



# Comparative molecular dynamics study of the receptor-binding domains in SARS-CoV-2 and SARS-CoV and the effects of mutations on the binding affinity

Shokouh Rezaei<sup>a</sup>, Yahya Sefidbakht<sup>a</sup>  and Vuk Uskoković<sup>b</sup> 

<sup>a</sup>Protein Research Center, Shahid Behesti University, Tehran, Iran; <sup>b</sup>Advanced Materials and Nanobiotechnology Laboratory, TardigradeNano, Irvine, CA, USA

Communicated by Ramaswamy H. Sarma

## ABSTRACT

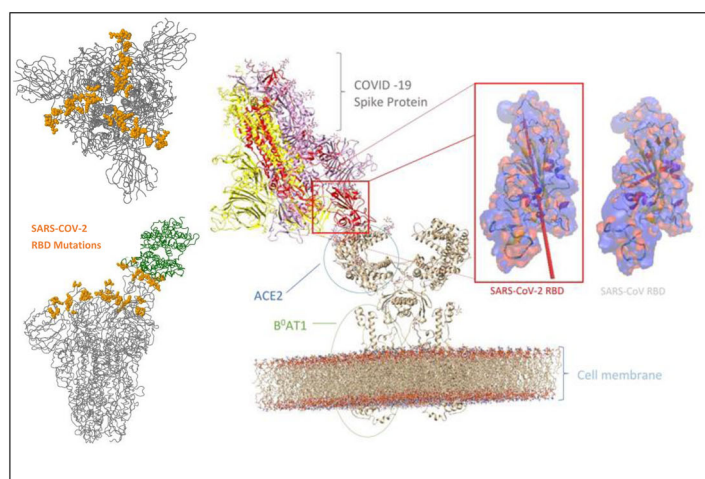
Here, we report on a computational comparison of the receptor-binding domains (RBDs) on the spike proteins of severe respiratory syndrome coronavirus-2 (SARS-CoV-2) and SARS-CoV in free forms and as complexes with angiotensin-converting enzyme 2 (ACE2) as their receptor in humans. The impact of 42 mutations discovered so far on the structure and thermodynamics of SARS-CoV-2 RBD was also assessed. The binding affinity of SARS-CoV-2 RBD for ACE2 is higher than that of SARS-CoV RBD. The binding of COVA2-04 antibody to SARS-CoV-2 RBD is more energetically favorable than the binding of COVA2-39, but also less favorable than the formation of SARS-CoV-2 RBD-ACE2 complex. The net charge, the dipole moment and hydrophilicity of SARS-CoV-2 RBD are higher than those of SARS-CoV RBD, producing lower solvation and surface free energies and thus lower stability. The structure of SARS-CoV-2 RBD is also more flexible and more open, with a larger solvent-accessible surface area than that of SARS-CoV RBD. Single-point mutations have a dramatic effect on distribution of charges, most prominently at the site of substitution and its immediate vicinity. These charge alterations alter the free energy landscape, while X→F mutations exhibit a stabilizing effect on the RBD structure through  $\pi$  stacking. F456 and W436 emerge as two key residues governing the stability and affinity of the spike protein for its ACE2 receptor. These analyses of the structural differences and the impact of mutations on different viral strains and members of the coronavirus genera are an essential aid in the development of effective therapeutic strategies.

## ARTICLE HISTORY

Received 7 November 2020  
Accepted 3 December 2020

## KEYWORDS

Receptor-binding domain; MD simulation; SARS-CoV-2; SARS-CoV; mutations; residue interactions network; binding free energy




## 1. Introduction

Betacoronaviruses ( $\beta$ -CoVs) are the genera of coronaviruses (CoVs) that are of particular clinical significance due to their virulence and the ability to jump across species. Prior  $\beta$ -CoVs that proved infectious to humans included severe acute respiratory syndrome coronavirus (SARS-CoV) and middle east respiratory syndrome coronavirus (MERS-CoV) (Yin & Wunderink, 2018),

which were both highly pathogenic. The symptoms of infection with any of these two  $\beta$ -CoVs were severe and the corresponding mortalities relatively high, which allowed the medical teams to stop the spread of the virus before it reached pandemic proportions. In December 2019, however, in China, severe respiratory syndrome coronavirus-2 (SARS-CoV-2) was identified as the third pathogenic virus of the  $\beta$ -CoV lineage (Liu et al., 2020).

**CONTACT** Yahya Sefidbakht  [y\\_sefidbakht@sbu.ac.ir](mailto:y_sefidbakht@sbu.ac.ir)  Protein Research Center, Shahid Behesti University, Tehran, Iran

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/07391102.2020.1860829>.

© 2020 Informa UK Limited, trading as Taylor & Francis Group

Only a month after it had been reported in China, the spread of SARS-CoV-2 reached global proportions and in March 2020 the World Health Organization declared the pandemic, which is still unfolding. Unlike SARS-CoV and MERS-CoV, SARS-CoV-2 is typified by a broad range of symptoms produced in humans, spanning from total asymptomatic conditions to a bedridden and life-threatening state. This has allowed the virus to rapidly spread through the population, using seemingly healthy individuals as carriers. As of November 2020, the pandemic has not wound down and the virus continues to spread across the population at an accelerating rate. This state of affairs has led to the recruitment of the scientific community in the effort to provide scientific findings that would assist in curbing this ongoing pandemic.

One of the effects of the rapid spread and exchange of the viral material between people is the equally rapid mutability of the virus, which SARS-CoV-2 itself is particularly prone to in the first place owing to its ribonucleic genetic material. RNA viruses tend to be less stable than their DNA counterparts, but their mutability is significantly higher (Uskoković, 2020). This has led to concerns that the virus may mutate into even less treatable forms than those that have spread through the community until now. Recent epidemiological analyses do confirm these expectations, indicating that mutated SARS-CoV-2 appearing in the most recent waves of infection in urban areas may be more contagious than the earlier versions (Long et al., 2020). Detailed analyses of the effects of mutations on the molecular characteristics of SARS-CoV-2 are therefore of benefit, like the one we undertake here. Also, because any rational design of drugs targeting and neutralizing SARS-CoV-2 would be based on achieving a precise molecular recognition effect between the drug molecule and a receptor on the viral surface (or the interior under special circumstances), understanding these molecular interactions is of paramount importance.

CoVs in general contain 16 non-structural (nsp1-16) proteins and four structural ones. The structural ones include spike (S), membrane (M), nucleocapsid (N), and envelope (E) proteins (Jiang et al., 2020). The envelope-anchored spike protein is an essential determinant of the specificity of the viral interaction with the host because it protrudes the viral particle most prominently and is the first of its components to come into contact with the host cell. The spike protein can be cleaved by host proteases into the N-terminal of S1 subunit and the C-terminal of S2 subunit (Huang et al., 2020). The molecular recognition and attachment of the viral  $\beta$ -CoV particle to the host cell owes itself to the binding of the receptor-binding domain (RBD) on the S1 subunit of the spike protein to human angiotensin-converting enzyme 2 (ACE2) (Gui et al., 2017). The RBD possesses two conformations: 'up' and 'down'. While the 'up' conformation corresponds to the receptor-accessible state, the 'down' conformation corresponds to the receptor-inaccessible state. Out of the two conformations, the 'up' conformation is thought to be less stable (Roy et al., 2020; Wrapp et al., 2020). In addition, the RBD contains a receptor-binding motif (RBM), which is its most functional segment because of its direct involvement in the attachment to ACE2 (Li et al., 2005)

and whose physicochemical properties are the direct determinants of this binding affinity. Based on previous studies, it is clear that both SARS-CoV-2 and SARS-CoV infect host cells by the binding of the RBD to ACE2 (Lan et al., 2020; Lukassen et al., 2020).

In the last few decades, computational approaches have developed into a multitude of tools for forecasting the antigenic evolution of viruses. Accordingly, researchers investigate amino acid substitutions in protein regions corresponding to the antigenic change, after which conclusions of relevance for the efficient method of vaccine development are drawn (Klingen et al., 2018). With this in mind, this study is based on one such computational assessment of the structural comparison of the RBDs of SARS-CoV-2 and SARS-CoV, with and without the interface with ACE2. In addition, over 40 mutations detected so far in the RBD structure of SARS-CoV-2 were investigated in terms of their effect on selected physicochemical properties of these molecules. With regards to the method, the RBD structures of SARS-CoV-2 and SARS-CoV were subjected to 50 ns molecular dynamics (MD) minimizations. In addition, the residue interaction networks (RINs) approach was applied to investigate protein-protein interactions and mutation effects on the affinity of the RBD toward ACE2. In the RIN model, a network is computed and analyzed, such that its nodes and edges represent residues and bonds between the residue side chains, respectively (Piovesan et al., 2016). This approach can give insight into the mechanisms of protein folding, conformational alterations, and the effects of amino acid mutations on the protein structure and function (Dehury et al., 2020; Fonseca et al., 2020), with the major implications on the binding of antibodies such as COVA2-04 and COVA2-39 to the given mutated RBDs (Brouwer et al., 2020). As a result, these results are of significance for understanding the immune response to the infection, but also for developing novel therapies based on precisely targeted pharmacodynamics.

## 2. Computational methods

### 2.1. Sequence and structure alignment

The sequences of wild type spike proteins isolated from the SARS-CoV-2 and SARS-CoV viruses were retrieved from UniProtKB (<http://www.uniprot.org/>), under the identifier numbers of P0DTC2 and P59594, respectively. SARS-CoV-2 and SARS-CoV RBDs consist of 223 (319-541) and 222 (306-527) residues, respectively (Wang et al., 2020). The automated sequence alignment was performed using the EMBOSS Needleman-Wunsch method (Needleman & Wunsch, 1970; Rice et al., 2000). This tool indicates the percentage of identity and similarity between RBDs. I-TASSER server (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) was utilized to build the 3D structure of RBDs. I-TASSER, an online server for automated protein structure prediction, was used earlier for several biological studies (Yang & Zhang, 2015). The predicted structures (5 models) were saved in the PDB format and sorted according to the C-scores. The best model was selected based on its C-score. The quality of the predicted structure was confirmed by PROCHECK (Laskowski

et al., 1993, 1996) and ProSA, which calculate an overall quality score for 3D structures (S1 (A-B)) (Wiederstein & Sippl, 2007). The local environment of amino acids was checked through the WHAT IF coarse packing quality control, which should stay above  $-5$ . Mutations were also predicted with the use of the WHAT IF mutation prediction tool (Vriend & Sander, 1993).

## 2.2. Mutation and their effects

A complete list of SARS-CoV-2 and SARS-CoV RBD mutations in the GISAID database (<https://www.gisaid.org/>) was utilized. Then, multiple sequence alignments between the wild type and the variants were performed to obtain sequence variations in the SARS-CoV-2 RBD region. In addition, a list of SARS-CoV RBD mutations in UniProt was considered and one of the mutations reported therein and not reported elsewhere (R426G) was included in the analysis.

### 2.2.1. Effect of mutations on thermodynamic parameters

**2.2.1.1. Free energy binding and  $\Delta\Delta G$  (affinity and destabilizing).** The structures of SARS-CoV-2 and SARS-CoV RBD-ACE2 complexes (PDB codes 6M0J and 2AJF) (Lan et al., 2020) and the structure of RBD-COVA2-04 and RBD-COVA2-39 complexes (PDB codes 7JMO and 7JMP) were downloaded from the Protein Data Bank (Wu et al., 2020). Then, the PRODIGY (PROtein binDing enerGY) web server was applied to predict the binding affinity of protein-protein complexes (Vangone & Bonvin, 2015). DynaMut (Rodrigues et al., 2018) and SAAMBE (Pahari et al., 2020) were used to predict  $\Delta\Delta G$  (destabilization and affinity) for mutations with respect to the given protein-protein complexes.

**2.2.1.2. Solvation free energy.** ProWaVE (<https://www.pro-wave.org/>) was used to calculate the solvation free energy (SFE) of the RBD and its variants. ProWaVE is an automated server for calculating the SFE based on the molecular theory of solvation (Chong et al., 2011; Chong & Ham, 2014). ProWaVE investigates the structural properties and solvation thermodynamics for the wild and mutant types using fully atomistic, explicit-water MD simulations as well as three-dimensional reference interaction site model (3D-RISM) theory. Thereupon, the following aspects were tried to be elucidated with the use of the 3D-RISM theory: 1) the protein structure and the folding kinetics alteration after inducing a single-point mutation; 2) the impact of mutations on the hydrophobicity of the protein.

**2.2.1.3. Free energy surface.** Protein processes such as folding or aggregation can be explained in terms of the molecular free energy,  $\Delta G$ :

$$DG(R) = -k_B T [\ln P(R) - \ln P_{max}]$$

where  $k_B$  is the Boltzmann constant,  $P$  is the probability distribution of the molecular system along the coordinate  $R$ , and  $P_{max}$  denotes its maximum, which is subtracted from  $P(R)$  to ensure that  $\Delta G=0$  for the lowest free energy

minimum. Common choices for  $R$  are the root-mean-square deviation of atomic positions (RMSD), radius of gyration ( $R_g$ ), number of hydrogen bonds (H-bonds) or native contacts, or principal components. Here, the free energy was typically plotted at 310K along axes defined by two parameters, e.g.  $R_g$  and RMSD.

### 2.2.2. Effect of mutations on the electrostatic potential

The degrees of protonation of titrable side chains at pH 7.4 were predicted through the H++ server (<http://biophysics.cs.vt.edu/H++>), using known pKa values (Anandakrishnan et al., 2012). The electrostatic representation of the molecule was calculated using UCSF Chimera 1.14 (Pettersen et al., 2004).

## 2.3. Molecular dynamics simulation

Energy minimization and MD simulations were carried out by GROMACS (2020.2) utilizing the Optimized Potential for Liquid Simulations (OPLS) force field (Jorgensen & Tirado-rives, 1988). The structure produced by I-TASSER was used as an input. Water box was created with at least 1 nm (10 Å) distances from the protein using the SPC water model and applying boundary conditions. The system neutralization was done by adding  $\text{Na}^+$  and  $\text{Cl}^-$  ions at the concentration of 0.1 M. The MD simulation was carried out to examine the quality of the model structures by checking their stability via performing 50 ns simulations at a constant temperature 310 K (NVT). Energy minimization was performed in 50,000 steps to avoid any bad contacts generated while solvating the system. Then the NPT optimization was done for 100 ps. To increase the likelihood of achieving the appropriate structure, the MD simulation was performed for 50 ns using the OPLS force field. Finally, GROMACS tools were applied for the trajectory analysis, including the parameters such as RMSD, RMSF,  $R_g$ , SASA, and the number of H-bonds.

