



Pierce into Structural Changes of Interactions Between Mutated Spike Glycoproteins and ACE2 to Evaluate Its Potential Biological and Therapeutic Consequences

Zahra Sadat Hashemi¹ · Mahboubeh Zarei² · Shaden M. H. Mubarak³ · Anahita Hessami⁴ · Maysam Mard-Soltani⁵ · Bahman Khalesi⁶ · Alireza Zakeri⁷ · Mohammad Reza Rahbar² · Abolfazl Jahangiri⁸ · Navid Pourzardosht⁹ · Saeed Khalili⁷

Accepted: 8 December 2021

© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

The structural consequences of ongoing mutations on the SARS-CoV-2 spike-protein remains to be fully elucidated. These mutations could change the binding affinity between the virus and its target cell. Moreover, obtaining new mutations would also change the therapeutic efficacy of the designed drug candidates. To evaluate these consequences, 3D structure of a mutant spike protein was predicted and checked for stability, cavity sites, and residue depth. The docking analyses were performed between the 3D model of the mutated spike protein and the ACE2 protein and an engineered therapeutic ACE2 against COVID-19. The obtained results revealed that the N501Y substitution has altered the interaction orientation, augmented the number of interface bonds, and increased the affinity against the ACE2. On the other hand, the P681H mutation contributed to the increased cavity size and relatively higher residue depth. The binding affinity between the engineered therapeutic ACE2 and the mutant spike was significantly higher with a distinguished binding orientation. It could be concluded that the mutant spike protein increased the affinity, preserved the location, changed the orientation, and altered the interface amino acids of its interaction with both the ACE2 and its therapeutic engineered version. The obtained results corroborate the more aggressive nature of mutated SARS-CoV-2 due to their higher binding affinity. Moreover, designed ACE2-based therapeutics would be still highly effective against covid-19, which could be the result of conserved nature of cellular ACE2.

Keywords SARS-CoV-2 · Lineage B.1.1.7 · Mutation · Spike protein · Structural analyses

✉ Navid Pourzardosht
pourzardosht@gums.ac.ir

✉ Saeed Khalili
saeed.khalili@sru.ac.ir

¹ ATMP Department, Breast Cancer Research Center, Motamed Cancer Institute, ACECR, Tehran, Iran

² Pharmaceutical Sciences Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

³ Department of Clinical Laboratory Science, Faculty of Pharmacy, University of Kufa, Najaf, Iraq

⁴ School of Pharmacy, Shiraz University of medical sciences, Shiraz, Iran

⁵ Department of Clinical Biochemistry, Faculty of Medical Sciences, Dezful University of Medical Sciences, Dezful, Iran

⁶ Department of Research and Production of Poultry Viral Vaccine, Razi Vaccine and Serum Research Institute, Agricultural Research Education and Extension Organization, Karaj, Iran

⁷ Department of Biology Sciences, Shahid Rajaei Teacher Training University, Tehran, Iran

⁸ Applied Microbiology Research Center, Systems Biology and Poisonings Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran

⁹ Biochemistry Department, Guilan University of Medical Sciences, Rasht, Iran

Introduction

Since December 2019, the world had faced a pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a highly successful novel coronavirus with a mortality rate of about 2–5% (Zhang et al. 2021). Coronaviruses are positive-sense single-stranded RNA viruses that lead to enhanced flexibility in genetic modifications, differences in the severity of the disease, and some cases, even the change in host range (Toyoshima et al. 2020; Wu et al. 2020). On December 14, 2020, the WHO reported a new strain of SARS-CoV-2 identified in the United Kingdom through viral genomic sequencing (known as UK variant, British variant, or English variant) (Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations 2020). The report revealed a specific phylogenetic cluster of COVID-19 disease in this country. This cluster has been proliferating ever since, which indicates its further expansion. Various aspects of this cluster are significant for epidemiological and biological reasons. This newly identified species of coronavirus, called lineage B.1.1.7, has a large number of genetic changes, especially in the spike protein. Three of these mutations have potential biological effects that have already been described with varying amplitudes. The first mutation is the N501Y that is one of the six central contact residues in the receptor-binding domain (RBD) of the spike protein. This mutation can affect the binding affinity of the spike protein to the Angiotensin-converting enzyme II (ACE2 (Boehm and Nabel 2002; Donoghue et al. 2000)). The second mutation observed in the lineage B.1.1.7 of the UK variant is the deletion mutation within the spike protein (69-70 deletion). This condition has been described in the context of escaping from the human immune response and has also occurred several times in association with other RBD changes (Davies et al. 2021; Kemp et al. 2020). The third mutation, P681H, is located immediately near the furin cleavage site. These mutations are of biological importance and, as mentioned, are more common in the locations where the virus binds to the host cell. Moreover, treatment options and therapeutics could be affected by these mutations. The rapid pandemic of this new lineage of the virus indicates the need for advanced genomic and epidemiological surveillance worldwide and laboratory studies of infectious diseases (Investigation of novel SARS-CoV-2 variant: Variant of Concern 202012/01 2020; SARS-CoV-2 Variant – United Kingdom of Great Britain and Northern Ireland 2020).

The mutagenesis methods have significantly contributed to understanding the structural-functional relationship of proteins and modification of their physicochemical properties. These studies have shown that specificity,

binding affinity, and immunogenicity of proteins can be altered by replacing amino acids in the contact sites. Since several amino acids are responsible for the binding affinity of antigenic proteins to their receptors, substitution of such essential amino acids can significantly alter the binding affinity (Clark et al. 2006; Xu et al. 2009). These amino acids could be determined in vitro or silico (using bioinformatics tools). Bioinformatics is an expanding field applied in rapid and accurate diagnosis of diseases using protein biomarkers (Hashemi et al. 2019; Jamadi et al. 2020; Khalili et al. 2017; Khodashenas et al. 2019; Mard-Soltani et al. 2018; Rahbar et al. 2019; Ramezani et al. 2019), cost-effective production of vaccines (Asadi-Ghalehni et al. 2018; Asadi-Ghalehni et al. 2017), design of newer and more effective drugs (Hashemi et al. 2021; Pourzardosht et al. 2020), improving the quality and quantity of agricultural products, and a better understanding of the biological mechanisms (Noori et al. 2020). Recently, an engineered ACE2 was designed and evaluated as a novel therapeutics against COVID-19 (Payandeh et al. 2020). ACE2 receptor plays the most significant role in the virus fusion into the lung cells. Therefore, the current study was conducted to investigate the structural effects of new mutations in the spike protein of mutant spike protein on its interaction with the ACE2 receptor and its engineered therapeutic variant.

Methods

Sequence Analyses

The protein sequence of the mutant spike protein was obtained from the genomic sequence of the SARS-CoV-2 variant. The GISAID Initiative (Global Initiative on Sharing All Influenza Data) (<https://www.gisaid.org/>) is established to share data from all influenza viruses and the coronavirus causing COVID-19. This public domain was searched for the genomic sequence of the England mutated variant. The obtained variant was fed into the ORFfinder tool (<https://www.ncbi.nlm.nih.gov/orffinder/>) of the NCBI database. The ORF belonging to the spike protein was found using the Smart BLAST tool of the ORFfinder, and its reverse-transcribed protein sequence was selected for further analyses. The CLUSTALW (<https://www.genome.jp/tools-bin/clustalw>) multiple sequence alignment tool was used to compare the mutant and reference spike protein sequences. The reference spike sequence was obtained from the UniProt (<https://www.uniprot.org/>) database.

