

Article

# Site Density Functional Theory and Structural Bioinformatics Analysis of the SARS-CoV Spike Protein and hACE2 Complex

Nitesh Kumawat <sup>1,†</sup>, Andrejs Tucs <sup>2,†</sup>, Soumen Bera <sup>3</sup>, Gennady N. Chuev <sup>4,\*</sup>, Marat Valiev <sup>5</sup>,  
Marina V. Fedotova <sup>6</sup>, Sergey E. Kruchinin <sup>6</sup>, Koji Tsuda <sup>2,7,8</sup>, Adnan Sljoka <sup>7,9,\*</sup> and Amit Chakraborty <sup>10,\*</sup>

<sup>1</sup> School of Mathematics, Statistics and Computational Sciences, Central University of Rajasthan, Ajmer 305817, India; 2018phdmt008@curaj.ac.in

<sup>2</sup> Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa 277-8568, Japan; a.tucs@edu.k.u-tokyo.ac.jp (A.T.); tsuda@k.u-tokyo.ac.jp (K.T.)

<sup>3</sup> Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA; soumenmath4189@gmail.com

<sup>4</sup> Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, 142290 Pushchino, Russia

<sup>5</sup> Molecular Sciences Software Group, Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352, USA; marat.valiev@pnl.gov

<sup>6</sup> G.A. Krestov Institute of Solution Chemistry, Russian Academy of Sciences, 153045 Ivanovo, Russia; hebrus@mail.ru (M.V.F.); kruchininse@gmail.com (S.E.K.)

<sup>7</sup> RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027, Japan

<sup>8</sup> Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, Tsukuba 305-0044, Japan

<sup>9</sup> Department of Chemistry, York University, Toronto, ON M3J 1P3, Canada

<sup>10</sup> Department of Mathematics, School of Physical Sciences, Sikkim University, Gangtok 737102, India

\* Correspondence: genchuev@rambler.ru (G.N.C.); adnan.sljoka@riken.jp (A.S.); achakraborty@cus.ac.in (A.C.)

† These authors contributed equally to this work.



**Citation:** Kumawat, N.; Tucs, A.; Bera, S.; Chuev, G.N.; Valiev, M.; Fedotova, M.V.; Kruchinin, S.E.; Tsuda, K.; Sljoka, A.; Chakraborty, A. Site Density Functional Theory and Structural Bioinformatics Analysis of the SARS-CoV Spike Protein and hACE2 Complex. *Molecules* **2022**, *27*, 799. <https://doi.org/10.3390/molecules27030799>

Academic Editors: Mihai V. Putz, Ottorino Ori and Savas Kaya

Received: 17 December 2021

Accepted: 21 January 2022

Published: 26 January 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The entry of the SARS-CoV-2, a causative agent of COVID-19, into human host cells is mediated by the SARS-CoV-2 spike (S) glycoprotein, which critically depends on the formation of complexes involving the spike protein receptor-binding domain (RBD) and the human cellular membrane receptor angiotensin-converting enzyme 2 (hACE2). Using classical site density functional theory (SDFT) and structural bioinformatics methods, we investigate binding and conformational properties of these complexes and study the overlooked role of water-mediated interactions. Analysis of the three-dimensional reference interaction site model (3DRISM) of SDFT indicates that water mediated interactions in the form of additional water bridges strongly increases the binding between SARS-CoV-2 spike protein and hACE2 compared to SARS-CoV-1-hACE2 complex. By analyzing structures of SARS-CoV-2 and SARS-CoV-1, we find that the homotrimer SARS-CoV-2 S receptor-binding domain (RBD) has expanded in size, indicating large conformational change relative to SARS-CoV-1 S protein. Protomer with the up-conformational form of RBD, which binds with hACE2, exhibits stronger intermolecular interactions at the RBD-ACE2 interface, with differential distributions and the inclusion of specific H-bonds in the CoV-2 complex. Further interface analysis has shown that interfacial water promotes and stabilizes the formation of CoV-2/hACE2 complex. This interaction causes a significant structural rigidification of the spike protein, favoring proteolytic processing of the S protein for the fusion of the viral and cellular membrane. Moreover, conformational dynamics simulations of RBD motions in SARS-CoV-2 and SARS-CoV-1 point to the role in modification of the RBD dynamics and their impact on infectivity.