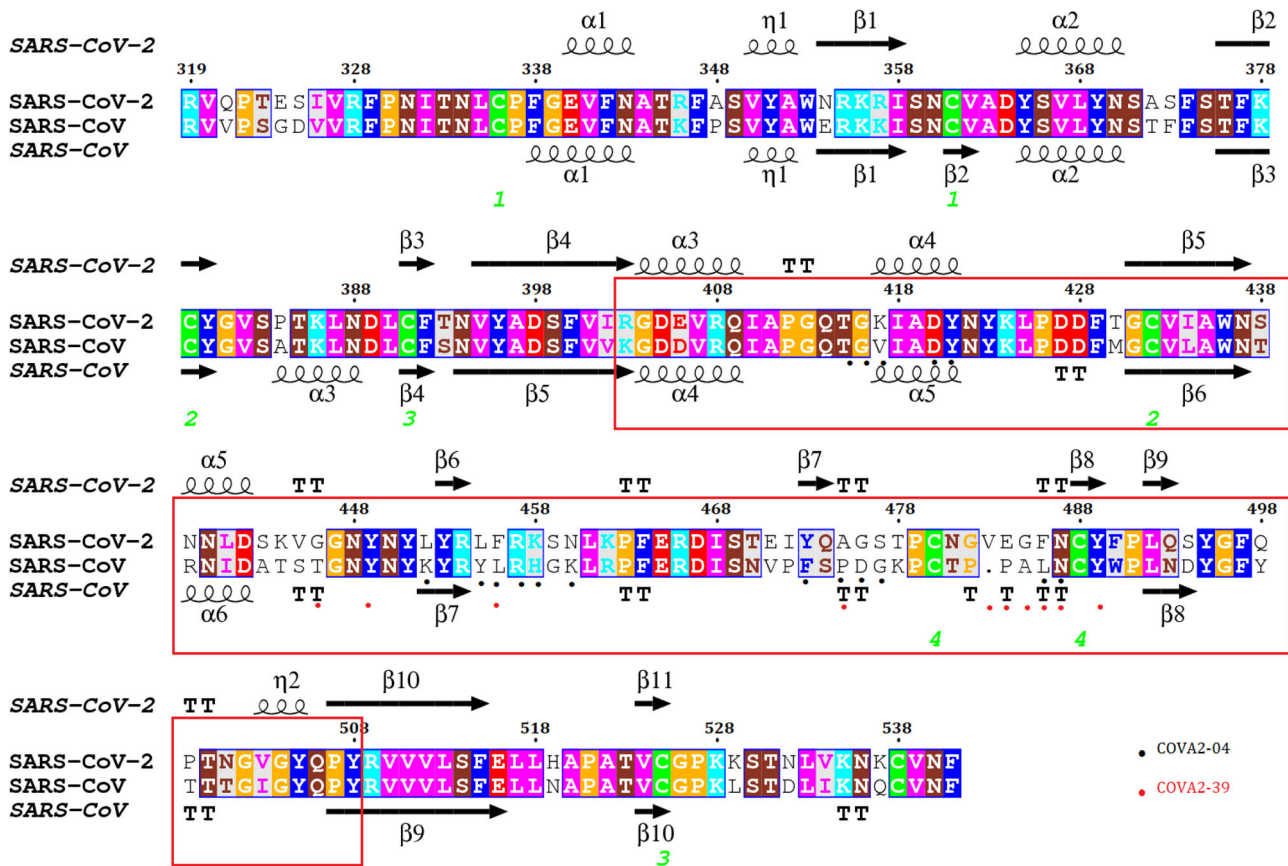
## 2.4. Principal component analysis

Eigenvectors of the covariance matrices and the projections of the first two principal components were calculated. The principal component analysis (PCA) was performed using GROMACS to reduce the dimensionality of the MD simulations data and identify the configuration space of a harmonic motion with only a few degrees of freedom.

## 2.5. RINs analysis

RING 2.0 (<http://old.protein.bio.unipd.it/ring/>) (Martin et al., 2011; Piovesan et al., 2016) and SARS-CoV-2 RBD-ACE2 complex (PDB ID: 6M0J) were used to generate the SARS-CoV-2 RBD-ACE2 RIN. Then, the RING output was visualized in the Cytoscape. MCODE (Bader & Hogue, 2003) plugin in Cytoscape was used to find clusters of the network. In general, clusters in the RINs for RBD-ACE2 and free RBD (including five or more nodes), stress and betweenness were used to identify the key residues. In addition, interactions at the contact surface of the RBD and ACE2 were investigated.





**Figure 1.** Structure and sequence alignment between the RBDs of SARS-CoV-2 and SARS-CoV. Red box indicates the RBM. Conservation of epitope residues is marked by red and black dots.

### 2.5.1. Key residues in the RBD-ACE2 complex and free RBD

Key residues were defined using stress and betweenness as the two elementary parameters of the local metrics. High betweenness values here indicate that the node is a mediator of interactions with other nodes and that it may be a key structural residue (Hu et al., 2014). In contrast, stress is defined as the number of shortest paths passing through a node.

### 2.6. Calculation of the membrane topology and the dipole moment

CHARMM-GUI Membrane Builder was applied to simulate the complex biological membrane and produce the representation of its topology (Allouche, 2011). Also, APBS-1.5 was used to calculate the dipole moment (Jurrus et al., 2018).

## 3. Results and discussion

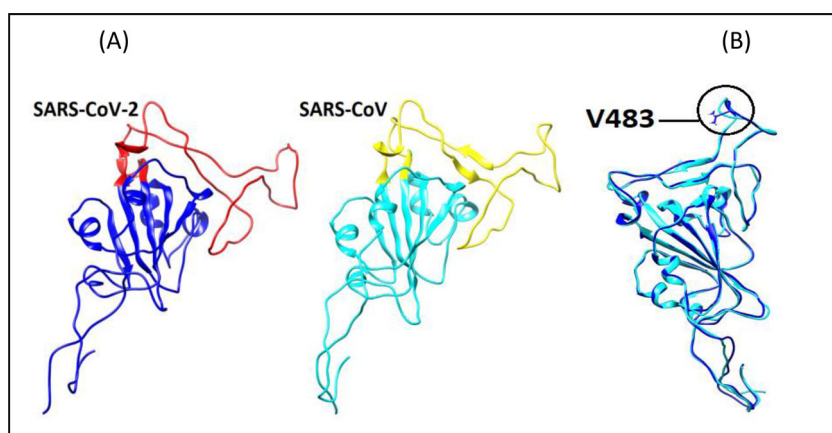
Investigating the effects of sequence variations in SARS-CoV-2 RBD is essential for the understanding of pathogenesis and the development of safe and effective prevention and treatment strategies for COVID-19, as through viral detection, vaccine design, and development of drugs (Zhou et al., 2020). One such study, gaining an insight into the basic aspects of thermodynamics and pharmacodynamics of the interaction

between wild-type and mutated SARS-CoV-2 RBDs and its receptor in the host cell has been undertaken here.

### 3.1. Sequence and structure alignment

The results of the amino acid sequence alignment revealed that the RBD of SARS-CoV-2 shares 73.1% identity and 82.1% similarity with the RBD of SARS-CoV (Figure 1). The secondary structure of the SARS-CoV-2 RBD contains five antiparallel  $\beta$  sheets ( $\beta 1$ ,  $\beta 2$ ,  $\beta 3$ ,  $\beta 4$  and  $\beta 7$ ) with short connecting helices and loops. The extended region between the  $\beta 4$  and  $\beta 7$  strands is the receptor-binding motif (RBM), stretching from R403 to Y508 in SARS-CoV-2 and from K403 to Y508 in SARS-CoV. The SARS-CoV-2 RBD has two  $\beta$  sheets less than the SARS-CoV RBD (Figure 2(A)), which is the first indicator of the higher structural flexibility of SARS-CoV-2 RBD compared to that of SARS-CoV RBD. Also, as shown in Figure 1, interestingly, the  $\alpha$ -helix ( $\alpha 3$ ) is formed at two different positions in the RBDs of SARS-CoV-2 and SARS-CoV. It appears that the presence of P384 in the (-VSPTKLN-) region prevents the formation of  $\alpha$ -helix in SARS-CoV-2 RBD.

Further, a total of nine cysteine residues were detected in the RBD of both viruses, eight of which involved four pairs of disulfide bonds. Three of those pairs aid in the stabilization of the  $\beta$  sheet structure, including Cys336–Cys361, Cys379–Cys432 and Cys391–Cys525; the remaining pair (Cys480–Cys488) is located in the distal end of the RBM loop. Also, an additional insertion (V483) was found in the



**Figure 2.** (A) Tertiary structure of the RBDs of SARS-CoV-2 and SARS-CoV. The RBM is colored red in the structure of SARS-CoV-2 and yellow in the structure of SARS-CoV. (B) Superimposition of the RBDs of SARS-CoV-2 and SARS-CoV.

(-AGSTPCNGVEGF-) loop of the SARS-CoV-2 RBD that is not found in the same region of the SARS-CoV RBD (-PDGKPCPPAL-) (Figure 2(B)). Sequence comparisons showed that some of the loop residues, such as V483, E484, G485, F486 and N487, bind to antibodies. Therefore, it is likely that V483 provides an additional site for the binding of antibodies, such as COVA2-04. It lies in-between the epitope residues A475 and F486, which were previously delineated as the binding sites for neutralizing antibodies (Yi et al., 2020).

### 3.2. Mutations and their effects

Figure 3(A–C) shows 42 different mutations identified in the various rare variants of the SARS-CoV-2 RBD. In Figure 3 they are listed according to the frequency of their appearance and include the following: N439K, S477N, T478I, G485R, G485S, G476S, E484Q, N501Y, N501S, N501T, S494P, R403K, G446V, E484K, A475V, S477I, L455F, F490S, F490L, Y453F, K417N, Q493L, F456L, T478K, T478A, Q506K, Y505W, G504D, E484A, G446S, S477G, V445A, S494L, F486L, S477R, V445I, E484D, G476A, V503F, V503I, S477T, and N437S. There is a set of mutations that occur specifically in the RBD of SARS-CoV, including G311R, K344R, F360S, R426G, N437D, L472P, N479K, D480G, T487S, and F501Y.

In general, investigating the impact of mutations, particularly on the conformation of the interacting proteins, is essential because mutations in proteins can have a dramatic effect on the protein folding and stability. Also, mutations alter the kinetics and thermodynamics of protein-protein interactions (Zhou et al., 2020), and these mutations can be either selectively beneficial to the organism via evolution or straightforwardly detrimental (Jubb et al., 2017). Therefore, the study of the effect of mutations is essential for different biomedical applications, including personalized medicine stemming from disease-associated mutation analyses or highly specific drug design and remedial interventions.

#### 3.2.1. Effect of mutations on thermodynamic parameters

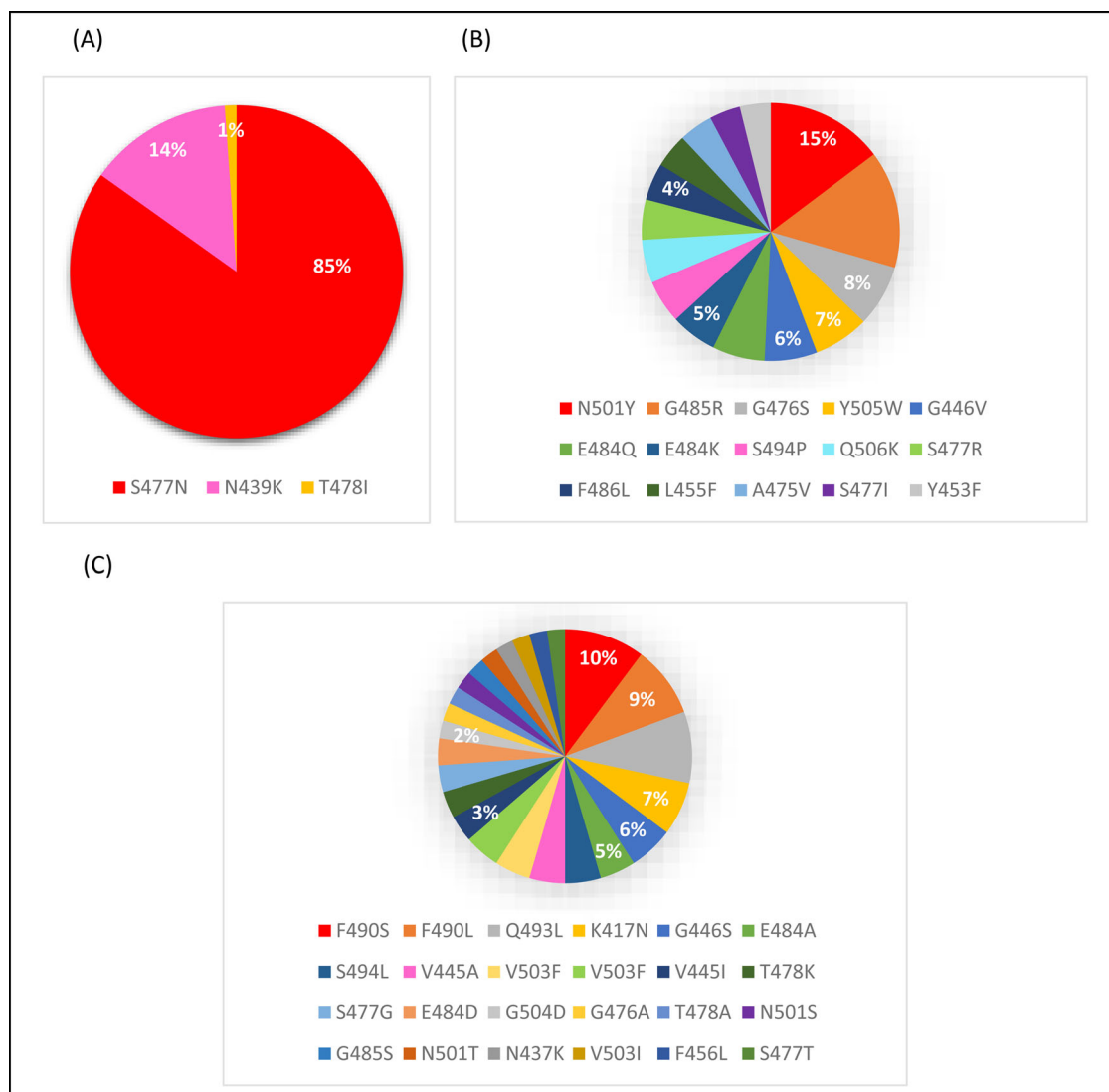
**3.2.1.1. Binding free energy and  $\Delta\Delta G$  (affinity and destabilizing).** Charting the thermodynamics of protein-protein

interactions is an important precondition for revealing their mechanisms of function and understanding the effects of disease-related conditions. The total binding free energies ( $\Delta G^{\text{Binding}}$ ) for SARS-CoV-2 and SARS-CoV RBD-ACE2 complexes were calculated. As it is shown in Table 1,  $\Delta G^{\text{Binding}}$  for SARS-CoV-2 RBD-ACE2 was  $-11.9 \text{ kcal mol}^{-1}$ ,  $1.1 \text{ kcal mol}^{-1}$  lower than that of SARS-CoV RBD-ACE2 ( $-10.8 \text{ kcal mol}^{-1}$ ), clearly indicating that the binding affinity of the SARS-CoV-2 RBD to ACE2 is higher than that of the SARS-CoV RBD.