Protein Structure Prediction

Due to the high number of mutations, the 3D structure of the spike protein should be determined. The following structure prediction servers were employed to predict the 3D structure of the mutated spike protein: Robetta (<http://rosetta.bakerlab.org/>) server which uses a relatively fast and accurate deep learning-based method, ITASSER (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) server which uses a hierarchical approach to protein structure prediction and structure-based function annotation, LOMETS (<https://zhanglab.ccmb.med.umich.edu/LOMETS/>) server which is a new generation of meta-server approach to template-based protein structure prediction and structure-based function annotation, and SWISS-MODEL (<https://swissmodel.expasy.org/>) server which is a fully automated protein structure homology-modeling server. The quality of the obtained models was determined using the QMEAN (<https://swissmodel.expasy.org/qmean/>) and Prosa (<https://prosa.services.came.sbg.ac.at/prosa.php>) servers. The 3Drefine (<http://sysbio.rnet.missouri.edu/3Drefine/index.html>) protein structure refinement server was used to improve the quality of the obtained structural models.

Molecular Dynamics

MD simulation was performed on the monomeric form of the mutant spike protein with the GROMACS 5.0.1 simulation package (Abraham et al. 2015), using the GROMOS 96 53a6 force field (Oostenbrink et al. 2004). The protein was placed in a cube box large enough to contain the protein and 1 nm of solvent on all sides and solvated with the simple point charge (SPC) water molecules. The system was neutralized by adding counter-ions Na^+ and Cl^- (0.15 M), and then was energy minimized through the steepest descent integrator with a maximum of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ and or 50,000 steps. Subsequently, the system was equilibrated under NVT and NPT ensemble. The temperature and the pressure were maintained at 300 K and 1 atm using the v-rescale algorithm and isotropic Parrinello–Rahman algorithm, with coupling times of $\tau^T = 0.1 \text{ ps}$ and $\tau^P = 0.5 \text{ ps}$, respectively. After equilibration, the MD production run was performed for 50 ns under the periodic boundary conditions with the 2 fs time step. The particle mesh Ewald (PME) method was applied to compute the long-range electrostatic interactions (Darden et al. 1993). The distance cut-off for van der Waals and Coulombic interactions was 1.4 nm. The LINCS algorithm was used to constrain the bond lengths (Hess et al. 1997). The radius of gyration (Rg) analysis was performed to evaluate the compactness of the structure.

Protein Docking

The structure of mutant spike protein (after MD simulation) and the ACE2 receptor were used to analyze their molecular interactions via molecular docking methods. To investigate the changes, the interaction between the RBD region of the spike protein and the ACE2 receptor under the PDB ID of 6M17 was used as the control structure. The HADDOCK 2.5 server (<https://wenmr.science.uu.nl/haddock2.4/>), Clus-Pro server (<https://cluspro.bu.edu/login.php>), and ZDOCK server (<http://zdock.umassmed.edu/>) was used to perform the docking analyses. The 319–541 residues of the spike protein and the 30–41, 82–84, and 353–357 residues of the ACE2 protein were set to be the interacting residues in the HADDOCK 2.5 server.

Refinement for Docked Complexes

FireDock (<http://bioinfo3d.cs.tau.ac.il/FireDock/>) refinement server utilizes a coarse refinement method to optimize the interaction in molecular docking studies. The software was used to refine the docked complexes. The results of the FireDock software were fed to rigid-body orientation and side-chain conformations optimization by the RosettaDock server (<http://rosettadock.graylab.jhu.edu/>). The amino acids in the interface of the docked protein structures were extracted using UCSF ChimeraX version 1.1. The docked systems of the wild-type and mutant spike proteins and the ACE2 protein were checked for their binding affinity using the PRODIGY server (<https://wenmr.science.uu.nl/prodigy/>). This server is widely used to predict the binding affinity in PROTEIN-PROTEIN complexes.

Generation of 2D Interaction Plots

LigPlus software (v.2.1) was used to generate the 2D plot of the interactions between the ACE2 and spike protein (both wild type and mutated). The molecular docking result between the ACE2 and mutant spike protein, and the 6M17 complex was used to draw the 2D interaction plots. This analysis would determine the amino acids involved in the interaction between the ACE2 and spike protein.

Residue Centrality Analyses

Residue Interaction Network Generator (<http://protein.bio.unipd.it/>) was used to extract the residue interaction network of ACE2 and spike protein (both wild type and mutated) using their docked tertiary structures. A strict distance threshold was selected to generate strong and reliable interactions. Calculation of the distance between mass centers of two interacting residues was set to be the network policy to identify all types of interactions between the interacting

pairs. Cytoscape 3.7.2 was exploited to analyze further and visualize the generated network. RINspecter application of Cytoscape was used to carry out a centrality analysis and investigate the critical residues in the correct folding and function of the protein. Change of average shortest path length following the removal of each residue (node in the network) was the base of the performed centrality analysis. The minimum number of edges required to connect the first node to the second one in the network is defined as the shortest path length. The mean of these numbers is defined as the average shortest path. The significance of the average shortest path length is determined by the calculation of Zscore. The Z-scores higher than two is considered relevant.

Cavity Search and Depth Analyses

The Molegro Virtual Docker V.6 was used to detect the cavities within the structures of the wild-type and mutant spike proteins. The Detect Cavity tool of the software was used to detect the top 10 cavities within the structures. The DEPTH server at <http://cospi.iiserpune.ac.in/depth/> was used to compute/predict the DEPTH, Cavity Sizes, and Ligand Binding Sites of the wild type and mutant spike proteins. Depth is defined as the distance of an atom/residue to its closest molecule of bulk solvent. Residue depth correlates significantly better than accessibility with effects of mutations on protein stability and protein–protein interactions.

The Interactions Between Mutated Spike and the Engineered ACE2 Proteins

The ZDOCK, ClusPro, and HADDOCK servers were used to perform a docking analysis between the mutated spike protein and the engineered ACE2 protein to be compared

with the docking results between the mutated spike protein and the wild type ACE2 protein. Both mutated spike and engineered ACE2 structures had undergone the MD simulations to get closer to the native like structures. The engineered ACE2 structure was adapted from a previous study conducted by Payandeh et al. (Payandeh et al. 2020). Following the docking analyses, the FireDock refinement server was exploited to optimize the interaction within the docked complexes. Rigid-body orientation and side-chain conformations optimization of the docking results was done by the RosettaDock server. The PRODEGY server was used to calculate the binding affinity between the docked complexes of the engineered ACE2 and both mutant and wild type spike proteins. The amino acids involved in the interface of the engineered ACE2 and both mutant and wild type spike proteins were extracted using UCSF ChimeraX version 1.1.

Results

Mutated Spike Sequence

The genomic sequence of the hCoV-19/England/MILK-9E05B3/2020 variant was stored under the GISAID accession number EPI_ISL_601443. The variant information is listed in Table 1. The ORF sequence of the spike protein was found spanning the 21500 to 25312 nucleotides of the genomic sequence. Comparing the sequences of the spike protein from the EPI_ISL_601443 variant and the reference sequence from UniProt (ID: PODTC2) has confirmed the reported mutations. According to the results of the multiple sequence alignment, the mutations of the EPI_ISL_601443 variant were as follows: H69 deletion, V70 deletion, Y144 deletion, N501Y substitution, A570D substitution, D614G

Table 1 Genomic information for EPI_ISL_601443

Virus detail		Sample information	
Virus name	hCoV-19/England/MILK-9E05B3/2020	Collection date	2020-09-20
Accession ID	EPI_ISL_601443	Location	Europe/United Kingdom/England
Variant	VUI202012/01 GR/501Y.V1 (B.1.1.7)	Host	Human
Passage details/history	Original	Sequencing technology	Illumina NovaSeq
Institute information		Submitter information	
Originating lab	Lighthouse Lab in Milton Keynes	Submitter	Davies, Robert
Sample ID given by the originating laboratory	MILK-9E05B3	Submission date	2020-10-27
Submitting lab	Wellcome Sanger Institute for the COVID-19 Genomics UK (COG-UK) consortium	–	–
Sample ID given by the submitting laboratory	MILK-9E05B3	–	–

substitution, P681H substitution, T716I substitution, S982A substitution, and D1118H substitution.