**Keywords:** site density functional theory; 3DRISM; spike protein binding to human receptor; coronavirus; structural rigidity; normal mode analysis; conformational dynamics

## 1. Introduction

The coronavirus pandemic COVID-19, caused by Severe Acute Respiratory Syndrome coronavirus (SARS-CoV-2), continues to pose a serious threat across continents. Despite

large volume of data on molecular structures of SARS-CoV-2 [1], essential details of molecular mechanism of interaction between SARS-CoV-2 and human cellular membrane receptor angiotensin-converting enzyme 2 (hACE2) are still under investigation. To better understand virus transmissibility, a large number of molecular studies have focused on the viral entry processes that are mediated by the spike glycoprotein (S protein), which is responsible for the receptor recognition and membrane fusion [2,3]. With the use of S protein the coronavirus hijacks the hACE2, which is highly expressed in lungs, heart, kidneys and intestine cells [2]. The S protein protomer is made of two subunits S1 and S2. The former unit, which comprises the receptor-binding domain (RBD), binds to the peptidase domain of hACE2 and contributes to stabilization of the prefusion conformational state. The SARS-CoV uses ectodomain trimer to mediate this viral entry [4,5]. Both SARS-CoV-1 and SARS-CoV-2 recognize hACE2 through its RBD, which is positioned within the flexible S1 unit of S-protein protomer.

Various computational studies have been conducted to understand the mechanism of viral infection and how it relates to inhibition of spike RBD–hACE2 binding. In order to have deeper understanding of this process, we will focus on applications of the density functional theory (DFT). It is an efficient tool for treating many-body phenomena in condensed matter systems. As a general method, DFT has been applied to problems in both quantum and classical domains. However, its impact of the applications to these areas is dramatically different. Advantages and merits of quantum DFT are well documented (see for example, [6,7]), whereas benefits and achievements of classical DFT are less evident. There are brilliant examples of DFT applications to simple fluids (see, for example [8]). However, the DFT treatment of molecular liquids remains a challenge. The bottleneck is that it requires evaluations of six-dimensional density profiles  $\rho(\mathbf{r}, \mathbf{w})$  in the case of rigid molecules:

$$\rho(\mathbf{r}, \mathbf{w}) = \left\langle \sum_{s=1}^N \delta(\mathbf{r} - \mathbf{x}_s) \delta(\mathbf{w} - \mathbf{w}_i) \right\rangle \quad (1)$$

where brackets mean statistical averaging, while  $\mathbf{x}_s$  denotes position of solvent molecule  $i$  (e.g., center of mass) and  $\mathbf{w}_i$  its angular orientation (e.g., Euler angles). Despite of a recent progress in this field [9,10] the problem is still not solved for biological macromolecules due to the “curse of dimensionality”.

An alternative has been proposed by Chandler [11] who introduced site density representation, i.e., density of individual atoms in solvent molecules. The idea has led to the development of the site density functional theory (SDFT) based on the atomistic or site density description:

$$\rho_i(\mathbf{r}) = \left\langle \sum_{s=1}^N \delta(\mathbf{r} - \mathbf{r}_{is}) \right\rangle \quad (2)$$

where indices  $i$  and  $s$  refer to atomic site and molecule indices respectively. This site based view is well adapted for problems where atomic level of details is of importance. One of the main benefits of the SDFT is ability to analyze contributions down to individual atoms and treatment of chemical bonds. The latter are strong but localized interactions, which have to be considered in conjunction with softer long-range inter-molecular forces. Simultaneous rigorous treatment of these two very different interactions scales poses a significant problem, constituting a single biggest obstacle in site density models of molecular liquids. This has been recognized in the original formulation by Chandler, McCoy, Singer (CMS) [12,13] and follow-up developments [14–21]. Despite a recent progress in the SDFT treatment of simple solutes [22–24], its practical application to biological molecules has turned out to be difficult. Underlying reason of this difficulty is related to treatment of chemical bonds, that necessitates to perform angular averaging directly. There is a simplified SDFT version referred to as the three-dimensional reference interaction site model (3DRISM) [25,26]. It treats molecular liquids as an “effective” atomic mixture, which makes it possible to utilize many of the techniques from simple fluid theory. Namely, the 3DRISM reduces the problem

to numerical solution of Ornstein-Zernike integral equations similar to those for simple fluids. The model is very popular for treatment of small bio-organic solutes [27–54].