In addition,  $\Delta G^{\text{Binding}}$  for two SARS-CoV-2 RBD-antibody complexes was computed. The results showed that  $\Delta G^{\text{Binding}}$  for RBD-COVA2-04 ( $-11.6 \text{ kcal mol}^{-1}$ ) was higher than that for RBD-COVA2-39 ( $-7.3 \text{ kcal mol}^{-1}$ ) (Table 2). Interestingly,  $\Delta G^{\text{Binding}}$  of the RBD-COVA2-04 complex was very close to the free energy of the SARS-CoV-2 RBD-ACE2 complex. Still, based on these thermodynamic considerations, the binding of the SARS-CoV-2 RBD to ACE2 is more energetically favorable than its binding to COVA2-04 or COVA2-39, which is only one aspect where the virus has an advantage over the immune protection. In addition to this, the dissociation constant of the SARS-CoV-2-ACE2 complex is relatively low at  $4\text{E}-09$  (Table 1), especially in comparison with that of the RBD-COVA2-39 complex, which is almost three orders of magnitude higher (Table 2).

A difference in the binding affinity to the epitope between the wild protein domain ( $\Delta G_{\text{w}}$ ) and the domain containing mutant residues ( $\Delta G_{\text{m}}$ ) is defined as  $\Delta\Delta G^{\text{Affinity}}$  ( $\Delta\Delta G = \Delta G_{\text{m}} - \Delta G_{\text{w}}$ ). This free energy difference can be a descriptor of how mutations influence the protein stability. Overall,  $\Delta\Delta G$  values below zero indicate that the mutation has caused destabilization of the protein; otherwise, it has caused its stabilization. In this analysis, the change in the binding affinity upon the mutation ( $\Delta\Delta G^{\text{Affinity}}$ ) and in the folding free energy ( $\Delta\Delta G^{\text{Destabilizing}}$ ) at the RBD-ACE2 (Figure 4(A,B)) and RBD-antibody (Figure 6(A,B)) interfaces were considered.

Among 25 analyzed mutations of the SARS-CoV-2 RBD-ACE2 complex, 14 had  $\Delta\Delta G < 0$  (Table S1). Interestingly, all of the X→F mutations (L445F, V503F, and Y453F) had  $\Delta\Delta G^{\text{Destabilizing}} > 0$  (Table S1) (Figures 4(A) and 5). This



**Figure 3.** Mutations in the RBD of SARS-CoV-2 that occurred between (A) 100 and 4000 times, (B) 10 and 50 times, and (C) 1 and 10 times.

**Table 1.** Binding free energy ( $\Delta G^{\text{Binding}}$ ) and dissociation constant ( $K_d$ ) predicted for the RBD-ACE2 complexes. \* NIS: Non-Interaction Surface.

Protein-protein complex	$\Delta G$ (kcal mol <sup>-1</sup> )	$K_d$ (M) at 37.0 °C	NIS charged	NIS apolar
SARS-CoV COMPLEX	-10.8	2.4E-08	26.32%	35.96%
SARS-CoV-2 COMPLEX	-11.9	4.0E-09	25.73%	35.06%

NIS = non-interaction surface.

**Table 2.** Binding free energies ( $\Delta G^{\text{Binding}}$ ) and dissociation constants ( $K_d$ ) predicted for the RBD-antibody complexes.

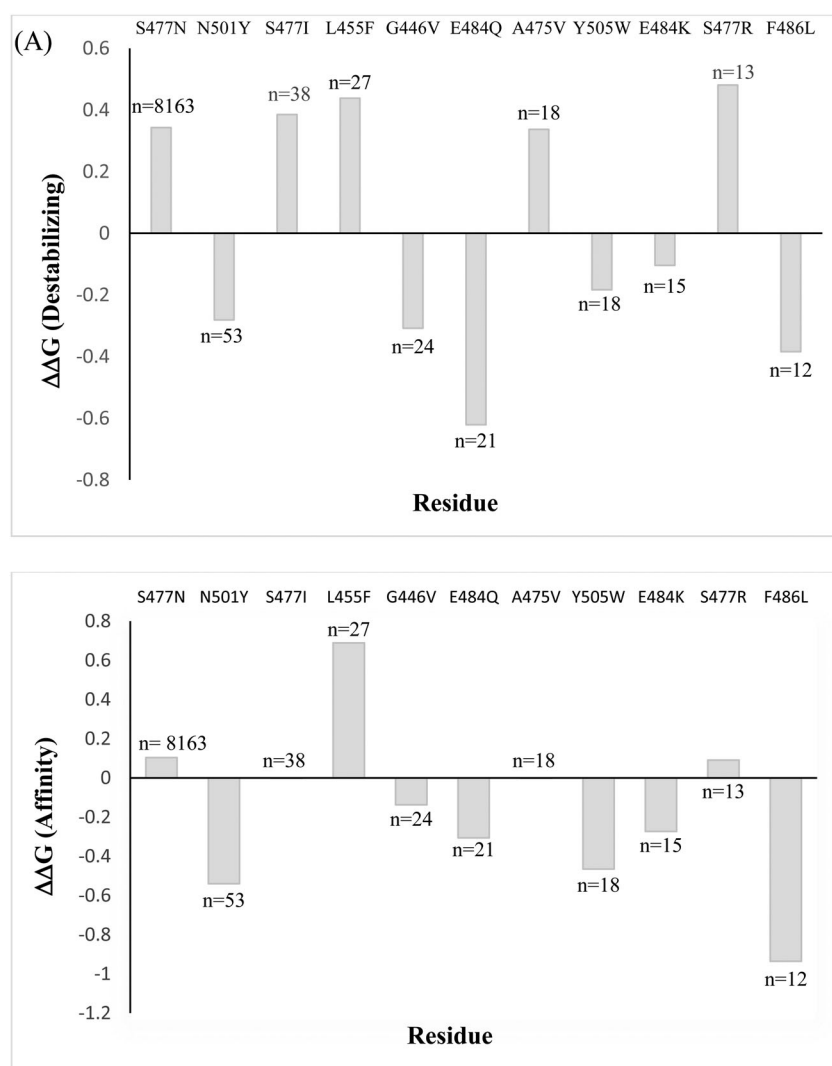
Complex	$\Delta G$ (kcal mol <sup>-1</sup> )	$K_d$ (M) at 37.0 °C	NIS charged	NIS apolar
RBD-COVA2-04	-11.6	7.2E-09	17.13%	40.19%
RBD-COVA2-39	-7.3	6.6E-06	16.88%	42.04%

NIS = non-interaction surface.

positive value may be due to the  $\pi$ -electrons of the phenyl ring, which can undergo stacking with the neighboring aromatic rings of amino acids such as F, W or Y. Therefore, promoting hydrophobic interactions, as in this case, can lead to protein folding and structural stability. In general, Figure 4(A) shows mutations leading to changes in protein interactions, which affect the stability of the structure.

Investigation of 21 mutations in the two RBD-antibody complexes, namely RBD-COVA2-04 or RBD-COVA2-39, indicated an equal variability of  $\Delta\Delta G^{\text{Affinity}}$  values (Figure 6(A,B)) as that found in the mutations of the SARS-CoV-2 RBD-ACE2 complex. These different values indicate a difference in the contribution of individual amino acids to protein-protein interaction. Equally, they suggest that not all mutations are equally important. Overall, as expected based on previous arguments, positive  $\Delta\Delta G^{\text{Affinity}}$  values caused a decrease in the protein-protein binding affinity (Table S1 and S2).

**3.2.1.2. Solvation free energy.** Since most proteins operate under the aqueous solvation, information about the interaction of the protein and its domains with water can be useful for predicting the stability of the protein and its specific interaction with ligands or the nonspecific binding to surfaces such as the cell membrane (König et al., 2013; Martini et al., 2013). In general, the information obtained from the investigation of the effect of side chains on the SFE can be important for understanding the protein folding and its stability (König et al., 2013). Moreover, defining the most appropriate mode of interaction of antibodies with antigens is a



**Figure 4.** (A)  $\Delta\Delta G^{\text{Destabilizing}}$  and (B)  $\Delta\Delta G^{\text{Affinity}}$  of mutations in the SARS-CoV-2 RBD-ACE2 complex.

prerequisite for an effective antibody design. For this purpose, physicochemical properties of wild and mutant type antibodies, including the nature of the interaction energy where water molecules by default play a key role, must be considered (Webster et al., 1994).

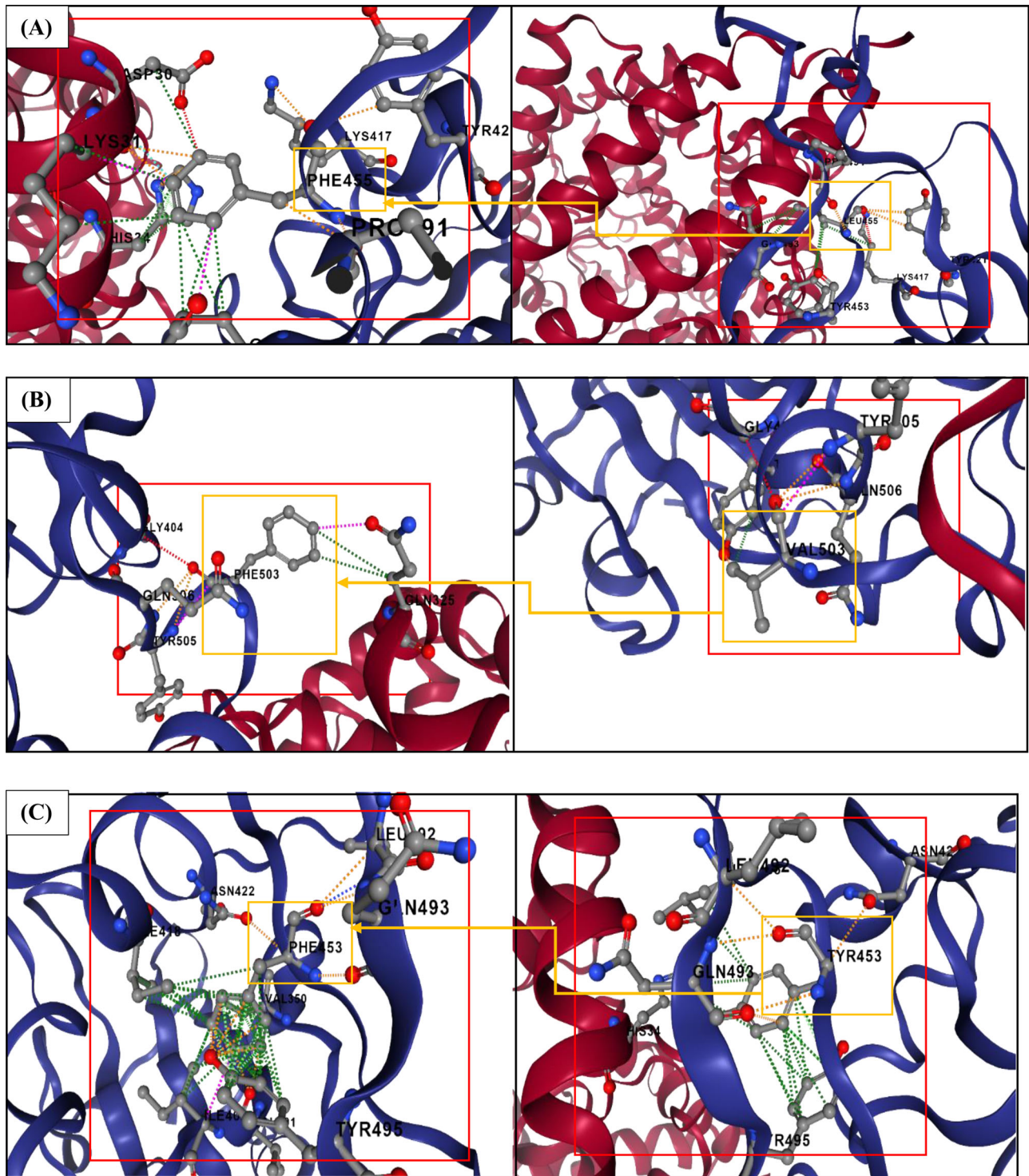
SFE ( $G_{\text{sol}}$ ) values calculated for the wild-type and the variants of SARS-CoV-2 RBD and SARS-CoV RBD are shown in Figure S2(A,B) and Table 3. By comparing these values, it can be noted that SARS-CoV RBD has a higher SFE and, thus, hydrophobicity than SARS-CoV-2 RBD. The higher charge of the SARS-CoV-2 RBD compared to that of SARS-CoV RBD can be one contributing factor to the lower SFE of the former protein domain (Table 3).