Protein Structure Modeling

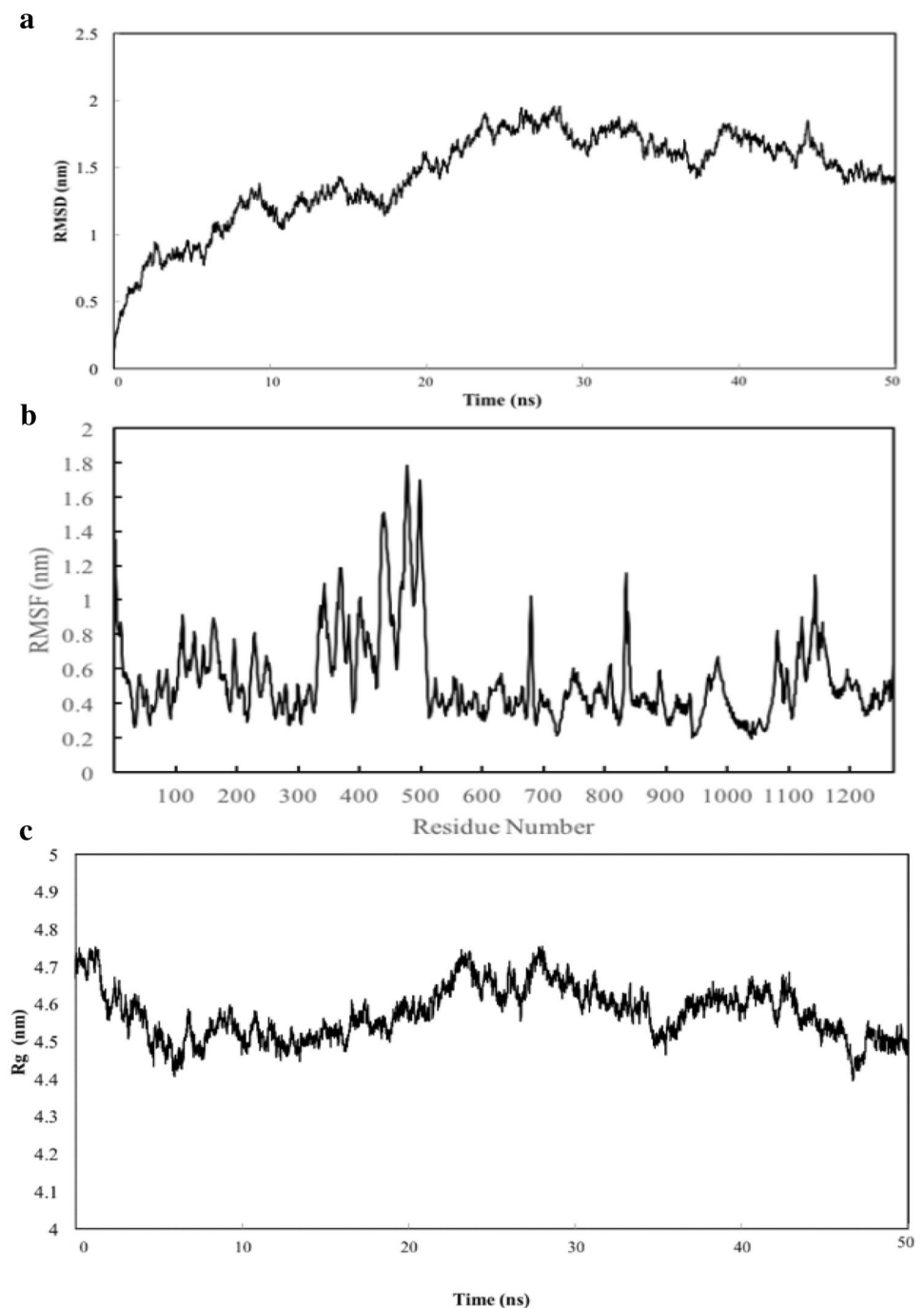
All of the protein structure prediction servers managed to build 3D models of the mutated spike protein. The model quality assessment results indicated that the predicted models are of acceptable structural quality. Given these results, the model built by the SWISS-MODEL server has

the highest quality scores (Supplementary Table 1). The N-terminus and the C-terminus regions of the model were associated with the lowest quality scores. The protein structure refinement server had improved the quality of the model predicted by the SWISS-MODEL server.

MD Simulation

The stability and dynamics of the simulated protein in the solvent environment were evaluated through

Fig. 1 The RMSD and RMSF plots of the MD simulation. **a** Backbone RMSD concerning the initial structure of the monomeric form of mutant spike protein during the simulation time. **b** $C\alpha$ RMSF of the monomeric form of mutant spike protein during the simulation time. **c** R_g versus time plot of the mutant spike protein during 50 ns MD simulation



root-mean-square deviation (RMSD) (fitted to the initial structure), and the root-mean-square fluctuation (RMSF) as

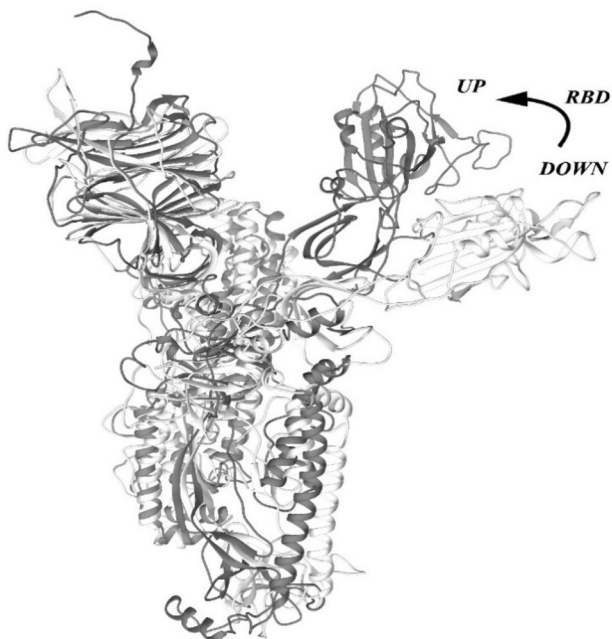


Fig. 2 Superimposition of spike glycoprotein before and after MD simulation. Protein structures are represented as cartoons. The gray ribbon is the mutant spike protein before MD simulation and white ribbons are the molecule after 50 ns MD simulation

a function of simulation time. As shown in Fig. 1a, the backbone RMSD of the mutant spike protein gradually increased to around 1.8 nm during the 20 ns, and remained relatively constant until the 47 ns and then decreased around 1.5 nm in the last three ns. These changes indicate that the stability was attained within the first 20 ns of the simulation time and the systems are well equilibrated. The protein flexibility was assessed using the RMSF of C-alpha ($C\alpha$) atoms per residue (Fig. 1b). The highest RMSF values belonged to the residues 329–510, which corresponded to the RBD domain of the protein. These results correlate with previous studies (He et al. 2020; Kalathiya et al. 2020). The structural dynamic for spike protein is illustrated in Fig. 2 before and after MD simulation. The monomeric spike protein in the solvent environment represented a movement from the “up” towards the “down” for the RBD domain. A relatively steady value of R_g over the simulation time shows that a protein is stably folded. The R_g of the mutant spike protein was relatively steady and averaged around 4.7 nm (Fig. 1c).

