emro.who.int/health-topics/mers-cov/mers-outbreaks.html>
5. <https://www.worldometers.info/coronavirus/>
6. Li F. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. *J Virol.* 2015;89(4):1954-1964.
7. Yan R, Zhang Y, Li Y, et al. Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science.* 2020;367(6485):1444-1448.
8. Liu Z, Xiao X, Wei X, et al. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. [published online ahead of print February 26, 2020]. *J Med Virol.* [https://doi: 10.1002/jmv.25726](https://doi.org/10.1002/jmv.25726)
9. Wan Y, Shang J, Graham R, et al. Receptor recognition by novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS. *J virol.* 2020;94(7).
10. Lu G, Wang Q, Gao GF. Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol.* 2015;23(8):468-478.
11. Goodsell DS, Zardecki C, Di Costanzo L, et al. RCSB Protein Data Bank: Enabling biomedical research and drug discovery. *Protein Sci.* 2020;29(1):52-65.
12. Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, et al. Ensembl variation resources. *Database.* 2018;2018:1-12.
13. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285-291.
14. Li F, Li W, Farzan M, et al. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science.* 2005;309(5742):1864-1868.
15. Li W, Zhang C, Sui J, et al. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J.* 2005;24(8):1634-1643.
16. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* 2005;33:W306-W310.
17. Sim NL, Kumar P, Hu J, et al. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* 2012;40(W1):W452-W457.

18. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.
19. Rentzsch P, Witten D, Cooper GM, et al. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47(D1):D886-D894.
20. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*. 2016;99(4):877-885.
21. Webb B, Sali A. Comparative protein structure modeling using MODELLER. *Curr Protoc Bioinformatics*. 2016;54(1):5-6.
22. Johansson MU, Zoete V, Michielin O, et al. Defining and searching for structural motifs using DeepView/Swiss-PdbViewer. *BMC Bioinformatics*. 2012;13(1):173.
23. Xue LC, Rodrigues JP, Kastritis PL, et al. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*. 2016;32(23):3676-3678.
24. Towler P, Staker B, Prasad SG, et al. ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J Biol Chem*. 2004;279(17):17996-18007.
25. Hofmann H, Geier M, Marzi A, et al. Susceptibility to SARS coronavirus S protein-driven infection correlates with expression of angiotensin converting enzyme 2 and infection can be blocked by soluble receptor. *Biochem Biophys Res Comm*. 2004;319(4):1216-1221.
26. Cao Y, Li L, Feng Z, et al. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discov*. 2020;6(1):1-4
27. Luo Y, Liu C, Guan T, et al. Association of ACE2 genetic polymorphisms with hypertension-related target organ damages in south Xinjiang. *Hypertens Res*. 2019 ;42(5):681-689.
28. Meng N, Zhang Y, Ma J, et al. Association of polymorphisms of angiotensin I converting enzyme 2 with retinopathy in type 2 diabetes mellitus among Chinese individuals. *Eye*. 2015;29(2):266-271.

29. Dong S, Sun J, Mao Z, et al. A guide line for homology modeling of the proteins from newly discovered betacoronavirus, 2019 novel coronavirus (2019-nCoV). [published online ahead of print 17 March 2020]. *J Med Virol*. <https://doi.org/10.1002/jmv.25768>
30. Salmaso V, Moro S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Front Pharmacol*. 2018;9:923.
31. Zeug A, Woehler A, Neher E, et al. Quantitative intensity-based FRET approaches—a comparative snapshot. *Biophys J*. 2012;103(9):1821-1827.
32. Marmor M, Hertzmark K, Thomas SM, et al. Resistance to HIV infection. *J Urban Health*. 2006;83(1):5-17.
33. Oyugi JO, Vouriot FC, Alimonti J, et al. A common CD4 gene variant is associated with an increased risk of HIV-1 infection in Kenyan female commercial sex workers. *J Infect Dis*. 2009;199(9):1327-1334.
34. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*. 2020;181:1-10.
35. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020; 395 (10229):1054-1062.

**Table 1. Comparison of structural and functional consequences of selected ACE2 allelic variants**

Variants	A.A	Allele Frequency	RMSD (Å)	DDG		SIFT	PolyPhen-2	CADD	REVEL
				25°C	37°C				
rs73635825	S19P	3.13E-4	0.24	0.39	0.41	T	PD	LB	LB
rs1299103394	K26E	5.45E-6	0.29	0.67	0.66	T	B	LB	LB

rs4646116	K26R	3.88E-3	0.24	-0.34	-0.30	T	B	LB	LB
rs781255386	T27A	1.09E-5	0.19	-1.86	-1.72	T	B	LB	LB
rs778500138	E35D	N/A	0.24	-0.51	-0.49	T	B	LB	LB
rs1348114695	E35K	1.64E-5	0.24	-1.46	-1.44	T	B	LB	LB
rs146676783	E37K	3.9E-5	0.27	-1.03	-1.02	D	PD	LB	LB
rs755691167	K68E	1.09E-5	0.20	-0.67	-0.65	T	B	LB	LB
rs766996587	M82I	2.44E-5	0.27	0.73	0.77	T	B	LB	LB
rs759134032	P84T	5.47E-6	0.27	-1.40	-1.32	T	B	LB	LB
rs143936283	E329G	3.44E-5	0.17	-0.46	-0.32	T	B	LB	LB
rs961360700	D355N	1.17E-5	0.58	-0.86	-0.64	D	PD	LB	LB
rs1396769231	M383T	N/A	0.21	-1.57	-1.49	D	PD	LB	LDC
rs762890235	P389H	3.83E-5	0.21	-1.27	-1.19	D	PD	LB	LB
rs1238146879	P426A	5.47E-6	0.25	-1.57	-1.42	T	B	LB	LB
rs1316056737	D427Y	1.09E-5	0.21	-0.18	-	D	PD	LB	LB