In addition to this crude comparison, three groups of mutations in the SARS-CoV-2 RBD were considered for a more selective study of the SFE. They included mutations converting (i) a charged amino acid to a neutral one (E484A, E484Q, K417N), (ii) a neutral amino acid to a charged one (N493K, G485R, S477R, T478K, G504D), and (iii) a non-hydrophobic amino acid to a hydrophobic one (Q493L, Y453F, S477I, S494L) (Table 4). Here, two mutations particularly stand out because they had a most dramatic effect on the SFE. They include K417N, which increased the SFE by nearly 50 kcal

$\text{mol}^{-1}$ , and S477R, which reduced it by over 30 kcal  $\text{mol}^{-1}$  (Table 3). With K417N substituting a neutral residue for a charged one and S477R substituting a charged residue for a neutral one, the key effect of charge on the SFE and the hydrophobicity, the latter of which is inversely proportional to charge, becomes obvious. Overall, the comparison of the wild-type SARS-CoV-2 RBD with the variants from the groups (i) and (ii), including S477R and K417N (Figure 7), shows the effects different mutations can have on charge, SFE, hydrophobicity, and electrostatic potential of the protein. What is common to all these single-point mutations, including S477R and K417N, is that they are most likely to affect the SFE by altering the folding and the stability of the protein structure. Needless to add, considering these changes made to the RBD due to mutations, their effect on the function and the properties of this domain can be harnessed for various therapeutic aims.

**3.2.1.3. Free energy surface.** The plots in Figure 8 provide a comparative view of the free surface energy landscape as a function of RMSD and  $R_g$  for the RBDs of SARS-CoV-2 and SARS-CoV. The general trend applying to both RBDs is that the higher the value of the two parameters (RMSD and  $R_g$ ),





**Figure 5.** Three mutations, namely (A) L445F, (B) V503F, and (C) Y453F, presented with a  $\Delta\Delta G$  value above zero. Orange boxes indicate mutant (left) and the wild type (right). The interface between RBD (blue) and ACE2 (red) is outlined by the red boxes (Interaction colors: Hydrophobic (Green), Hydrogen Bond (Red), Polar (Orange), van der Waals (Blue)).

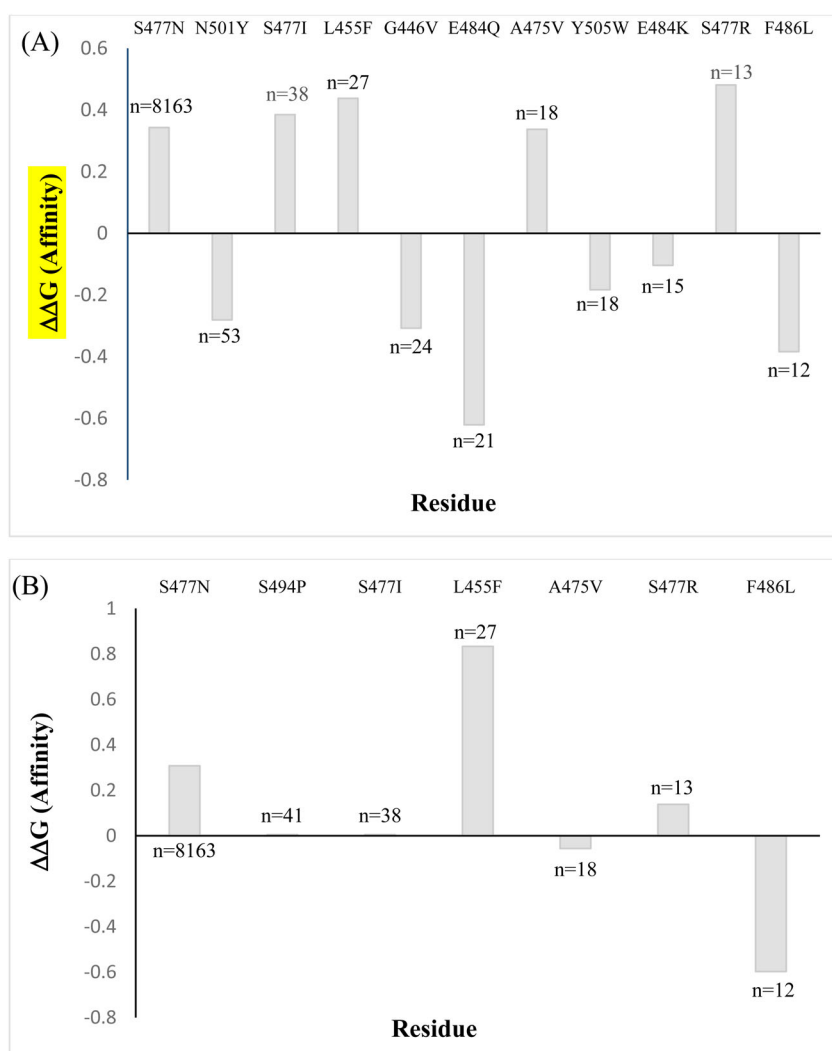
the lower the free energy surface. Also, given that the more stable structure possesses a lower free energy, the SARS-CoV RBD structure appears to be more stable than that of SARS-CoV-2 RBD (Figure 8(A,B)).

### 3.2.2. Effect of mutations on the electrostatic potential

Electrostatic interactions between protein residues and the solvent are one of the key determinants of protein folding

and stability (Strickler et al., 2006; Zhou & Pang, 2018). The electrostatics also regulates protein interactions with other proteins as well as with molecules other than proteins, such as small-molecule drugs (Voet et al., 2013). The electrostatic features of the protein are determined by the distribution of whole and partial charges across the 3D protein structure (Vascon et al., 2020). Here, Coulomb's law is not appropriate for describing electrostatic effects in proteins because it applies to a system with uniform dielectric properties, whilst





**Figure 6.**  $\Delta\Delta G^{\text{Affinity}}$  of mutations of (A) the RBD-COVA2-04 complex, and (B) the RBD-COVA2-39 complex.

**Table 3.** Comparison of the solvation free energy for the RBDs of SARS-CoV-2 and SARS-CoV.

RBD	Solvation free energy (kcal mol <sup>-1</sup> )	Charge
SARS-CoV	3082.49	4
SARS-CoV-2	2829.47	7

proteins dispersed in a medium display a gradient dielectric constant, with a hydrophobic core surrounded by the hydrophilic surface and covered by the solvent. Hence, electrostatic calculations for proteins were carried out using the Poisson-Boltzmann equation (Vascon et al., 2020).

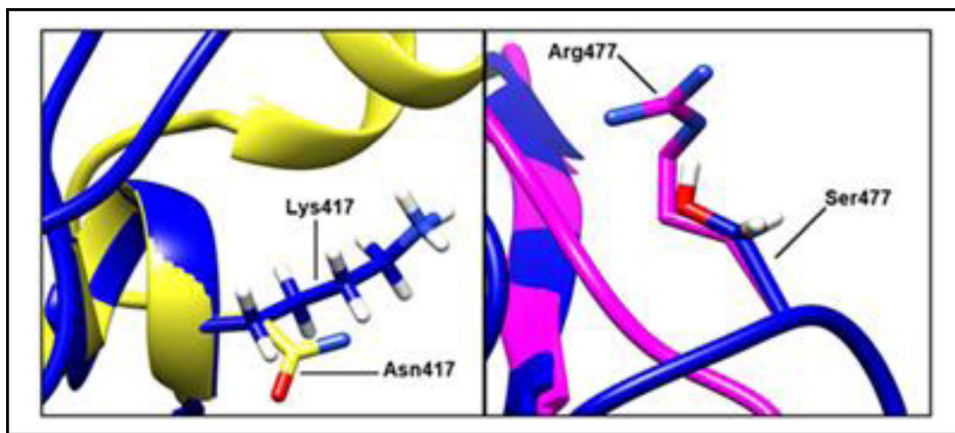
As in agreement with the earlier demonstrated effect of charged residues on the SFE, electrostatic interactions between the solvent (water or physiological media) and the protein were evidenced as determinants of the degree of hydrophobicity. The distribution of the electrostatic potential on the surfaces of the RBDs of SARS-CoV-2 and SARS-CoV was evaluated too. Figure 9(A,B) shows this distribution, with blue and red surfaces indicating electropositive and electro-negative surfaces, respectively. Investigation of the amino acid sequence of RBDs showed that the net charge for SARS-CoV-2 and SARS-CoV is 7 and 4, respectively. This difference

**Table 4.** Effect of selected mutations on the solvation free energy of the SARS-CoV-2 RBD.

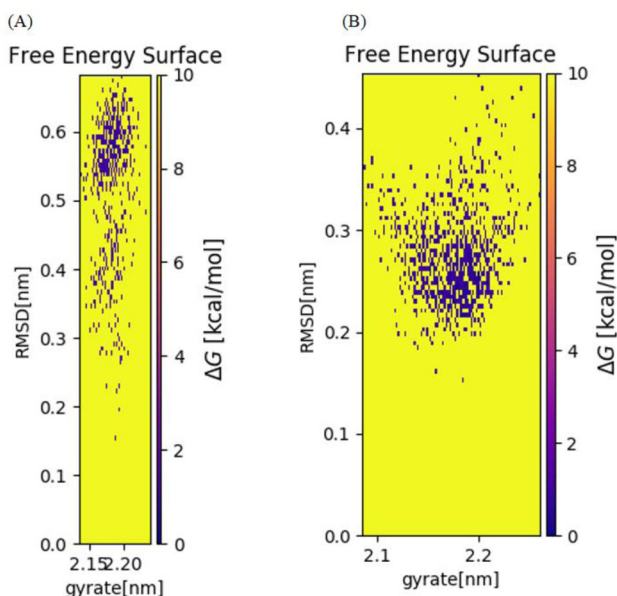
RBD of SARS-CoV-2	Solvation free energy ( $\Delta G_{\text{solv}}$ ) (kcal mol <sup>-1</sup> )	Charge
Wild-type	2829.46	7
K417N	2878.24	6
S477R	2698.92	8
Q493L	2847.17	7
Y505W	2835.44	7
Y453F	2835.91	7
S477I	2822.92	7
S494L	2854.42	7

in the net charge is explained by the larger number of charged amino acid residues in the RBD of SARS-CoV-2 than in the RBD of SARS-CoV. Charged residues that particularly stand out are K445 in the SARS-CoV-2 RBM and R426 and D463 in the SARS-CoV RBM. Since these residues are involved in binding to ACE2, they are important to examine in more detail.

With the electrostatic potential having an essential role in various biochemical reactions, such as the binding of a ligand to a specific protein target, electrostatic effects of mutations are important to consider (Figure 9(C-G)). Here we



**Figure 7.** Schematic representation of two mutations in SARS-CoV-2 RBD (blue, yellow, and pink indicate wild type, K417N, and S477R variants, respectively).



**Figure 8.** Free energy surface plots of (A) SARS-CoV-2 RBD and (B) SARS-CoV RBD as functions of RMSD and  $R_g$ .

consider five different mutations, namely N439K, G485R, S477R, G504D and T478K, and demonstrate that the distribution of charges on the RBD profoundly changes with each of these single-point mutations. This can be evidenced by noticing intense changes in the two colors representing positive and negative surface charges entailing each of the mutations and affecting most segments in the domain. Still, the distribution of charges is most drastically affected at the site of substitution and its immediate vicinity on the surface of the folded protein. In view of this, the effects of mutations must be manifold and may critically change the specific interactions with drugs, antibodies or ACE2. These structural changes allow the virus to potentially evade the therapy through timely mutation. These findings may be of benefit not only for delineating the most critical amino acids for the RBD function, but also for finding out how specific mutations may lead to disorder or improvement in the function of the RBD and its binding affinity through alterations in the electrostatic potential.

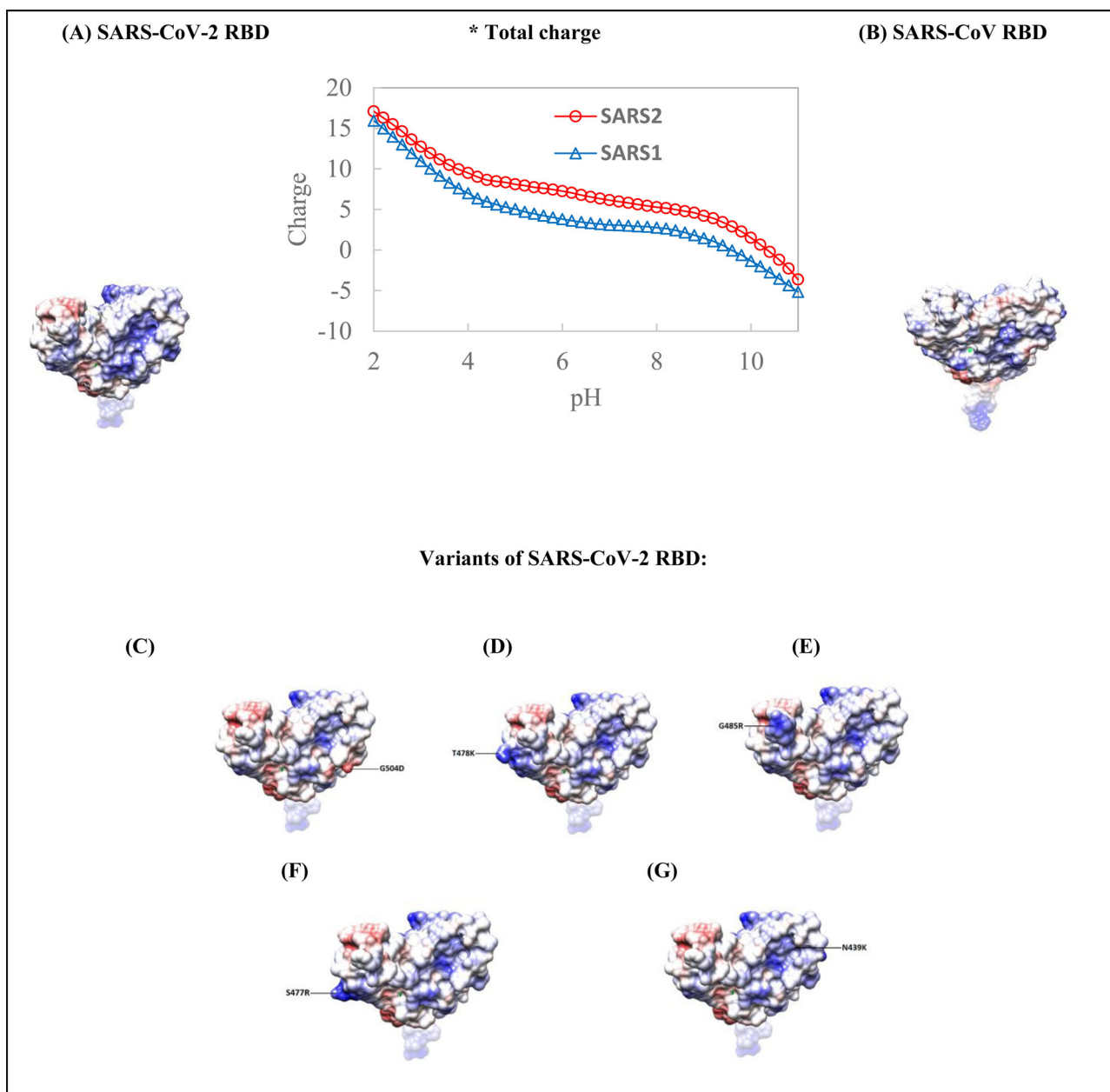
### 3.2.3. Effect of mutations on the antibody design

R426 is known to be one of the key residues for the binding of antibodies to the SARS-CoV RBD (Amin et al., 2020). This residue is positively charged under the physiological conditions, forming a salt bridge in interaction with receptors and antibodies. As a result, mutations at this site can have a large effect on protein-protein interactions. The R426G change, for example, causes the removal of this positively charged amino acid in the SARS-CoV-2 RBD, meaning that no salt bridge can be created by the RBD at this position. The significant decrease in the binding affinity of SARS-CoV-2 RBD to antibodies may be due to the lack of this salt bridge formation.