Protein–Protein Docking

The results of protein–protein docking analyses showed that the interaction between the ACE2 receptor and the spike protein occurs at the RBD site of the mutant spike protein. Despite the structural changes of the mutant spike protein, the orientation of interaction with the ACE2 receptor is

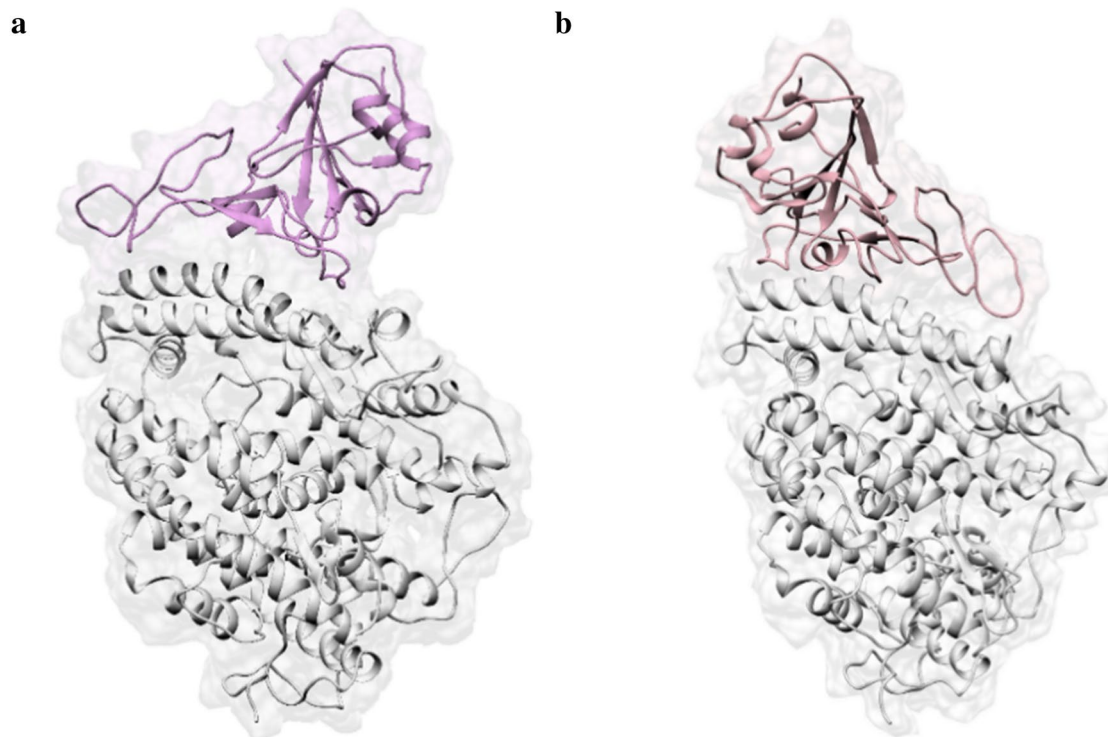


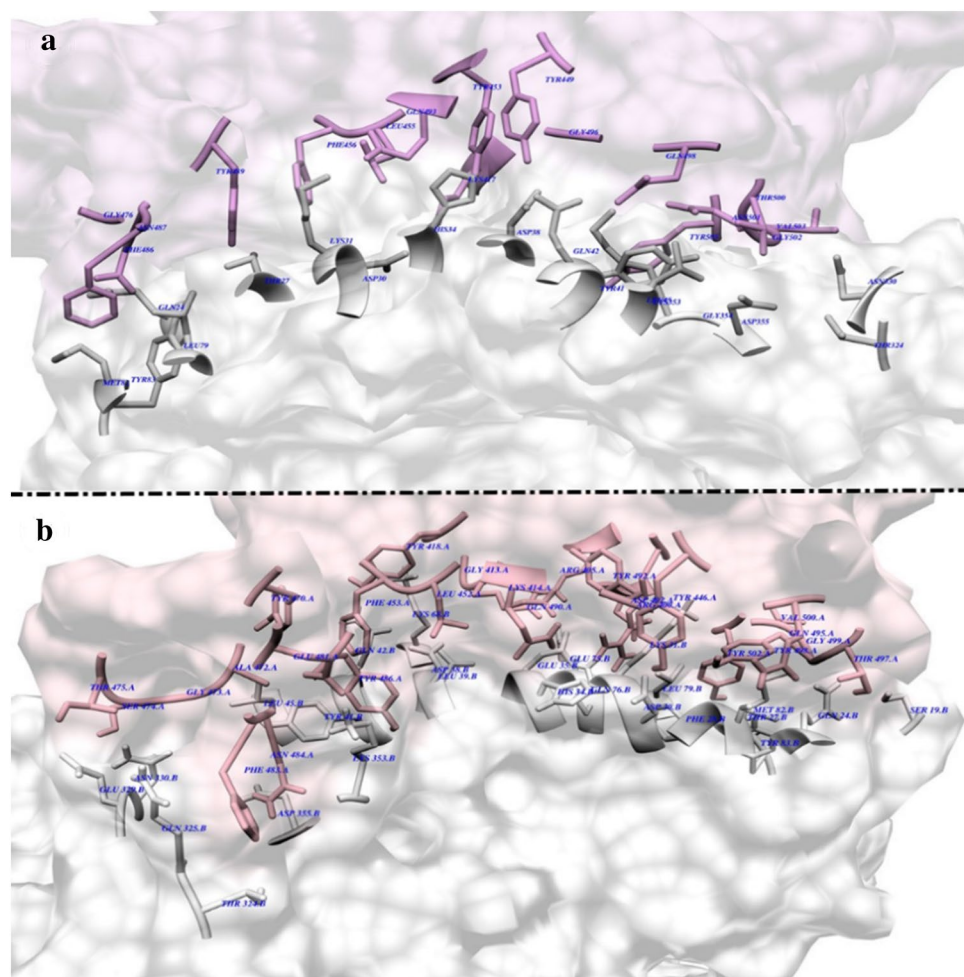
Fig. 3 The interaction orientation of ACE2 (in grey) with the wild type (a: magenta) and mutant (b: violet red) RBD of the spike proteins (Color figure online)

mainly preserved (Fig. 3). The interaction site for both complexes exactly includes the RBD of the spike protein and two N-terminal helices regions of the ACE2 protein. However, the mutant spike protein seems to be tilted clockwise to interact with the same region of the ACE2 protein. This means that although the interaction site is shared between wild type and mutant spike proteins, the involving amino acids and the binding energy should be different.

Refined Docked Complexes

The docked complex was refined by the refinement servers. Since the docked complex is refined by the refinement server, the involved amino acids could be derived more accurately. Although the overall interactions occurred with the same orientation, the amino acids at the protein–protein interface were different (Fig. 4). Supplementary Table 2 lists the amino acids involved in the interface of the ACE2 and spike protein for mutant protein and the 6M17 complex. To have a reliable comparison between the docked proteins and the control complex, the structures were trimmed to include the ACE2 receptor and the RBD region of the spike protein.

Fig. 4 The interaction orientation of ACE2 wild type and mutant RBDs of the Spike protein and interface residues. The upper panel **a** shows the wild type RBD (plum) engaged in ACE2 (white). The lower panel **b** illustrates the engagement of mutant RBD (pink) in ACE2 (white). In both panels the interface residues are presented as ribbons the whole molecules are presented as transplant surfaces. Labels are including three letter code of amino acids, residue number, and the related chain. In both complexes RBDs are chain A and ACE2 are chain B. The different orientations of RBDs in complex with ACE2 are evidenced in the images (Color figure online)



The binding ΔG was calculated to be $-11.8 \text{ kcal mol}^{-1}$ for the wild-type spike and the ACE2 complex and $-13.1 \text{ kcal mol}^{-1}$ for the mutant spike and the ACE2 complex. This means that the interaction between the mutant spike and the ACE2 is more stable.

The 2D Interaction Plot

The 2D interaction plot (Fig. 5) of the ACE2 and spike protein indicates that several amino acids from both proteins are involved in hydrogen bonds and hydrophobic interactions. The number of hydrogen bonds and hydrophobic interactions between the ACE2 (13 amino acids with hydrogen bonds and 13 amino acids with hydrophobic interactions) and the mutant spike protein (12 amino acids with hydrogen bonds and 16 amino acids with hydrophobic interactions) is higher than the interactions between the ACE2 (8 amino acids with hydrogen bonds and six amino acids with hydrophobic interactions) and the wild type spike protein (eight amino acids with hydrogen bonds and eight amino acids with hydrophobic interactions).