Despite certain success of the 3DRISM applications, only few of them are devoted to coronaviruses [55,56]. The reason for that lies not only in the large size of viruses but also a difficulty in treatment of the flexibility of coronavirus structure. Although there is a possibility to address this issue by combining 3DRISM and conventional molecular mechanics treatment [57], the computational costs of such combined method makes its applications to virus systems impractical. One possible way to overcome this bottleneck is to integrate 3DRISM with bioinformatics methods. The latter is widely used to evaluate protein-protein and protein-ligand interactions. The structural bioinformatics approach involves a wide range of tools such as molecular docking, evaluations of affinity and binding constants, constrained geometric simulations (CGS) for conformational sampling, normal mode analysis (NMA), etc. [58]. Typically, to reduce computational costs, the bioinformatics approach limits molecular level description to the interacting protein and ligands, and treats environmental effects in a simplified manner (e.g., generalized Born surface area (GBSA) model [59]). The latter is accurate only when electrostatic interactions dominate. However, short ranged protein-solvent interactions involving hydrogen bond formation, steric repulsion, and Van-der-Waals attraction play an essential role in the protein-protein complexation. The 3DRISM seems to be a suitable methodological choice to reveal a role of electrostatic and short-ranged protein-solvent interactions in solvation thermodynamics, as the theory can be used to detect water-mediated contacts in protein complexes and identify favorable or unfavorable contributions to the binding free energy. As such, integration of 3DRISM calculations with structural bioinformatics treatments represents an attractive strategy. The latter may be used to select “regions of interests” that can be analyzed by 3DRISM, providing an insight into the role of water-mediated interactions. Furthermore bioinformatics analysis can help to evaluate the role of rigidity and flexibility of the complexes under consideration. The goal of this paper is to test the utility of integrated classical SDFT and bioinformatics methodology for simulating COV spike protein. In particular, we investigate formation and stability of SARS-CoV-2-hACE2 and SARS-CoV-1-hACE2 complexes. The scope of the paper is the following. First, we provide the preliminary bioinformatics treatment to reveal main characteristics of binding between the RBD domain of spike protein and hACE2 receptor (Section 2.1). Then, we analyze the role of water-mediated interactions in the formation of this complex using 3DRISM (Section 2.2). We also test how replacement of SARS-CoV-1 by SARS-CoV-2 will affect the binding of RBD domain to hACE2. To validate the results we provide further bioinformatics analysis using the NMA and CGS methods (Section 2.3). Finally, we summarize the revealed effects and discuss briefly benefits and bottleneck of the combined 3DRISM and structural bioinformatics tools. We outline the basics of the classical SDFT and the main features of the applied bioinformatics tools in Section 4. Essential details of the evaluations and additional evidences can be found in Supporting Information.

## 2. Results

### 2.1. Preliminary Bioinformatics Treatment

SARS-CoV-2-S protein holds 98% sequence similarity with the bat coronavirus RaTG13. A most critical variation observed in CoV-2 is an insertion “RRAR” at the furin recognition site at S1/S2 junction, while SARS-CoV-1 has single arginine at that site. Besides this insertion, 60% residual substitutions are noted at the RBD domain [60]. Examining how such differences contribute to higher recognition capability of hACE2 receptor is important to underpin the therapeutic target that can prevent virus entry. Here we analyze respective protomers of four S-protein cryo-EM structures: SARS-CoV-1 (pdb id: 6CRZ) [61], SARS-CoV-2 (pdb id: 6VYB) [62], SARS-CoV-RBD-hACE2 complex (pdb id: 2AJF) [63], and SARS-CoV-2-RBD-hACE2 complex (pdb id: 6M17) [64]. Recent work has revealed that the S-protein in the open state with at least one RBD in the “up” conformation, corresponds to the receptor accessible conformation that can bind to hACE2. In comparison to the CoV-1-RBD up state,