The N439K mutation swapping an uncharged group for a positively charged one may elicit the opposite effect, possibly inducing the formation of a salt bridge at a site that previously did not form them. Here, it is important to add that N439K is the mutation with the second highest frequency of occurrence (Figure 3(A)). The effects of this mutation may be such that it induces the creation of a salt bridge as in the SARS-CoV RBD, thus increasing the strength of the binding of the RBD to the receptors or antibodies. This increased affinity, of course, is a double-edged sword because it may improve the interaction with drugs and antibodies, thus increasing their efficacy, but it also strengthens the interaction of the vital particle with the host cell.

### 3.3. Molecular dynamics study

MD simulations were performed for the SARS-COV-2 RBD, the SARS-CoV RBD, and the SARS-CoV-2 RBD-ACE2 complex in order to understand the difference in the structure of the two RBDs, alongside the implications for the stability of the SARS-CoV-2 RBD induced by these changes. To determine the stability and convergence of structure, RMSD and  $R_g$  were calculated (Sargsyan et al., 2017). RMSDs for the SARS-CoV-2 RBD and the SARS-CoV RBD were found to be  $\sim 0.5$  and  $\sim 0.36$  nm, respectively. The RMSD plot indicates that both RBDs reached equilibrium after 40 ns and that the SARS-CoV RBD is more stable than the SARS-CoV-2 RBD (Figure 10(A)). The RMSD of the RBD-ACE2 complex during 10 ns of the MD simulation time displayed the lowest RMSD



**Figure 9.** Main: Electrostatic properties of the RBDs of SARS-CoV-2 (A), SARS-CoV (B) and selected mutated SARS-CoV-2 RBD variants: (C) G504D, (D) T478K, (E) G485R, (F) S477R, and (G) N439K. Blue and red represent positive and negative charges, respectively. Inset: Net charge for the RBDs of SARS-CoV-2 and SARS-CoV as a function of pH.

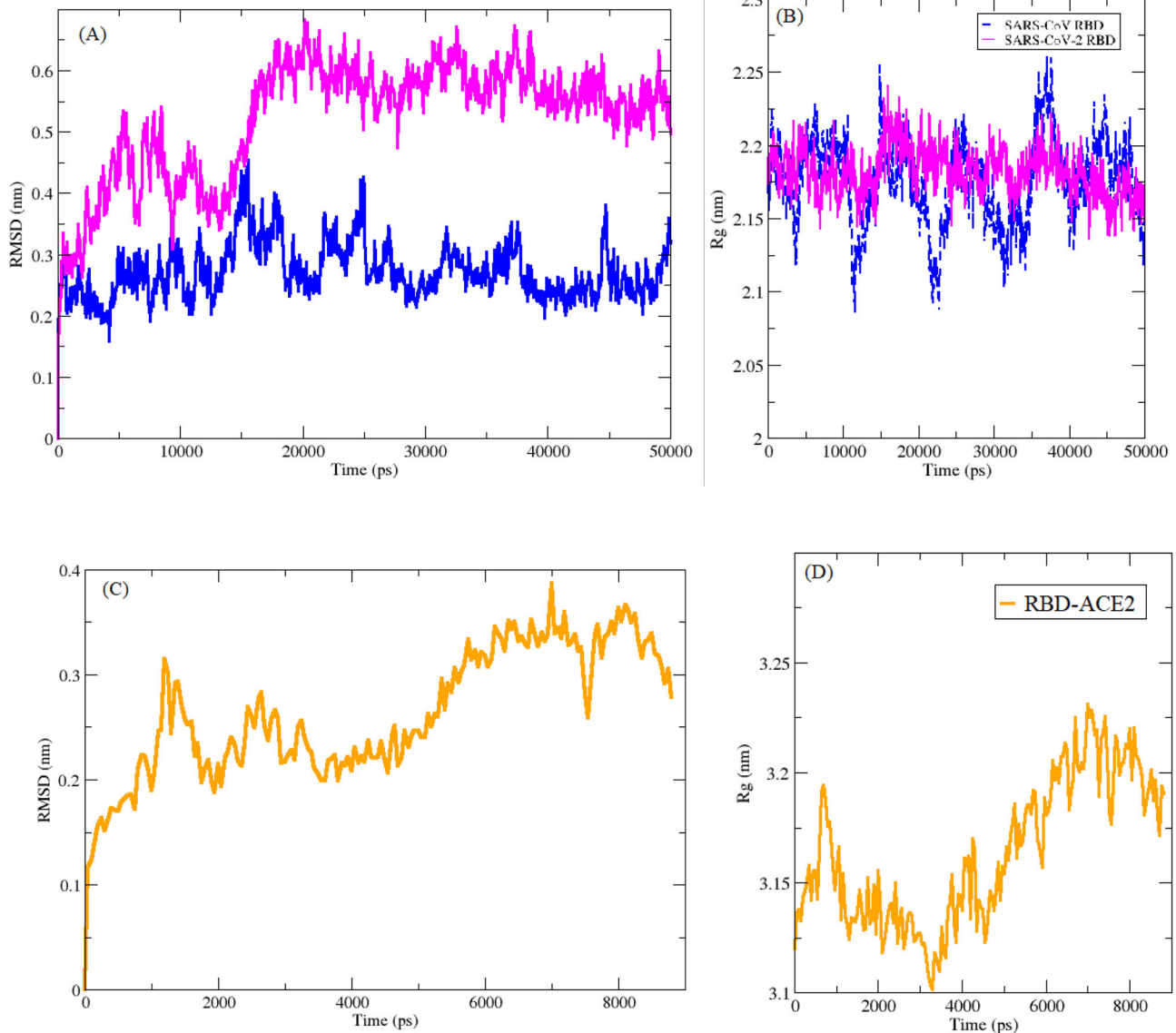
( $\sim 0.25$  nm) (Figure 10(C)), which can be due to the presence of the stabilizing interactions in the protein-protein complex.

$R_g$  was calculated from the 50 ns trajectory to gain insight on the compactness and rigidity of the RBDs. The largest deviation was observed in the SARS-CoV-2 RBD, which can indicate functional movements and structural 'breathing' due to the lesser compactness of the structure (Figure 10(B)).  $R_g$  was also calculated for the SARS-CoV-2 RBD-ACE2 complex (Figure 10(D)), showing deviation within  $\sim 4.4$  ns ( $\sim 3.17$  to  $\sim 3.12$  Å). In addition, the residue-based root mean square deviation (RMSF) plot was constructed from the 50 ns data in order to understand the deviation of each RBD. It was shown that the RMSF of the backbone for the SARS-CoV-2 RBD displays more flexible residues at residue numbers 477 to 487 as compared to the SARS-CoV RBD (Figure 11). This

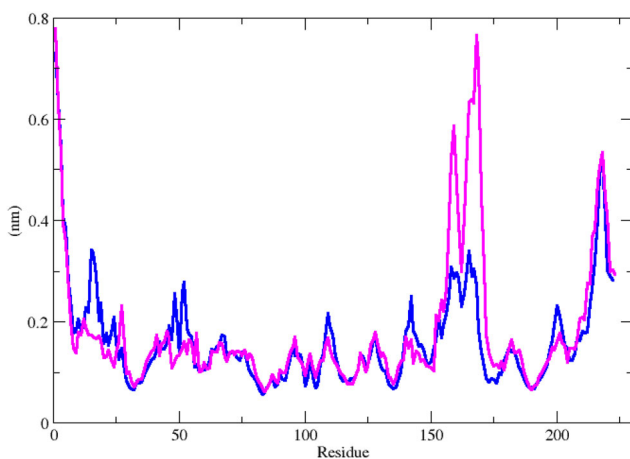
conformational flexibility of the binding domain may be one of the key factors in ensuring the high infectivity of SARS-CoV-2, as in analogy with other active sites in proteins (Obadan et al., 2019; Penfold et al., 2004).

In order to analyze H-bond interactions during the 50 ns and 10 ns simulation times, intramolecular H-bond plots were constructed for the RBDs of SARS-CoV-2 and SARS-CoV (Figure 12(A)). From the comparative plots, it could be deduced that the SARS-CoV-2 RBD has a similar number of intramolecular H-bonds as the SARS-CoV RBD (Figure 12(A)). In addition, Figure 12(C) indicates the number of H-bond in the SARS-CoV-2 RBD-ACE2 protein complex. The high number of H-bonds here indicates the relatively high stability of the interaction between the RBD and its receptor.





**Figure 10.** (A) RMSD of the RBDs. (B)  $R_g$  of the SARS-CoV-2 RBD and SARS-CoV RBD. (C) RMSD of SARS-CoV-2 RBD-ACE2. (D)  $R_g$  of SARS-CoV-2 RBD-ACE2.



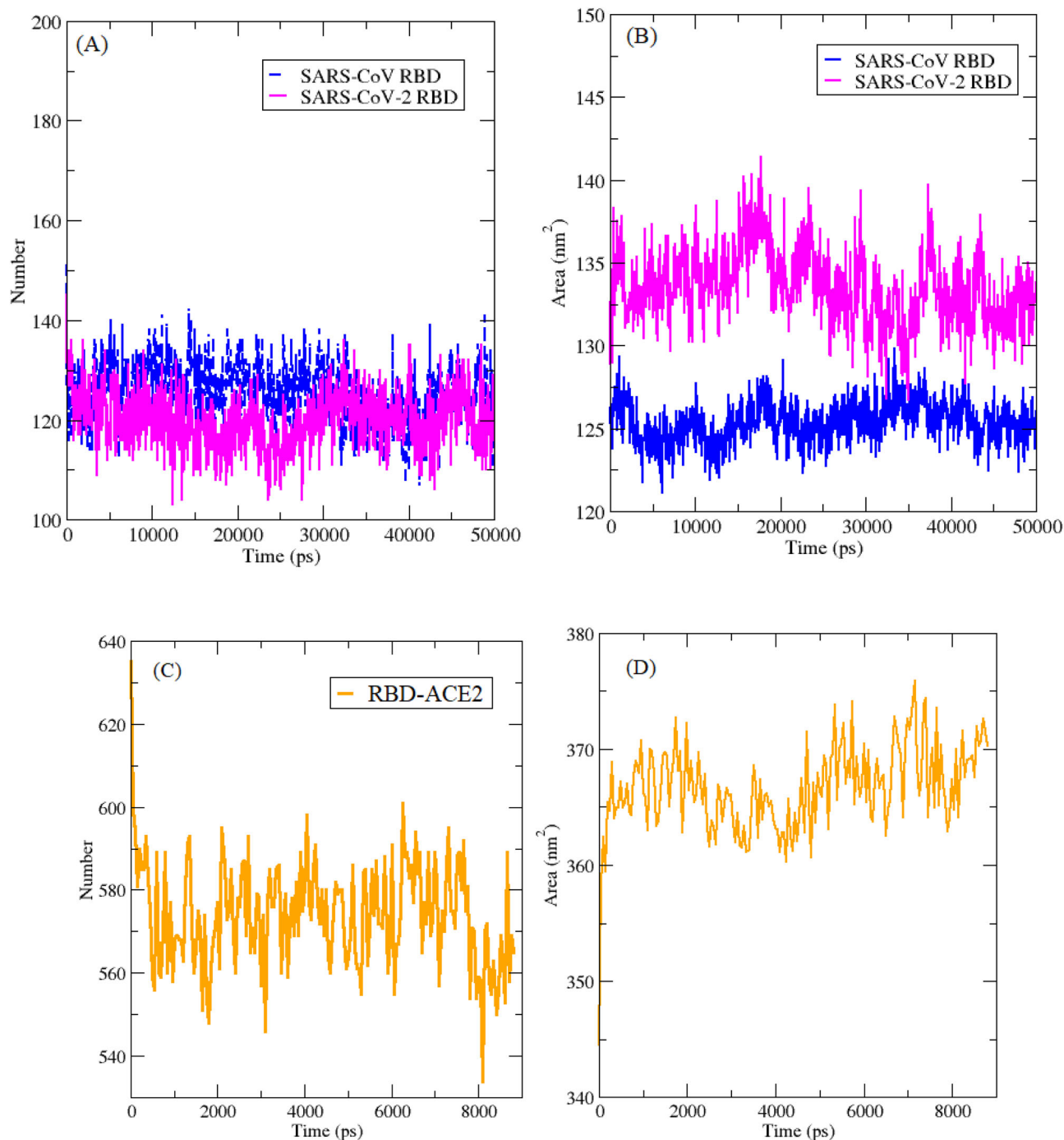
**Figure 11.** RMSF of the SARS-CoV-2 and SARS-CoV RBD. Color pattern: SARS-CoV-2 (Magenta) and SARS-CoV (Blue).

The solvent-accessible surface area (SASA) as the surface region of the protein that is accessible to the solvent was calculated for the RBDs of SARS-CoV-2 and SARS-CoV and for the

SARS-CoV-2 RBD-ACE2 complex as a function of time and the resulting plots are shown in Figure 12(B,D). The results obviously indicate that the SARS-CoV-2 RBD has a higher SASA ( $\sim 134 \text{ nm}^2$ ) than the SARS-CoV RBD ( $\sim 126 \text{ nm}^2$ ). In addition, based on these plots, the SASA for the SARS-CoV-2 was  $\sim 370 \text{ nm}^2$ . With SASA being a measure of the degree to which an amino acid is exposed to the environment, the higher SASA for the RBD of SARS-CoV-2 indicates its more open and diffused structure than that of the more compact RBD of SARS-CoV. Any increase or decrease in the SASA in response to structural changes indicates alterations in the conformation of the protein.