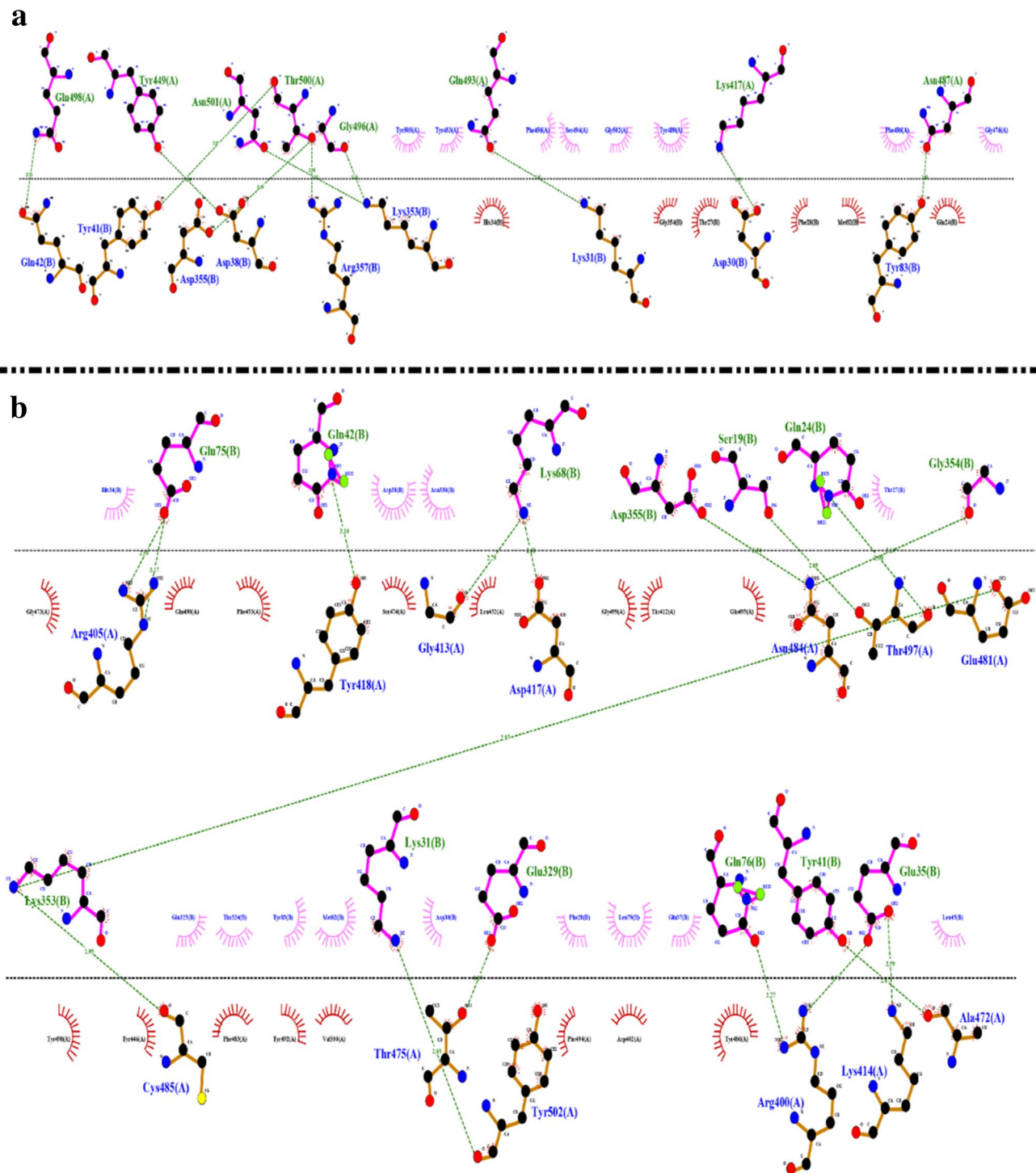


Fig. 5 The 2D plot of interactions between the ACE2 and wild type (a) and mutant (b) spike proteins (the amino acids with chain ID of B are from ACE2 and the amino acids with chain ID of A are from spike protein)

Residue Centrality

The residue interaction network of the ACE2 and the spike protein (both wild type and mutant) was subjected to

centrality analysis to determine the functional residues of the interacting proteins. Due to the complexity of derived networks, the interface residues between ACE2 (Chain B) and spike proteins (chain A) were extracted from the whole

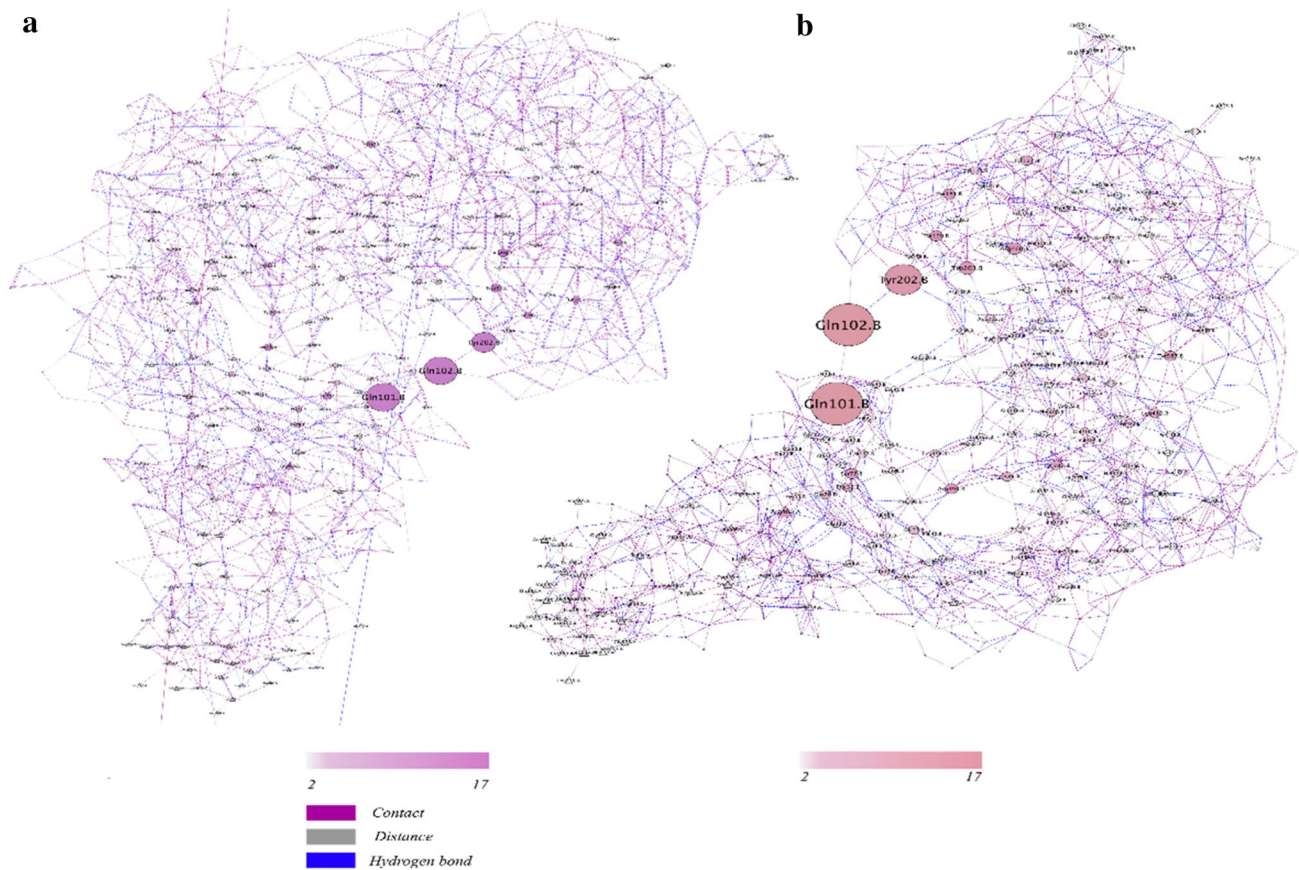


Fig. 6 The illustration of residues of significant centrality scores in the interaction network. The nodes are representative of residues in the structure; triangles are chain A (RBD) and ellipses are chain B (ACE2). The sizes of nodes are proportional to the Z-scores of cen-

trality. The color codes below each network define the Z-scores and edges. The left panel **a** is wildtype RBD and right panel **b** is the mutant RBD

network to have a more clear view (Fig. 6a). Moreover, the residues with significant centrality were separated and illustrated in Fig. 6b. Residues with Z-scores higher than two were considered significant and represented in Supplementary Table 3.