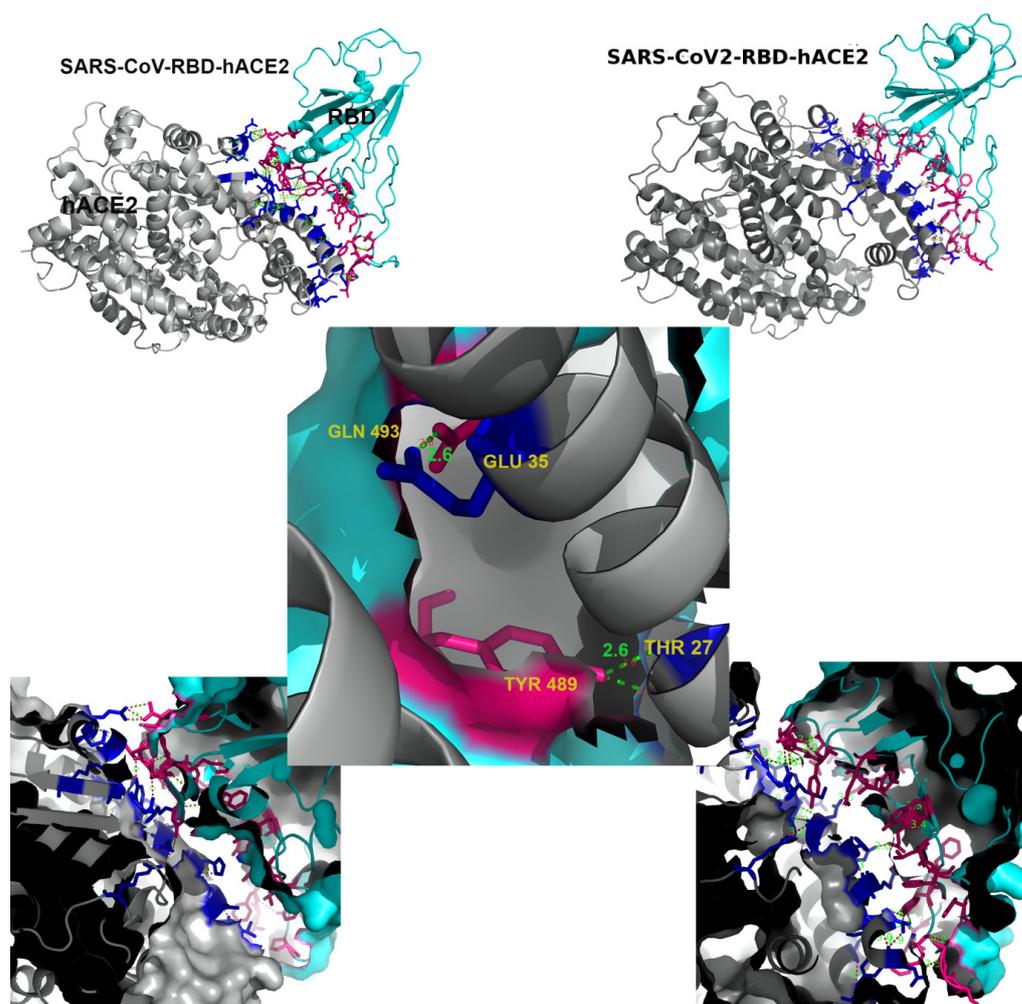
CoV-2-RBD has expanded its net surface area, undergoing a large local conformational change, and a relatively large amplitude anti-phase-like RBD motion in the slow-motion second normal mode. In order to recognize the host receptor, the RBD of S1 undergoes hinge-like conformational motions. In a recently reported structure of SARS-CoV-2, it was observed that hACE2 can only bind when the RBD (residues 336-518) adapts an open up-conformational state [65]. Peptidase domain of hACE2 clashes when all the RBD domains of the homotrimer SARS-CoV-2 S are in down conformational state. When compared with the SARS-CoV-1 RBD, an hACE2-binding ridge in SARS-CoV-2 has more compact conformation and several changes in the interface residues which stabilize two virus-binding hotspots at the RBD-hACE2 interface [64].

Interface interactions certainly play crucial role for binding and dynamics of domain motion. Initially, we use the Prodigy webserver to evaluate these interactions [66]. These calculations indicate that interface between CoV-2-RBD and hACE2 is relatively larger than the SARS-CoV-1 complex, with higher potential for large intermolecular interactions. The interface properties of the complexes are listed in Table 1. In both complexes, almost all interface molecular interactions involve loops and small parts of beta-sheets secondary structures of RBD and a single alpha helical structure of hACE2, indicating interface instability (Figure 1). We find uneven distribution of interface interactions along this single hACE2 alpha-helix. While the two end portions and the middle part of the helix hold most interactions, there are other interacting residues between CoV-1 and CoV-2 complex (see Tables S1 and S2 in the Supporting Information). In particular, we note that hACE2.GLU35[OE2]-RBD.GLN493[NE2] and hACE2.THR27[O]-RBD.TYR489[OH] come closer (2.6 Å) and form H-bonds in the CoV-2-RBD-hACE2 complex (Figure 1). Inclusion of these two additional polar interactions may affect subsequent proteolytic processing of S protein and membrane fusion with the S2 unit. The Prodigy treatment yields almost similar binding free energy for both complexes at T = 25 °C. The latter contradicts the experimental observations, for example, recent measurements by surface plasmon resonance at T = 37 °C provide evidence [67] that the binding energy is lower for the CoV-2-hACE2 complex than that for CoV-1-hACE2 complex. These experiments estimate the difference  $\Delta G_{exp}$  between the relevant binding energies as  $\Delta G_{exp} = -1.1$  kcal/mol.