### 3.4. Principal component analysis

PCA is a method for the analysis of the MD trajectory that extracts dominant modes in the overall molecular motion. Here, it was used to determine the dominant modes of molecular flexibility and functional motions in the RBDs. To perform the PCA, eigenvectors were obtained by diagonalization of the covariance matrix



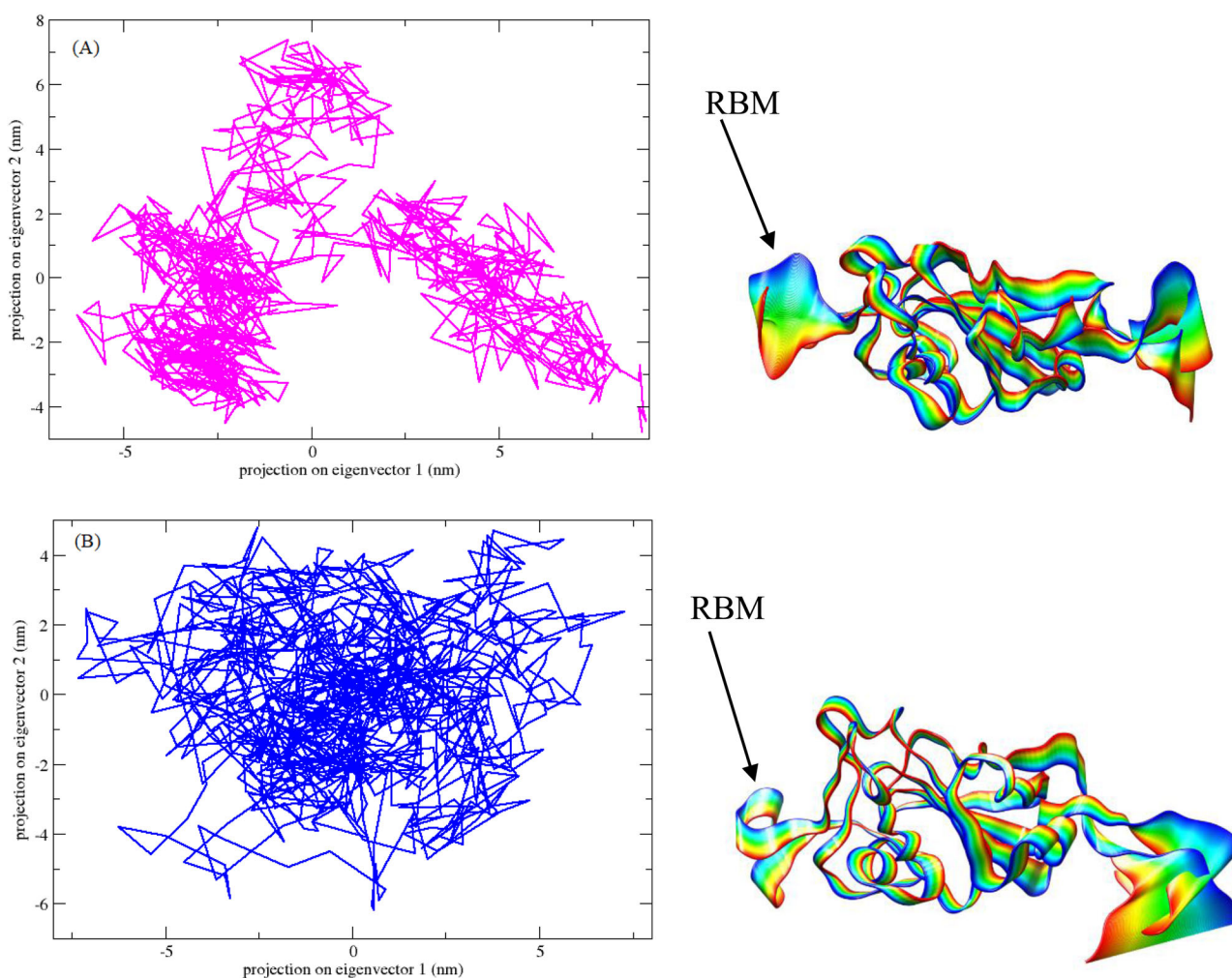
**Figure 12.** (A) Number of intramolecular H-bonds of SARS-CoV-2 and SARS-CoV RBDs during the 50 ns MD simulation time. (B) SASA analysis of SARS-CoV-2 and SARS-CoV RBDs. (C) Total number of intramolecular H-bonds formed between the SARS-CoV-2 RBD and ACE2 over the simulation time of 10 ns. (D) SASA analysis of the SARS-CoV-2 RBD-ACE2 complex.

of the coordinate fluctuations of the RBDs of SARS-CoV-2 and SARS-CoV. To understand the total motion of these structures in the phase space, the first two eigenvectors (1 and 2) were projected onto it. The comparison has shown that the motion properties described by these first two eigenvectors are different in the SARS-CoV-2 RBD as compared to the SARS-CoV RBD (Figure 13).

### 3.5. Analysis of residue interaction networks

The RIN corresponding to the receptor-bound form of the SARS-CoV-2 RBD is presented in Figure 14. The greater

number of clusters and the correspondingly higher MCODE score was obtained for the SARS-CoV-2 RBD complex with ACE2 than for the SARS-CoV-2 RBD without the receptor. The greatest overlap in cluster members in both the free and the complex form of the RBD was noted for the residues I358, C361, C336, F456, and F515. Among these residues, F456 appears particularly significant because the RIN in Figure 14 indicates its direct role in mediating the interaction between the RBD and ACE2. In addition, this residue was identified as one of the mutations (F456L) in the RBD of SARS-CoV-2 (Figure 3).



**Figure 13.** The principal component analysis. Projection of the most significant principal components of motion of the (A) SARS-CoV-2 RBD and (B) SARS-CoV RBD. The trajectory is projected onto the two dimensional space defined by eigenvectors 1 and 2.

Critical residues at the interface between the SARS-CoV-2 RBD and ACE2 in their complex are indicated in the RIN presented in Figure 14. Also, Table 5 shows the seemingly most important residues of the SARS-CoV-2 involved in the interaction between its RBD and ACE2 via H-bonds, van der Waals bonds, ionic and  $\pi$ - $\pi$  stacking interactions. Among these residues, K417 ( $\rightarrow$ 417N), Y453 ( $\rightarrow$ 453F), F486 ( $\rightarrow$ 486L), A475 ( $\rightarrow$ 475V), and F456 ( $\rightarrow$ 456L) are of utmost importance because their mutations could alter the ability of the RBD to bind to ACE2, which could lead to the loss of efficacy of antibodies/drugs targeting the RBD or competing with it for ACE2.

### 3.5.1. Key residues in the SARS-CoV 2 RBD with and without ACE2

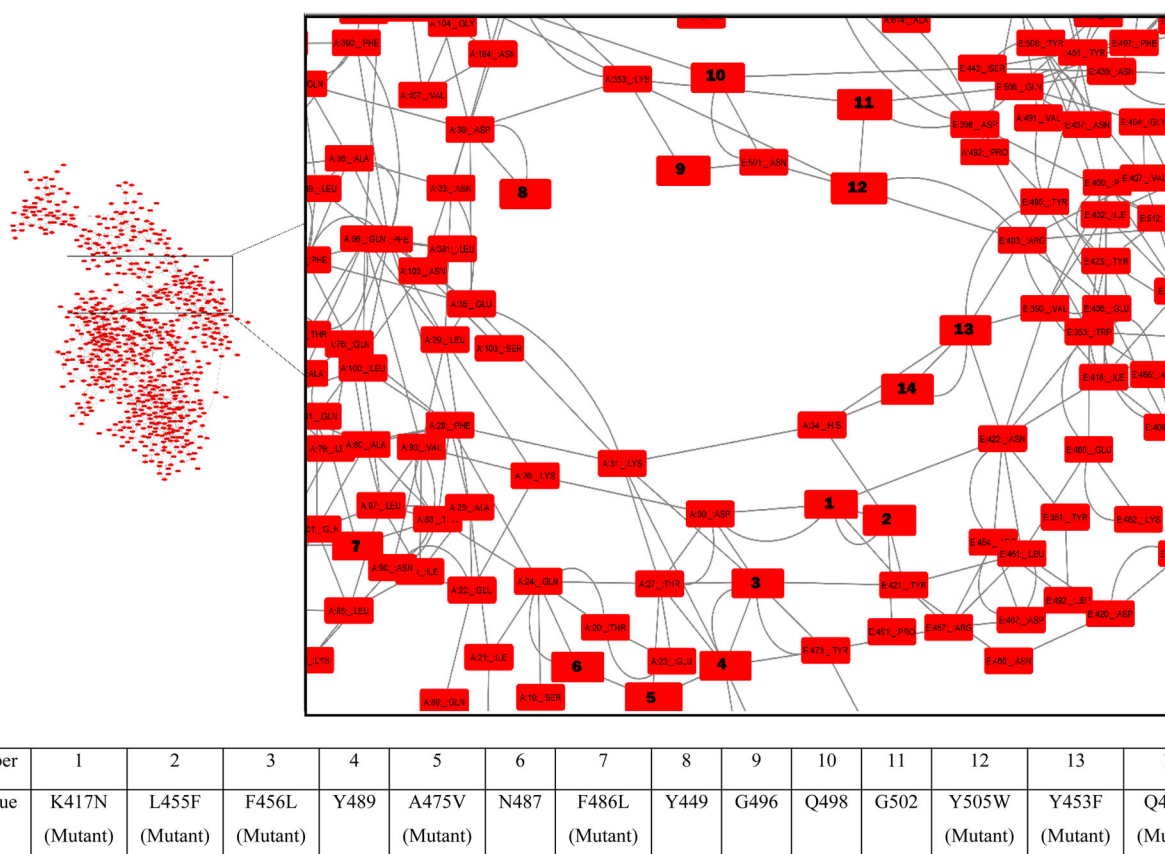
The betweenness analysis identified S349, A352, W436, and L452 as residues of interest in the SARS-CoV-2 RBD without ACE2 (Figure 15(A)). The stress metric, in contrast, identified N422, W436, Y486, W353, F347, F400, and L461 as residues of interest in the SARS-CoV-2 RBD without ACE2 (Figure 15(C)). As far as the SARS-CoV-2 RBD with ACE2 is concerned, the betweenness analysis identified L452, S349, and A352 as residues of interest (Figure 15(B)), while the stress metric

identified residues Y453, W436, Y505, R403, Y508, F342, and F338 as those of interest (Figure 15(D)). Different residues emerging as the results of the stress and betweenness analyses applied to the free RBD and the RBD-ACE2 complex may be indicative of significant structural alterations occurring upon the binding of the RBD to the receptor. Interestingly, by appearing in 3 out of 4 of these distinct analyses, W436 appears to be a key residue in both the free and the receptor-bound structures of the SARS-CoV-2 RBD and may play an essential role in ensuring both the stability and the binding of the RBD to its receptor.

### 3.6. Calculation of the dipole moment

Dipole moment can play an important role in the mechanism of interaction of proteins with receptor and drugs (Lien et al., 1982). The computed dipole moments of the SARS-CoV-2 RBD, the SARS-CoV RBD, and the SARS-CoV-2 RBD-ACE2 were 720 Debye (D), 204 D, and 760 D, respectively. As shown in Figure 16, it appears that the RBDs of SARS-CoV-2 and SARS-CoV, albeit having similar functions, have quite distinct dipole moments. This difference in the dipole moment reiterates the different charge distribution on the surface of the two RBDs and is expected in view of the fact that the





**Figure 14.** 3D and selected 2D region of the RIN descriptive of the interface between the SARS-CoV-2 RBD and ACE2.

**Table 5.** Molecular forces governing the interaction between different residues of the SARS-CoV-2 RBD and ACE2.

Residues of RBD	Residues of ACE2	Interaction
Asn487	Gln24	Hydrogen bond
Lys417	Asp30	Hydrogen bond
Tyr449	Asp38	Hydrogen bond
Gly496	Lys353	Hydrogen bond
Gly502	Lys353	Hydrogen bond
Lys417	Asp30	Ionic interaction
Tyr453	His34	$\pi$ - $\pi$ Stacking
Phe486	Tyr83	$\pi$ - $\pi$ Stacking
Asn487	Gln24	VDW
Phe456, Ala475, Tyr489	Thr27	VDW
Phe456	Asp30, Lys31	VDW
Leu455, Gln493	His34	VDW
Phe486	Leu79, Met82, Tyr83	VDW
Gln498	Tyr41	VDW
Tyr449	Asp38	VDW
Tyr505	Lys353	VDW