Cavities and Residue Depth

The search for cavities within the structure of the wild-type and mutant spike proteins was performed using the Molegro Virtual Docker. For each structure, ten cavities were detected at different locations and sizes. The position, volume, and surface of the found cavities are listed in Supplementary Table 4. The structural position of the cavities is depicted in Supplementary Fig. 1. The cavities of the wild-type spike protein are less expanded than the mutant spike protein. Although some of the cavities are at equivalent positions, some of them are located at different locations within the wild-type and mutant spike proteins. The results of the residue depth analyses are illustrated in Supplementary Fig. 2.

The obtained results indicate that the relative residue depth of mutant spike protein is higher than the wild-type spike protein.

Protein–Protein Interactions Between Mutated Spike and the Engineered ACE2

The structures of the engineered ACE2 and both mutant and wild type spike proteins were docked, and the resulting complexes were refined to get the most native-like interaction profiles. The obtained results indicated that the mutant and wild type RBD bind to the two N-terminal helices region of the engineered ACE2 protein. The two N-terminal helices region within the ACE2 structure is the target site for the interactions between the wild type ACE2 and wild type RBD proteins. As depicted in Fig. 7, both wild type and mutant RBDs occupy the same engineered ACE2 region with slightly differing orientations. The binding affinity analyses showed that the binding energy of both complexes is surprisingly identical. Both RBD/ACE2 complexes interact with

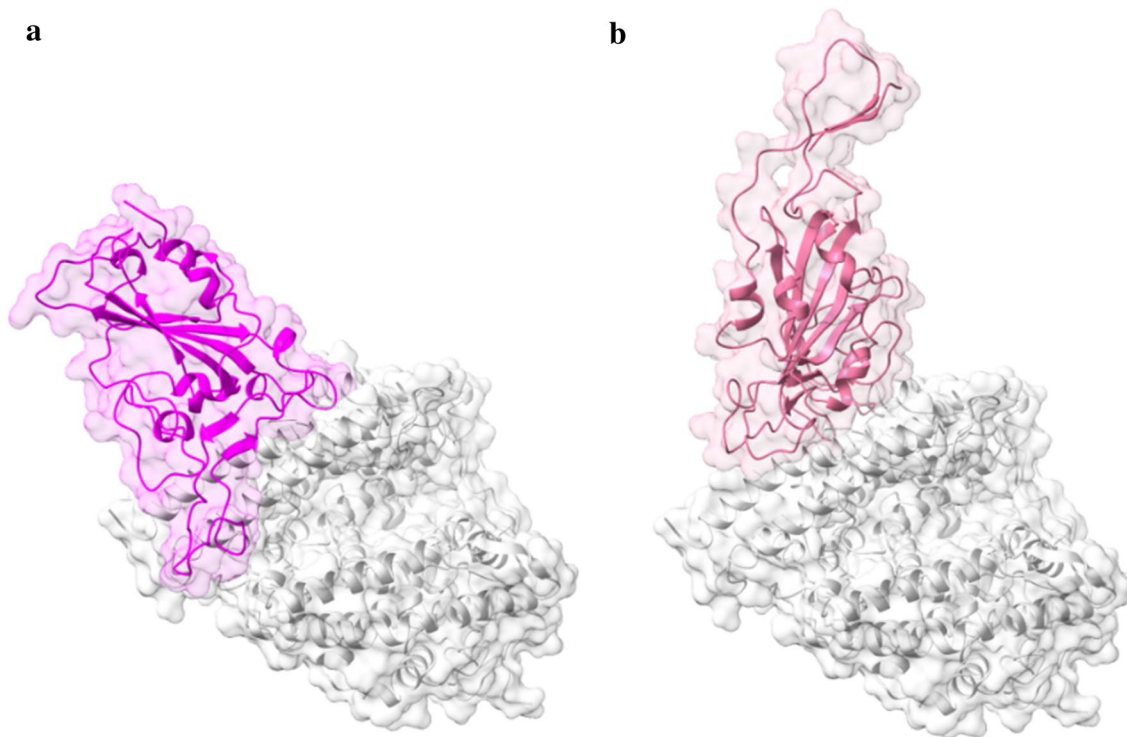


Fig. 7 The interaction orientation between the engineered ACE2 (light gray) and both wild type (**a**: magenta) and mutant (**b**: violet-red) spike proteins (Color figure online)

the binding affinity of $-18.5 \text{ kcal mol}^{-1}$. This is while the binding affinity between the wild-type RBD and the wild type ACE2 (complex adapted from PDB ID of 6M17) is calculated to be significantly lower ($-11.4 \text{ kcal mol}^{-1}$).

Discussion

Genome sequencing projects could pave the way for scientists to identify and characterize the novel viral variants for ongoing infections. Viral replication provides the ground for their mutation and the production of novel variants. This issue is even more concerning in SARS-CoV-2, due to an unprecedented number of individuals infected with this virus. The drastic number of infected individuals accentuates the possibility of SARS-CoV-2 replication and mutation. Therefore, emerging novel variants of SARS-CoV-2, with an accumulated number of mutations seem inevitable during different stages of the COVID19 pandemic (Weisblum et al. 2020). These amino acid mutations could have various implications ranging from slight structural changes in protein folding to imperative changes in intermolecular interactions with other proteins. The consequences of the adapted mutations would comply with the type of substituted amino acid and the location of the substitution within the protein structure (Gyulkhandanyan et al. 2020; Ittisoponpisan et al.

2019; Kucukkal et al. 2015; Thusberg and Vihinen 2009; Yates and Sternberg 2013). In the present study, we have studied the possible therapeutic and structural consequences of the mutations in the spike protein of the lineage B.1.1.7. This study considers all ten mutations of the spike protein in the lineage B.1.1.7.

As expected, the homology modeling approach was more successful in the prediction of 3D structure for the mutated spike protein. The highest RMSF values belonged to the residues 329-510, which corresponded to the RBD domain of the protein. These results are in line with the previous studies (He et al. 2020; Kalathiya et al. 2020). The monomeric spike protein in the solvent environment represented a movement from the “up” towards the “down” for the RBD domain. This movement aligns with the spike flexibility and its continuous up and down motion (Melero et al. 2020). Moreover, this movement indicates that the obtained mutations do not hamper the flexibility of the spike protein to alter between up and down conformation. The fully mutated structure of the spike protein was docked against the ACE2 protein to analyze the changes in their interactions. The obtained results indicate that despite the high number of mutations, the spike protein targets the same region of the ACE2 protein as the wild type spike protein (Ortega et al. 2020). Although the N-terminus helices of both wild-type and mutated spike