**Table 1.** Interface properties of RBD-hACE2 complexes for CoV-1 and CoV-2.

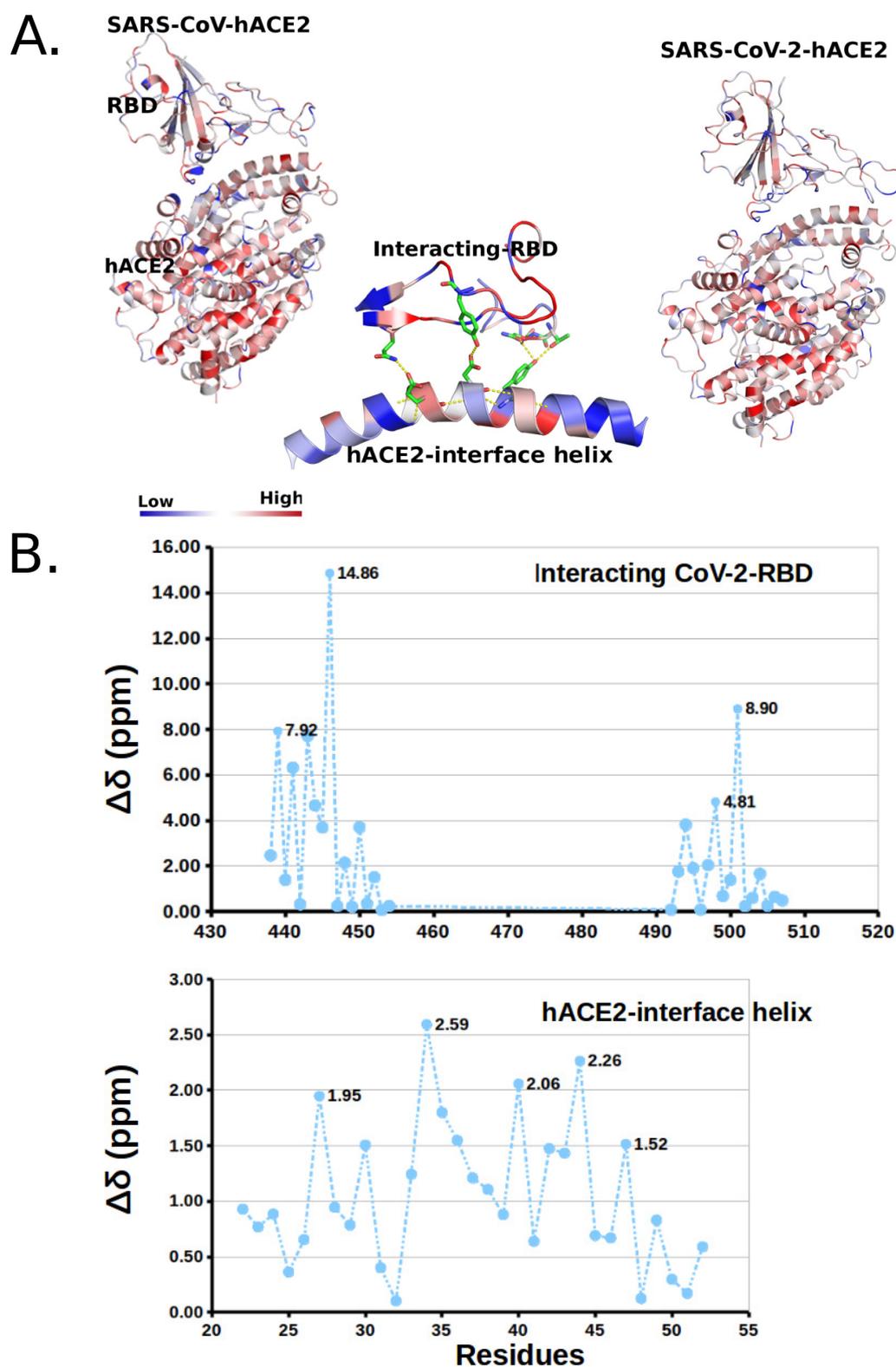
Parameter	CoV-1-RBD-hACE2	CoV-2-RBD-hACE2
	structure	
Bound-RBD surface area (Å <sup>2</sup> )	18,435.6	19,866.4
Interface area (Å <sup>2</sup> )	796.7	827.7
	thermodynamics	
Binding Energy (kcal/mol)		
PRODIGY	-11.1	-10.8
EXP [67]	-10.7	-11.8
MM/GBSA [68]	-10.0	-24.9
CG [69]	-66.4	-70.7
3DRISM	-50.1	-57.2
$U_{LJ}$ (kcal/mol) <sup>a</sup>	-281	-261
$U_C$ (kcal/mol) <sup>a</sup>	-4887	-5160

<sup>a</sup>  $U_{LJ}$  and  $U_C$ —Lennard-Jones and Coulomb parts of complex-water interactions.