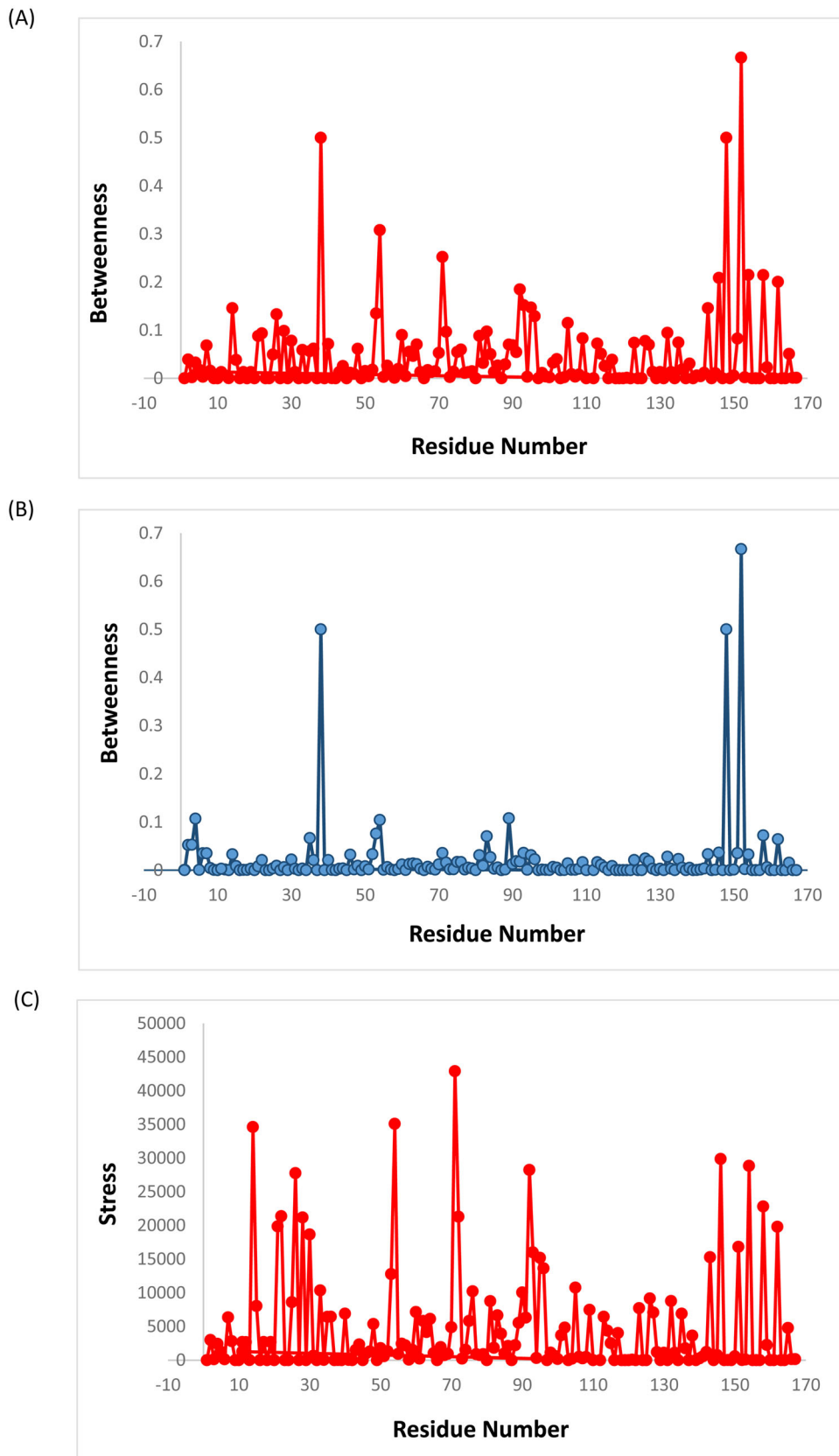
distribution of charges in the protein defined the dipole moment. Further, the dipole moment of the SARS-CoV-2 RBD-ACE2 complex is more inconstant than the relatively stable, albeit naturally fluctuant dipole moments of the RBDs of SARS-CoV-2 and SARS-CoV alone. Overall, the dipole moment values are significant because large dipole moments may promote binding and also control other aspects of the protein function (Felder et al., 2007). Large dipole moments lead to stabilization of the protein-protein complexes and are also of critical importance for achieving the correct orientation prior to the formation of the given complexes (Pichierri, 2013). For this reason, many studies consider the dipole moment as a key parameter governing the drug-receptor

interaction (Lien et al., 1982). Here, it should be kept in mind that this large of the dipole moment value was calculated for the RBD only and may differ from the dipole moment of the spike protein as a whole.

#### 4. Discussion

This computational study consisted of the analysis of the major and the minor structural factors of the SARS-CoV-2 RBD playing a role in determining the virulent function of the SARS-CoV-2 spike protein. The protein dynamics (Hollingsworth & Dror, 2018), electrostatics (Zhou & Pang, 2018) and residue networking (Gordon et al., 2020) were examined to elucidate this relationship between the protein structure and function. All-atom molecular dynamics (MD) simulations were carried out in the course of 50 ns at 310 K and a number of physiochemical parameters were subjected to analysis. A special emphasis was put on the assessment of the impact of amino acid mutations on the structure and thermodynamic properties of 42 variants of the SARS-CoV-2 RBD discovered so far. It goes beyond the scope of this paper to comment on the continuously evolving interrelation of different structural parameters investigated here, but the major findings pertaining to individual parameters could be compiled.

First of all, the binding affinity of SARS-CoV-2 RBD for ACE2 is shown to be higher than that of SARS-CoV RBD. Next, the binding of COVA2-04 antibody to SARS-CoV-2 RBD is more energetically favorable than the binding of COVA2-



**Figure 15.** Betweenness (a and c) and stress (b and d) centralities for the SARS-CoV-2 RBD structure without ACE2 (A and C) and with ACE2 (B and D).

39, but also less favorable than the formation of the SARS-CoV-2 RBD-ACE2 complex, hinting at a possible thermodynamic advantage that the virus has over the immune system. The net charge, the dipole moment and hydrophilicity

of SARS-CoV-2 RBD are higher than those of SARS-CoV, producing lower solvation and surface free energies and thus lower stability. This lower stability of SARS-CoV-2 was confirmed in the MD simulations of RMSD, RMSF, radius of

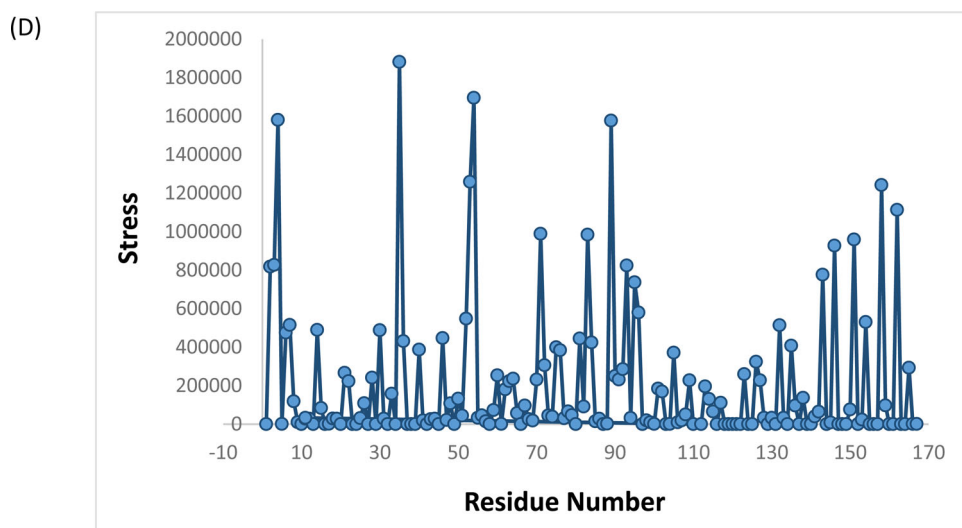


Figure 15. Continued.

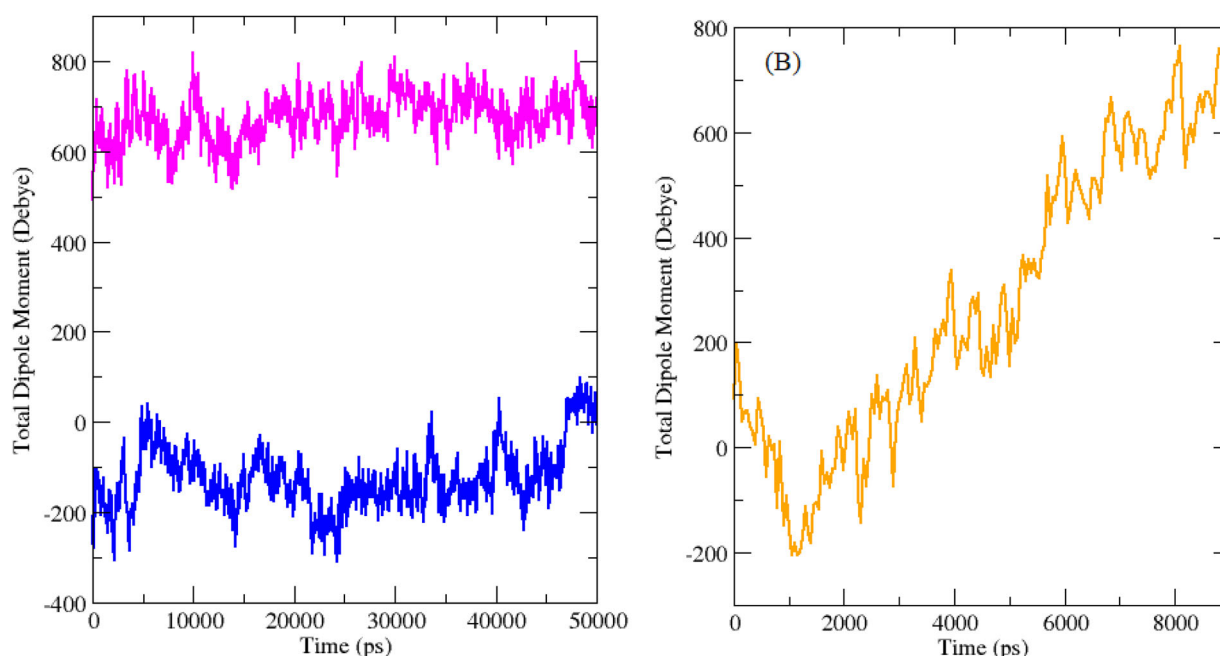


Figure 16. Total dipole moment of the RBDs of (A) SARS-CoV-2 (Magenta) and SARS-CoV (Blue). (B) Total dipole moment of the RBD-ACE2 complex (Orange).

gyration and H-bond distribution. It is paralleled by the higher flexibility of SARS-CoV-2 RBD compared to more rigid SARS-CoV RBD, owing in part to two less  $\beta$  sheets in the RBD of SARS-CoV-2 than in that of SARS-CoV. The structure of SARS-CoV-2 RBD is correspondingly more open, with the solvent-accessible surface area larger by  $8 \text{ nm}^2$  than that of SARS-CoV RBD. It also undergoes intense structural alterations upon binding to ACE2. Mutations have a dramatic effect on distribution of charges, most prominently at the site of substitution and its immediate vicinity. Charge alterations upon single-point mutation have a prominent effect on the free energy landscape, increasing the solvation free energy by  $\sim 50 \text{ kcal mol}^{-1}$  for a charge-promoting substitution such as K417N and reducing it by over  $30 \text{ kcal mol}^{-1}$  for a charge-neutralizing substitution such as S477R. Meanwhile, all X $\rightarrow$ F mutations exhibit a stabilizing effect on the RBD structure through  $\pi$  stacking. Considering the effect of

different mutations on the thermodynamics of SARS-CoV-2 RBD, some of them have imposed considerable changes to  $\Delta\Delta G^{\text{Destabilizing}}$  and  $\Delta\Delta G^{\text{Affinity}}$ . These changes proceed in such a way that with an increasing variety of mutations, the value of  $\Delta\Delta G^{\text{Affinity}}$  has tended to increase and the value of  $\Delta\Delta G^{\text{Destabilizing}}$  to decrease. Finally, F456 and W436 emerged from the RIN analysis of betweenness and stress as two key residues governing the stability and affinity of the spike protein for its ACE2 receptor.

Notwithstanding that the effect that varying physiological conditions and the presence of different drugs can have on the mechanism of interaction of the SARS-CoV-2 RBD with ACE2, the network analysis can provide a valuable tool for predicting changes to the interaction potential and the rate of interaction between the RBD and its receptor in response to structural changes, such as mutations. It can also help identify key mutations that may occur in the future and thus



facilitate the antiviral drug discovery from the fundamental standpoint of protein-protein interaction between the virus and the host (Ackerman et al., 2018; Brito & Pinney, 2017). This research route is worth examining in the future when the number of mutants reaches a statistically acceptable level, which would allow the determination of the direction in which the RBD mutations proceed with satisfactory levels of confidence.

## 5. Conclusion

SARS-CoV-2 has gripped the world in an ongoing pandemic and timely investigations of various aspects of its nature are of paramount importance. MD simulations were performed to structurally compare the RBDs on the spike proteins of SARS-CoV-2 and SARS-CoV, with and without binding to ACE2 and also in response to different mutations. Given the significant recent work reported on SARS-CoV-2, an attempt was made to examine parameters that were considered to a lesser extent in previous studies. The physicochemical properties of the RBD analyzed included the free energies of binding, destabilizing and solvation, the electrostatic potential and hydrophobicity; the effects of mutations on each of these properties was assessed. Some of the most critical mutations were highlighted, such as K417 and N439, and their effect on the physicochemical properties of the RBD discussed, along with the corollaries for the design of antibodies and drugs, most of which utilize the given RBD as the target. The study suggests that changes induced by these mutations may play a role in reducing or increasing the number of protein interactions and protein surface charges, consequently affecting the stability and function of the RBD and thus the infectivity of SARS-CoV-2. Therefore, the structural details and physicochemical parameters of the RBD reported here may help in the development of antibodies or drugs at the molecular level. Also, previous studies indicated that many disease-associated amino acid mutations are located at the protein-protein interface, for which reason we have also focused on understanding the structural changes occurring at the interface between SARS-CoV-2 and ACE2. Mutations are also known to affect protein-protein interaction by changing the binding affinity, which was analyzed in this study in detail. Information on the impact of mutations on protein-protein interaction may aid in classifying mutations into pathogenic and harmless. From there on, this information could prove essential for the accelerated development of appropriate therapeutic and drug design strategies. This study suggests that next to investigating the SARS-CoV-2 RBD structure, researchers dealing with drugs and antibodies field could benefit from the fast and precise computational tools to predict the binding affinity and its changes in response to mutation.