proteins interact with the same region of the ACE2, the orientation of the interactions is different. This change in the interaction orientation has led to the altered list of involved amino acids in the interaction surface. This change is also evident in the list of central amino acids. Along with shared central amino acids, there are several different central amino acids in the interface of the both wild-type and mutated spike proteins and the ACE2 protein. This means that the exerted mutation, in part, has changed the functional roles of the interface amino acids (Meng et al. 2021). It's noteworthy that the ACE2 amino acids with higher centrality are more conserved. This property indicates that the interaction between (wild-type and mutated) spike proteins and the ACE2 protein is mediated by similar amino acids on the ACE2 surface. Moreover, despite a large number of shared amino acids with similar interactions, the number of interacting amino acids (via hydrogen bonds or hydrophobic interactions) is higher in the mutated spike protein. This difference could be construed as higher affinity of the mutated spike protein against the ACE2 protein. This notion was confirmed by the binding affinity analyses. Although we have used a fully mutated version of spike protein, the obtained results are in line with the observations of prior studies with selected mutations of the lineage B.1.1.7 (Ahmed et al. 2021; Natarajan et al. 2021; Ortega et al. 2021; Shahhosseini et al. 2021; Villoutreix et al. 2021). The N501Y substitution is the only mutation within the RBD region of the spike protein. Although this mutation has substituted a hydrogen bond (via Asparagine) with a hydrophobic interaction (via Tyrosine), the overall number of hydrogen bonds and hydrophobic interaction is increased at the interface of the spike and ACE2. In line with our results, Yang et al. have confirmed that the mutated spike protein has an increased binding affinity against the ACE2 receptor. They also confirmed the key role of N501Y substitution in binding affinity enhancement. They have demonstrated that N501Y substitution introduces a π - π interaction that enhances RBD binding to ACE2 receptor. However, they raised a concern about the possible role of this substitution in abolishing the binding of a potent neutralizing antibody. They have shown that a cocktail of neutralizing antibodies could tackle this issue (Yang et al. 2021). Therefore, the structural changes enforced by all ten mutations of the spike protein, compensate for the loss of a hydrogen bond from N501Y substitution. This observation reveals that the mutated spike protein should be investigated in its fully mutated (10 mutations) status to understand the consequences of the obtained mutations. In line with this notion, Cai et al. have shown that a fully mutated spike protein is endowed with increased accessibility of its RBD and increased binding affinity for ACE2 receptor. These features could result in increased transmissibility

and risk of mortality. Higher binding affinity and RBD accessibility would lead to efficient infection of additional cell types, which express low levels of ACE2 (Cai et al. 2021; Cerutti et al. 2021).

Cavity sites are locations of protein structure where it binds to other proteins or biological macromolecules. Therefore, these sites are essential for the functional annotation of proteins (Naidu and Suneetha 2013). Although the cavities of the wild type and mutated spike proteins share some similarities regarding their location and volume, the central cavity of the mutated spike protein is much bigger than the wild type spike protein. This cavity size makes the mutated spike protein become less compact and relatively higher residue depth. We believe that the P681H mutation could be one of the contributing factors to this change. This residue is located between S1 and S2 domains in the spike protein and is one of the four residues comprising the furin cleavage site (Frampton et al. 2021). The change from a nonpolar amino acid to a positively charged one, increased the number of positively charged amino acids and improved the virus interaction with the furin cleavage site (Hoffmann et al. 2020). This property makes the mutant spike protein more accessible and prone to protease cleavage. This observation is concordant with the study conducted by Lubinski et al. (Lubinski et al. 2021). They have shown that the P681H mutation may increase the spike cleavage by furin-like proteases. This property coupled with higher affinity against ACE2 protein could lead to better viral fusion into the target cells and consequently better pathogenicity.

The interactions between the engineered ACE2 and mutant spike protein showed that the N501Y substitution has made structural changes within the RBD and, by extent RBM of the mutant spike protein. These changes were extended enough to alter the orientation in which the RBD chooses to dock on the engineered ACE2 molecule. However, since the target region (two N-terminal helices regions) to land on the engineered ACE2 protein is preserved, the primary backbone of interactions should be remained intact. Identical binding energies for engineered-ACE2/wild-type-RBD and engineered-ACE2/mutant-RBD complexes reveal that any loss of stabilizing interaction, due to the N501Y substitution, is compensated with novel and equivalent interactions. These results could be construed as the ability of the engineered ACE2 protein to prohibit the interactions between the cellular ACE2 and the mutant spike protein. This property turns out to be an appealing characteristic for the engineered ACE2, as it can block the main mechanism of SARS-CoV-2 to fuse with the target cell both for wild-type virus and the lineage B.1.1.7. The obtained results corroborate the promising nature of the engineered ACE2 to act as a therapeutic agent (Monteil et al. 2020; Payandeh et al. 2020; Yang et al. 2020) emerging mutant variants of SARS-CoV-2.

Conclusion

It could be concluded that the accumulation of mutations within the sequence of the spike protein have led to pronounced structural changes. These changes have brought about higher affinity against the ACE2 protein, more likely cleavage by furin-like proteases, and altered yet identically strong interaction with an ACE2 based therapeutic drug. These properties are concordant with higher transmissibility and higher virulent nature of mutant virus. This study delineates the potential mechanism behind the enhanced virulence of lineage B.1.1.7, increased frequency of reinfection, and possible blockade by engineered ACE2 based therapeutic drugs. The attained results could be supplemented with immunological analyses to pave the way to understand the mechanism behind the increased resistance of lineage B.1.1.7 against specific antibodies from the sera of convalescence patients and vaccinated individuals.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10989-021-10346-1>.

Acknowledgements The authors wish to thank Shahid Rajaei Teacher Training University for supporting the conduct of this research.

Funding Not applicable.

Data Availability The authors confirm that the *data* supporting the findings of this study are available within the article and supplementary materials.

Declarations

Conflict of interest The author declare that they have no conflicts of interest.

References

- Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1:19–25
- Ahmed W, Phillip AM, Biswas KH (2021) Stable interaction of the UK B.1.1.7 lineage SARS-CoV-2 S1 spike N501Y mutant with ACE2 revealed by molecular dynamics simulation. *bioRxiv*. <https://doi.org/10.1101/2021.01.07.425307>
- Asadi-Ghalehni M et al (2017) A novel recombinant anti-epidermal growth factor receptor peptide vaccine capable of active immunization and reduction of tumor volume in a mouse model. *Microbiol Immunol* 61:531–538
- Asadi-Ghalehni M, Rasaei MJ, Asl NN, Khosravani M, Rajabibazl M, Modjtahedi H, Sadroddiny E (2018) Construction of a recombinant phage-vaccine capable of reducing the growth rate of an established LL2 tumor model. *Iranian Journal of Allergy. Asthma Immunol* 17:240–249
- Boehm M, Nabel EG (2002) Angiotensin-converting enzyme 2—a new cardiac regulator. *N Engl J Med* 347:1795–1797
- Cai Y et al (2021) Structural basis for enhanced infectivity and immune evasion of SARS-CoV-2 variants. *bioRxiv*. <https://doi.org/10.1101/2021.04.13.439709>
- Cerutti G et al (2021) Structural basis for accommodation of emerging B.1.351 and B.1.1.7 variants by two potent SARS-CoV-2 neutralizing antibodies. *Structure*. <https://doi.org/10.1016/j.str.2021.05.014>
- Chand M et al (2020) Investigation of novel SARS-COV-2 variant: Variant of Concern 202012/01. *Public Health England* 21.
- Clark LA et al (2006) Affinity enhancement of an in vivo matured therapeutic antibody using structure-based computational design. *Protein Sci* 15:949–960
- Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N- \log (N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
- Davies NG et al (2021) Estimated transmissibility and severity of novel SARS-CoV-2 variant of concern 202012/01 in England. *MedRxiv*. <https://doi.org/10.1101/2020.12.24.20248822>
- Donoghue M et al (2000) A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circ Res* 87:e1–e9
- Frampton D et al (2021) Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B. 1.1. 7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *Lancet Infect Dis*. [https://doi.org/10.1016/S1473-3099\(21\)00170-5](https://doi.org/10.1016/S1473-3099(21)00170-5)
- Gyulkhanyan A, Rezaie AR, Roumenina L, Lagarde N, Fremaux-Bacchi V, Miteva MA, Villoutreix BO (2020) Analysis of protein missense alterations by combining sequence- and structure-based methods. *Mol Genet Genomic Med* 8:e1166
- Hashemi ZS et al (2019) Serum DKK1 is correlated with γ peak of serum electrophoresis in multiple myeloma: a multicenter biomarker study. *Biomark Med* 13:1297–1306
- Hashemi ZS et al (2021) In silico approaches for the design and optimization of interfering peptides against protein-protein interactions. *Front Mol Biosci* 8:282
- He J, Tao H, Yan Y, Huang S-Y, Xiao Y (2020) Molecular mechanism of evolution and human infection with SARS-CoV-2. *Viruses* 12:428
- Hess B, Bekker H, Berendsen HJ, Fraaije JG (1997) LINCS: a linear constraint solver for molecular simulations. *J Comput Chem* 18:1463–1472
- Hoffmann M et al (2020) SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181:271–280.e278
- Ittisoponpisan S, Islam SA, Khanna T, Alhuzimi E, David A, Sternberg MJ (2019) Can predicted protein 3D structures provide reliable insights into whether missense variants are disease associated? *J Mol Biol* 431:2197–2212
- Jamadi RH, Khalili S, Mirzapour T, Yaghoubi H, Hashemi ZS, Mard-Soltani M, Jafarizadeh M (2020) Anticancer activity of brevinin-2R peptide and its two analogues against myelogenous leukemia cell line as natural treatments: an in vitro study. *Int J Pept Res Ther* 26:1013–1020
- Kalathiya U et al (2020) Highly conserved homotrimer cavity formed by the SARS-CoV-2 spike glycoprotein: a novel binding site. *J Clin Med* 9:1473
- Khalili S, Rasaei MJ, Mousavi SL, Amani J, Jahangiri A, Borna H (2017) In silico prediction and in vitro verification of a novel multi-epitope antigen for HBV detection. *Mol Genet Microbiol Virol* 32:230–240
- Khodashenas S, Khalili S, Moghadam MF (2019) A cell ELISA based method for exosome detection in diagnostic and therapeutic applications. *Biotechnol Lett* 41:523–531