					0.10				
rs1016777825	R559S	N/A	0.21	-1.38	- 1.28	T	B	LB	LB

Abbreviations: B (Benign), D (Deleterious), LB (Likely Benign), PD (Probably Damaging) and T (Tolerated).

Figure 1. Structure of ACE2 allelic variants. Structural comparison of wild type (green) and 17 different allelic variants of ACE2 protein (differently coloured) in (A) ribbon conformation and (B) C $\alpha$  backbone. (C) Ball and stick representation of key amino acids critical for interaction with corona virus spike protein in wild type ACE2 (green) compared to the corresponding substituting amino acid (purple) in 17 different ACE2 allelic variants. (D) Enlarged view of the boxed region (C), amino acid changes with corresponding SNP id are labelled at respective positions.

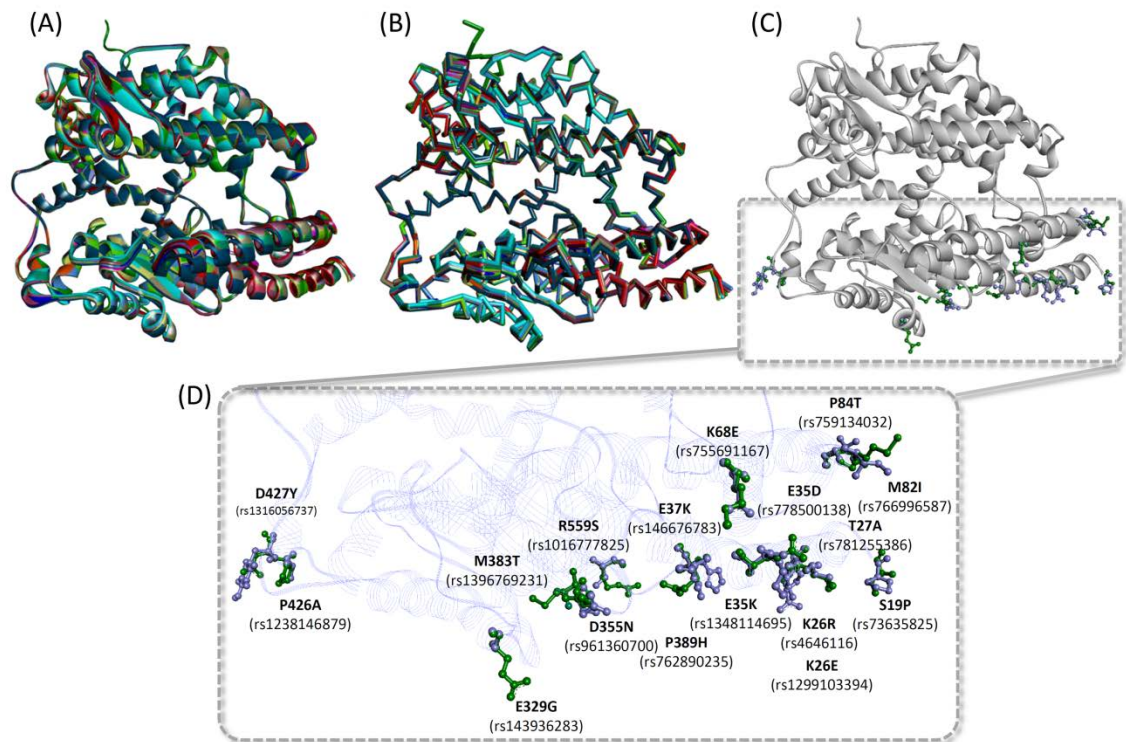


Figure 2. Interaction between ACE2 allelic variants and SARS-CoV-2 spike protein. (A) Ribbon representation of docking conformation of spike protein (gold) with wild type ACE2 (green) and its allelic variants (gray). (B) Bar graph showing number of different types of intermolecular contacts between ACE2 variants and SARS-CoV-2 spike protein. (C) Intermolecular interactions between wild type ACE2 (gray ribbon) and SARS-CoV-2 spike protein (gold ribbon). Amino acids of ACE2 and spike protein are represented by purple and orange sticks; red, blue and green dotted lines respectively correspond to hydrogen bonds, electrostatic interactions and hydrophobic interactions (see supplementary figure for all ACE2 variants). (D) Comparison of critical intermolecular interactions and binding affinity between different ACE2 variants complexed with SARS-CoV-2 spike protein, where green ticks and red crosses represent presence and absence of the interactions, respectively.