**Figure 1.** Interface interactions between CoV-1 and CoV-2 receptor-binding domain (RBD) and the hACE2. In both complexes, almost all interface molecular interactions involve loops and small parts of beta-sheets secondary structures of RBD and a single alpha helical structure of hACE2. The hACE2.GLU35–RBD.GLN493 and hACE2.THR27–RBD.TYR489 (2.6 Å) form H-bond interactions indicated by dashed lines. Pink and blue color represent RBD and hACE2 interfaces respectively.

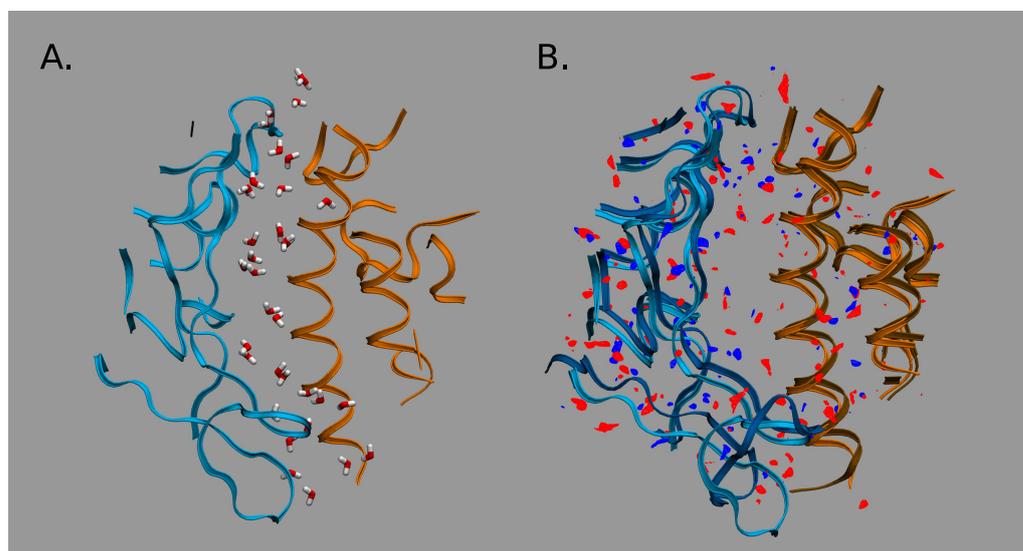
In order to check the reliability of the predicted interface interactions, we have calculated  $^{13}\text{C}$  chemical shifts ( $\delta$ ) using structural coordinates of both the CoV-1 and CoV-2 hACE2 bound complex. This calculation is carried out using SHIFTX2 [70] which combines ensemble machine learning approaches with the sequence alignment-based methods. We note that there is no overall significant differences in  $\delta$  (ppm) between the hACE2 bound CoV-1 and CoV-2 complex (Figure 2A). However, relative absolute differences  $\Delta\delta$  are prominent at the interface between hACE2 and RBD (Figure 2A middle). Detected uneven distribution and changes of interface interactions are shown to be strongly correlated to deviation in the chemical shifts ( $\Delta\delta$ ). In particular, the hACE2 interface helix and the interacting RBD domain of the CoV-2 RBD shows significant difference in  $\Delta\delta$  (Figure 2B). As expected, very high  $\Delta\delta$  values are noted at and nearby H-bond forming residues GLU 35, THR 27 on the hACE2 helices and GLN 493 and TYR 489 on the RBD interface (Figure 2B). A large  $\Delta\delta$  pick of 14.86 ppm on the RBD-interface and 2.50 ppm on the hACE2 helix are noted, with significant changes in the middle and end parts of the interacting interface helix. Furthermore, large picks in  $\Delta\delta$  corresponds to AA substitutions. For example,  $\Delta\delta$  pick of 14.86 and 8.90 ppm correspond to the substitution THR433GLY446 and THR487ASN501 on the interacting RBD domain.