- Kucukkal TG, Petukh M, Li L, Alexov E (2015) Structural and physico-chemical effects of disease and non-disease nsSNPs on proteins. *Curr Opin Struct Biol* 32:18–24
- Lubinski B, Tang T, Daniel S, Jaimes JA, Whittaker G (2021) Functional evaluation of proteolytic activation for the SARS-CoV-2 variant B. 1.1. 7: role of the P681H mutation. *bioRxiv*. <https://doi.org/10.1101/2021.04.06.438731>
- Mard-Soltani M, Rasaei MJ, Khalili S, Sheikhi A-K, Hedayati M, Ghaderi-Zefrehi H, Alasvand M (2018) The effect of differentially designed fusion proteins to elicit efficient anti-human thyroid stimulating hormone immune responses. *Iran J Allergy Asthma Immunol* 17:158–170
- Melero R et al. (2020) Continuous flexibility analysis of SARS-CoV-2 spike prefusion structures. *IUCr J* 7:1059–1069
- Meng B et al (2021) Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B. 1.1. 7. *Cell reports* 35:109292
- Meng X, Li W, Peng X, Li Y, Li M (2021) Protein interaction networks: centrality, modularity, dynamics, and applications. *Front Comput Sci* 15:156902. <https://doi.org/10.1007/s11704-020-8179-0>
- Monteil V et al (2020) Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2. *Cell* 181:905–913.e907
- Naidu CK, Suneetha Y (2013) Structure-based computational analysis of protein binding sites for function and druggability in macrophage infectivity potentiator (MIP) protein of legionella pneumophila. *Curr Trends Biotechnol Pharm* 7:847–853
- Natarajan MA, Javali PS, Pandian CJ, Ali MA, Srivastava V, Jeyaraman J (2021) Computational investigation of increased virulence and pathogenesis of SARS-CoV-2 lineage B. 1.1. 7. *bioRxiv*. <https://doi.org/10.1101/2021.01.25.428190>
- Noori E, Kazemi B, Bandehpour M, Zali H, Khalesi B, Khalili S (2020) Deciphering crucial genes in coeliac disease by bioinformatics analysis. *Autoimmunity* 53:102–113
- Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676
- Ortega JT, Serrano ML, Pujol FH, Rangel HR (2020) Role of changes in SARS-CoV-2 spike protein in the interaction with the human ACE2 receptor: an in silico analysis. *EXCLI J* 19:410–417. <https://doi.org/10.17179/excli2020-1167>
- Ortega JT, Pujol FH, Jastrzebska B, Rangel HR (2021) Mutations in the SARS-CoV-2 spike protein modulate the virus affinity to the human ACE2 receptor, an in silico analysis. *EXCLI J* 20:585
- Payandeh Z et al (2020) Design of an engineered ACE2 as a novel therapeutics against COVID-19. *J Theor Biol* 505:110425
- Pourzardosht N et al (2020) Liothyronine could block the programmed death-ligand 1 (PDL1) activity: an e-Pharmacophore modeling and virtual screening study. *J Recept Signal Transduction*: 1–9
- Rahbar MR et al (2019) Trimeric autotransporter adhesins in *Acinetobacter baumannii*, coincidental evolution at work. *Infection, Genetics and Evolution* 71:116–127
- Rambaut A et al. (2020) Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. *Genom Epidemiol* 1–5
- Ramezani A et al (2019) Structure based screening for inhibitory therapeutics of CTLA-4 unveiled new insights about biology of ACTH. *Int J Pept Res Ther* 1–11
- SARS-CoV-2 Variant—United Kingdom of Great Britain and Northern Ireland (2020)
- Shahhosseini N, Babuadze GG, Wong G, Kobinger GP (2021) Mutation signatures and in silico docking of novel SARS-CoV-2 variants of concern. *Microorganisms* 9:926
- Thusberg J, Vihinen M (2009) Pathogenic or not? And if so, then how? Studying the effects of missense mutations using bioinformatics methods. *Hum Mutat* 30:703–714
- Toyoshima Y, Nemoto K, Matsumoto S, Nakamura Y, Kiyotani K (2020) SARS-CoV-2 genomic variations associated with mortality rate of COVID-19. *J Hum Genet* 65:1075–1082. <https://doi.org/10.1038/s10038-020-0808-9>
- Villoutreix BO, Calvez V, Marcelin A-G, Khatib A-M (2021) In silico investigation of the new UK (B. 1.1. 7) and South African (501y.v2) SARS-CoV-2 variants with a focus at the ace2–spike rbd interface. *Int J Mol Sci* 22:1695
- Weisblum Y et al (2020) Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *Elife* 9:e61312
- Wu L et al (2020) Broad host range of SARS-CoV-2 and the molecular basis for SARS-CoV-2 binding to cat ACE2. *Cell Discov* 6:68. <https://doi.org/10.1038/s41421-020-00210-9>
- Xu J et al (2009) Improving GPX activity of selenium-containing human single-chain Fv antibody by site-directed mutation based on the structural analysis. *J Mol Recognit* 22:293–300
- Yang J et al (2020) Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nat Commun* 11:1–10
- Yang T-J et al (2021) Effect of SARS-CoV-2 B. 1.1. 7 mutations on spike protein structure and function. *Nat Struct Mol Biol* 28:731–739
- Yates CM, Sternberg MJ (2013) The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein–protein interactions. *J Mol Biol* 425:3949–3963
- Zhang Y, He X, Zhai J, Ji B, Man VH, Wang J (2021) In silico binding profile characterization of SARS-CoV-2 spike protein and its mutants bound to human ACE2 receptor. *Brief Bioinform*. <https://doi.org/10.1093/bib/bbab188>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.