## Acknowledgement

The support and resources from the Center for High Performance Computing at Shahid Beheshti University of Iran are gratefully acknowledged.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Yahya Sefidbakht  <http://orcid.org/0000-0003-0538-0252>  
Vuk Uskoković  <http://orcid.org/0000-0003-3256-1606>

## References

- Ackerman, E. E., Kawakami, E., Katoh, M., Watanabe, T., Watanabe, S., Tomita, Y., Lopes, T. J., Matsuoka, Y., Kitano, H., Shoemaker, J. E., & Kawaoka, Y. (2018). Network-guided discovery of influenza virus replication host factors. *mBio*, 9(6), e02002-18.
- Allouche, A. (2011). Software news and updates gabedit—A graphical user interface for computational chemistry softwares. *Journal of Computational Chemistry*, 32(1), 174–182. <https://doi.org/10.1002/jcc> <https://doi.org/10.1002/jcc.21600>
- Amin, M., Sorour, M. K., & Kasry, A. (2020). Comparing the binding interactions in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *The Journal of Physical Chemistry Letters*, 11(12), 4897–4900. <https://doi.org/10.1021/acs.jpcllett.0c01064>
- Anandkrishnan, R., Aguilar, B., & Onufriev, A. V. (2012). H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research*, 40(W1), W537–W541. <https://doi.org/10.1093/nar/gks375>
- Bader, G. D., & Hogue, C. W. V. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, 4, 2–27. <https://doi.org/10.1186/1471-2105-4-2>
- Bruto, A. F., & Pinney, J. W. (2017). Protein-protein interactions in virus-host systems. *Frontiers in Microbiology*, 8, 1557–1511. <https://doi.org/10.3389/fmicb.2017.01557>
- Brouwer, P. J. M., Caniels, T. G., van der Straten, K., Snitselaar, J. L., Aldon, Y., Bangaru, S., Torres, J. L., Okba, N. M. A., Claireaux, M., Kerster, G., Bentlage, A. E. H., van Haaren, M. M., Guerra, D., Burger, J. A., Schermer, E. E., Verheul, K. D., van der Velde, N., van der Kooij, A., van Schooten, J., ... van Gils, M. J. (2020). Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science (New York, N.Y.)*, 369(6504), 643–650. <https://doi.org/10.1126/science.abc5902>
- Chong, S. H., & Ham, S. (2014). Interaction with the surrounding water plays a key role in determining the aggregation propensity of proteins. *Angewandte Chemie (International ed. in English)*, 53(15), 3961–3964. <https://doi.org/10.1002/anie.201309317>
- Chong, S. H., Lee, C., Kang, G., Park, M., & Ham, S. (2011). Structural and thermodynamic investigations on the aggregation and folding of acylphosphatase by molecular dynamics simulations and solvation free energy analysis. *Journal of the American Chemical Society*, 133(18), 7075–7083. <https://doi.org/10.1021/ja1116233>
- Dehury, B., Raina, V., Misra, N., & Suar, M. (2020). Effect of mutation on structure, function and dynamics of receptor binding domain of human SARS-CoV-2 with host cell receptor ACE2: A molecular dynamics simulations study. *Journal of Biomolecular Structure and Dynamics*. <https://doi.org/10.1080/07391102.2020.1802348>
- Felder, C. E., Prilusky, J., Silman, I., & Sussman, J. L. (2007). A server and database for dipole moments of proteins. *Nucleic Acids Research*, 35(Web Server issue), W512–W521. <https://doi.org/10.1093/nar/gkm307>
- Fonseca, F. V., Júnior, R. O. S., Almeida, M. V. A., De, Thiago, D., Morais, D. A. A., Dalmolin, R. J. S., & Lima, J. P. M. S. (2020). CoRINs: A tool to compare residue interaction networks from homologous proteins and conformers. *BioRxiv*. <https://doi.org/10.1101/2020.06.29.178541>
- Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., ... Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug

- repurposing. *Nature*, 583(7816), 459–468. <https://doi.org/10.1038/s41586-020-2286-9>
- Gui, M., Song, W., Zhou, H., Xu, J., Chen, S., Xiang, Y., & Wang, X. (2017). Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Research*, 27(1), 119–129. <https://doi.org/10.1038/cr.2016.152>
- Hollingsworth, S. A., & Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron*, 99(6), 1129–1143. <https://doi.org/10.1016/j.neuron.2018.08.011>
- Hu, G., Yan, W., Zhou, J., & Shen, B. (2014). Residue interaction network analysis of Dronpa and a DNA clamp. *Journal of Theoretical Biology*, 348, 55–64. <https://doi.org/10.1016/j.jtbi.2014.01.023>
- Huang, Y., Yang, C., Xu, X. F., Xu, W., & Liu, S. W. (2020). Structural and functional properties of SARS-CoV-2 spike protein: Potential antiviral drug development for COVID-19. *Acta Pharmacologica Sinica*, 41, 1141–1149. <https://doi.org/10.1038/s41401-020-0485-4>
- Jiang, S., Hillyer, C., & Du, L. (2020). Neutralizing antibodies against SARS-CoV-2 and other human coronaviruses. *Trends in Immunology*, 41(5), 355–359. <https://doi.org/10.1016/j.it.2020.03.007>
- Jorgensen, W. L., & Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(4), 1657–1666.
- Jubb, H. C., Pandurangan, A. P., Turner, M. A., Ochoa-Montano, B., Blundell, T. L., & Ascher, D. B. (2017). Mutations at protein-protein interfaces: Small changes over big surfaces have large impacts on human health. *Progress in Biophysics and Molecular Biology*, 128, 3–13. <https://doi.org/10.1016/j.pbiomolbio.2016.10.002>
- Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., ... Baker, N. A. (2018). Improvements to the APBS biomolecular solvation software suite. *Protein Science*, 27(1), 112–128. <https://doi.org/10.1002/pro.3280>
- Klingen, T. R., Reimering, S., Guzmán, C. A., & McHardy, A. C. (2018). In silico vaccine strain prediction for human influenza viruses. *Trends in Microbiology*, 26(2), 119–131. <https://doi.org/10.1016/j.tim.2017.09.001>
- König, G., Bruckner, S., & Boresch, S. (2013). Absolute hydration free energies of blocked amino acids: Implications for protein solvation and stability. *Biophysical Journal*, 104(2), 453–462. <https://doi.org/10.1016/j.bpj.2012.12.008>
- Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., & Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, 581(7807), 215–220. <https://doi.org/10.1038/s41586-020-2180-5>
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291. <https://doi.org/10.1107/S0021889892009944>
- Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, 8(4), 477–486. <https://doi.org/10.1007/BF00228148>
- Li, F., Li, W., Farzan, M., & Harrison, S. C. (2005). Structural biology: Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science (New York, N.Y.)*, 309(5742), 1864–1868. <https://doi.org/10.1126/science.1116480>
- Lien, E. J., Guo, Z.-R., Li, R.-L., & Su, C.-T. (1982). Use of dipole moment as a parameter in drug-receptor interaction and quantitative structure-activity relationship studies. *Journal of Pharmaceutical Sciences*, 71(6), 641–655. <https://doi.org/10.1002/jps.2600710611>
- Liu, Z., Xiao, X., Wei, X., Li, J., Yang, J., Tan, H., Zhu, J., Zhang, Q., Wu, J., & Liu, L. (2020). Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *Journal of Medical Virology*, 92(6), 595–601. <https://doi.org/10.1002/jmv.25726>
- Long, S. W., Olsen, R. J., Christensen, P. A., Bernard, D. W., Davis, J. J., Shukla, M., Nguyen, M., Saavedra, M. O., Yerramilli, P., Pruiett, L., Subedi, S., Kuo, H., Hendrickson, H., Eskandari, G., Nguyen, H. A. T., Long, J. H., Kumaraswami, M., Goike, J., Boutz, D., ... Musser, J. M. (2020). Molecular architecture of early dissemination and massive second wave of the SARS-CoV-2 virus in a major metropolitan. *mbio*, 11(6), e02707-20.
- Lukassen, S., Chua, R. L., Trefzer, T., Kahn, N. C., Schneider, M. A., Muley, T., Winter, H., Meister, M., Veith, C., Boots, A. W., Hennig, B. P., Kreuter, M., Conrad, C., & Eils, R. (2020). SARS-CoV-2 receptor ACE 2 and TMPRSS 2 are primarily expressed in bronchial transient secretory cells. *The EMBO Journal*, 39(10), 1–15. <https://doi.org/10.15252/embj.20105114>
- Martin, A. J. M., Vidotto, M., Boscaroli, F., Di Domenico, T., Walsh, I., & Tosatto, S. C. E. (2011). RING: Networking interacting residues, evolutionary information and energetics in protein structures. *Bioinformatics (Oxford, England)*, 27(14), 2003–2005. <https://doi.org/10.1093/bioinformatics/btr191>
- Martini, S., Bonechi, C., Foletti, A., & Rossi, C. (2013). Water-protein interactions: The secret of protein dynamics. *TheScientificWorldJournal*, 2013, 138916. <https://doi.org/10.1155/2013/138916>
- Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48(3), 443–453. [https://doi.org/10.1016/0022-2836\(70\)90057-4](https://doi.org/10.1016/0022-2836(70)90057-4)
- Obadan, A. O., Santos, J., Ferreri, L., Thompson, A. J., Carnaccini, S., Geiger, G., Reiche, A. S. G., Rajão, D. S., Paulson, J. C., & Perez, R. (2019). Flexibility in vitro of amino acid 226 in the receptor-binding site of an H9 subtype influenza A virus and its effect in vivo on virus replication, tropism, and transmission. *Journal of Virology*, 93(6), e02011-18.
- Pahari, S., Li, G., Murthy, A. K., Liang, S., & Fragoza, R. (2020). SAAMBE-3D: Predicting effect of mutations on protein-protein interactions. *International Journal of Molecular Sciences*, 48(3), 2563.
- Penfold, C. N., Healy, B., Housden, N. G., Boetzel, R., Vankemmelbeke, M., Moore, G. R., Kleanthous, C., & James, R. (2004). Flexibility in the receptor-binding domain of the enzymatic colicin E9 is required for toxicity against *Escherichia coli* cells. *Journal of Bacteriology*, 186(14), 4520–4527. <https://doi.org/10.1128/JB.186.14.4520-4527.2004>
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. <https://doi.org/10.1002/jcc.20084>
- Pichierri, F. (2013). Dipole-dipole interactions in protein-protein complexes: A quantum mechanical study of the ubiquitin-Dsk2 complex.
- Piovesan, D., Minervini, G., & Tosatto, S. C. E. (2016). The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Research*, 44, 367–374. <https://doi.org/10.1093/nar/gkw315>
- Rice, P., Longden, L., & Bleasby, A. (2000). EMBOS: The European molecular biology open software suite. *Trends in Genetics: TIG*, 16(6), 276–277. [https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2)
- Rodrigues, C. H. M., Pires, D. E. V., & Ascher, D. B. (2018). DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Research*, 46(W1), W350–W355. <https://doi.org/10.1093/nar/gky300>
- Roy, S., Jaiswar, A., & Sarkar, R. (2020). Dynamic asymmetry exposes 2019-nCoV prefusion spike. *The Journal of Physical Chemistry Letters*, 11(17), 7021–7027. <https://doi.org/10.1021/acs.jpcl.0c01431>
- Sargsyan, K., Grauffel, C., & Lim, C. (2017). How molecular size impacts RMSD applications in molecular dynamics simulations. *Journal of Chemical Theory and Computation*, 13(4), 1518–1524. <https://doi.org/10.1021/acs.jctc.7b00028>
- Strickler, S. S., Gribenko, A. V., Gribenko, A. V., Keiffer, T. R., Tomlinson, J., Reihle, T., Loladze, V. V., & Makhatadze, G. I. (2006). Protein stability and surface electrostatics: A charged relationship. *Biochemistry*, 45(9), 2761–2766. <https://doi.org/10.1021/bi0600143>
- Uskoković, V. (2020). Why have nanotechnologies been underutilized in the global uprising against the coronavirus pandemic? *Nanomedicine (London, England)*, 15(17), 1719–1734. <https://doi.org/10.2217/nmm-2020-0163>

- Vangone, A., & Bonvin, A. M. J. J. (2015). Contacts-based prediction of binding affinity in protein–protein complexes. *eLife*, 4, 1–15. <https://doi.org/10.7554/eLife.07454>
- Vascon, F., Gasparotto, M., Giacomello, M., Cendron, L., Bergantino, E., Filippini, F., & Righetto, I. (2020). Protein electrostatics: From computational and structural analysis to discovery of functional fingerprints and biotechnological design. *Computational and Structural Biotechnology Journal*, 18, 1774–1789. <https://doi.org/10.1016/j.csbj.2020.06.029>
- Voet, A., Berenger, F., & Zhang, K. Y. J. (2013). Electrostatic similarities between protein and small molecule ligands facilitate the design of protein-protein interaction inhibitors. *PLoS One*, 8(10), e75762. <https://doi.org/10.1371/journal.pone.0075762>
- Vriend, G., & Sander, C. (1993). Quality control of protein models: Directional atomic contact analysis. *Journal of Applied Crystallography*, 26(1), 47–60. <https://doi.org/10.1107/S0021889892008240>
- Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., Lu, G., Qiao, C., Hu, Y., Yuen, K. Y., Wang, Q., Zhou, H., Yan, J., & Qi, J. (2020). Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell*, 181(4), 894–904.e9. <https://doi.org/10.1016/j.cell.2020.03.045>
- Webster, D. M., Henry, A. H., & Rees, A. R. (1994). Antibody-antigen interactions. *Current Opinion in Structural Biology*, 4(1), 123–129. [https://doi.org/10.1016/S0959-440X\(94\)90070-1](https://doi.org/10.1016/S0959-440X(94)90070-1)
- Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35(Web Server), W407–W410. <https://doi.org/10.1093/nar/gkm290>
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., Graham, B. S., & McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 367(6483), 1260–1263. <https://doi.org/10.1126/science.abb2507>
- Wu, N. C., Yuan, M., Liu, H., Lee, C. D., Zhu, X., Bangaru, S., Torres, J. L., Caniels, T. G., Brouwer, P. J. M., van Gils, M. J., Sanders, R. W., Ward, A. B., & Wilson, I. A. (2020). An alternative binding mode of IGHV3-53 antibodies to the SARS-CoV-2 receptor binding domain. *bioRxiv*, 2020.07.26.222232. <https://doi.org/10.1101/2020.07.26.222232>
- Yang, J., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, 43(W1), W174–W181. <https://doi.org/10.1093/nar/gkv342>
- Yi, C., Sun, X., Ye, J., Ding, L., Liu, M., Yang, Z., Lu, X., & Zhang, Y. (2020). Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. *Cellular & Molecular Immunology*, 17(6), 621–630. <https://doi.org/10.1038/s41423-020-0458-z>
- Yin, Y., & Wunderink, R. G. (2018). MERS, SARS and other coronaviruses as causes of pneumonia. *Respirology (Carlton, Vic.)*, 23(2), 130–137. <https://doi.org/10.1111/resp.13196>
- Zhou, G., Chen, M., Ju, C. J. T., Wang, Z., Jiang, J.-Y., & Wang, W. (2020). Mutation effect estimation on protein-protein interactions using deep contextualized representation learning. *NAR Genomics and Bioinformatics*, 2(2), lqaa015. <https://doi.org/10.1093/nargab/lqaa015>
- Zhou, H. X., & Pang, X. (2018). Electrostatic interactions in protein structure, folding, binding, and condensation. *Chemical Reviews*, 118(4), 1691–1741. <https://doi.org/10.1021/acs.chemrev.7b00305>