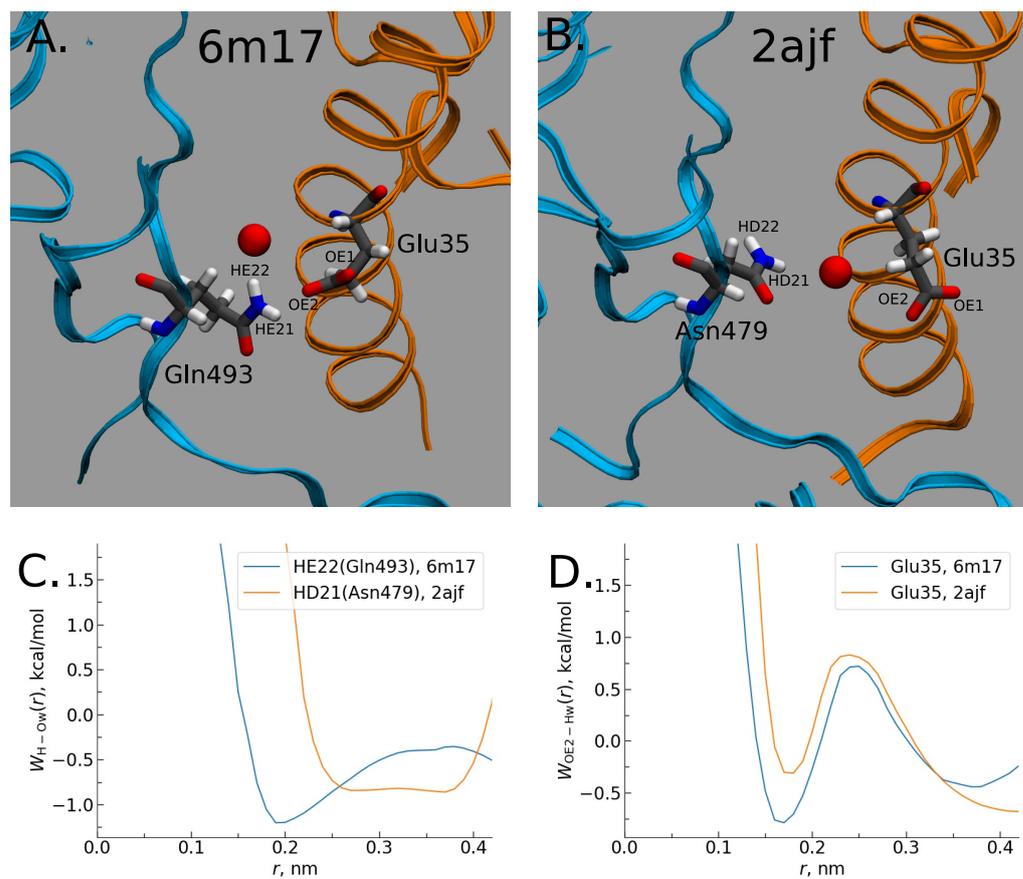
**Figure 2.**  $^{13}\text{C}$  Chemical shifts  $\delta$  caused by interface interactions between hACE2 and RBD. (A) Changes in  $\delta$  for the complexes (top two figures), the middle figure shows the deviation in chemical shifts ( $\Delta\delta$ ) at the interface interacting sites. (B) Change in  $\Delta\delta$  observed at H-bond forming residues GLU 35, THR 27 on the hACE2 helices and GLN 493 and TYR 489 on RBD interface.

### 2.2. 3DRISM Studies of Water-Mediated Interactions in SARS-CoV-1-hACE2 and SARS-CoV-2-hACE2 Complexes

As we have seen, the preliminary bioinformatics treatment reveals an essential role of interfacial interactions in the formation of RBD-hACE2 complexes. However, the simple Prodigy analysis can not evaluate the relative stability of SARS-CoV-hACE2 complexes. To provide further insight into the problem, we perform the 3DRISM treatment (see Section 4). Figure 3A demonstrates a water distribution in the interfacial region for CoV-2-hACE2 complex. Water molecules are depicted only in the most probable positions. As it is seen, the interfacial water interacts strongly with the RBD as well as with the hACE2 subunits, the latter confirms the previous molecular dynamic simulations [71]. However, our calculations yield the water-mediated interactions to be stronger for the CoV-2-hACE2 complex than those for the CoV-1-hACE2. See for example Figure 3B demonstrating a relative difference in distributions of water oxygens (blue) and hydrogens (red) between the two complexes. Moreover, the water interacts more strongly for CoV-2-hACE2 rather than for pure coronavirus (see Figure S4 in the Supporting Information). To confirm this effect we calculated protein-solvent interaction energies (Table 1). Although the Lennard-Jones energy between the CoV-1-hACE2 and water is stronger than that for the CoV-2-hACE2 complex, the total interaction energy has a more profit for the latter complex due to a strong decrease in the electrostatic energy. This observed effect can be explained by a polarization of interfacial water which is stronger for the CoV-2-hACE2 complex. Apart from it, the latter complex forms water bridges whose number is larger by 2 than in the case of CoV-1-hACE2. At the same time the water bridges is stronger in the CoV-2-hACE2 aggregate. To prove it, we have calculated the potential of mean force (pmf) between atoms of RBD and hACE2 receptor (Figure 4). It is clearly seen that bridging water molecule forms simultaneously hydrogen bonds with CoV-2 (Gln493) and hACE2 (Glu35) amino acids in the case of CoV-2. However, the similar hydrogen bond with bridging water is very weak and diffusive for the CoV-1-hACE2 complex (Figure 4). Therefore, we conclude that bridge water molecules play a significant and, perhaps, crucial role in stabilizing CoV-2-hACE2 complex.



**Figure 3.** Water distribution in the interfacial region of complexes: (A) For the CoV-2-hACE2 complex. (B) Differences in distributions of water oxygens (blue colour) and water hydrogens (red colour) between the CoV-2-hACE2 and CoV-hACE2 complexes. The RBD is indicated by blue ribbons, the hACE2 by orange ribbons, and the CoV-hACE2 is shown as background for the differences.



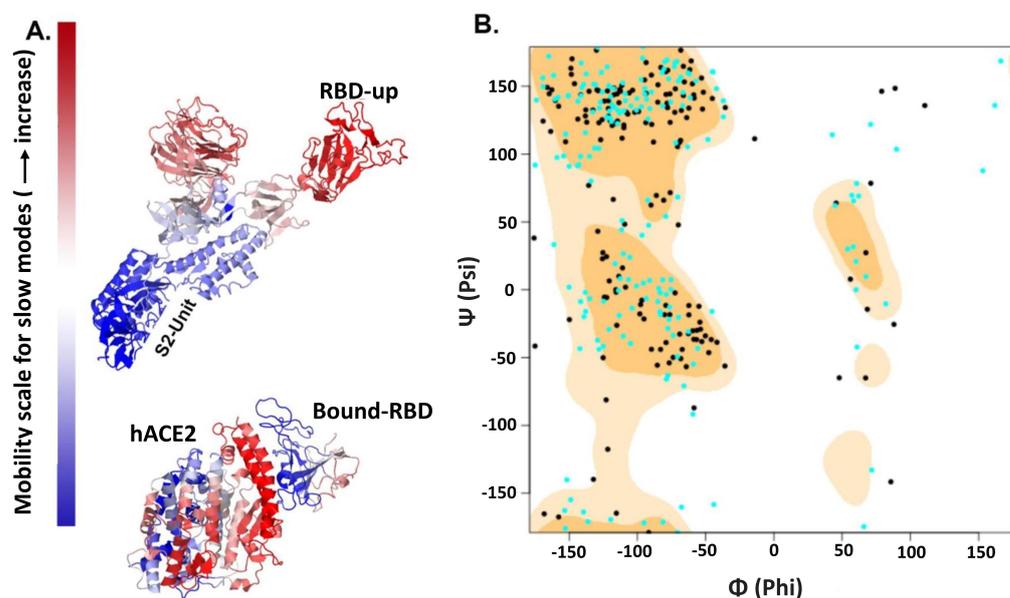
**Figure 4.** Water bridging the complexes: (A) Location of water oxygen bridging the CoV-2-hACE2 complex. (B) The same for the CoV-1-hACE2 complex. (C) The pmf of H-O distribution for the CoV-2 and CoV-water bridges. (D) The pmf of O-H for the hACE2-water bridges in the complexes.

We have also calculated the binding energies for the complexes under consideration. For this purpose, we split the RBD and hACE2 parts of the complexes and calculated them separately. The binding energy is evaluated as a difference between the sum of free energies of the subunits and that for the whole complex. The results are indicated in Table 1. We note that the absolute values of the calculated energies are varied as calculation box size changes, nevertheless the relative difference  $\Delta G$  between the binding energies remains the same  $\Delta G = -7.1$  kcal/mol. The latter is in an agreement with results obtained by coarse-grained (CG) simulations  $\Delta G_{CG} = -4.3$  kcal/mol [69] and with the use of molecular mechanics/GBSA calculations [68] yielding  $\Delta G_{GBSA} = -14.9$  kcal/mol. Simple 3DRISM analysis of contributions to the binding affinity indicates that the major difference come from electrostatic contributions. The latter correlates also with our estimation of Lennard-Jones and Coulomb parts in protein-water interactions. We note all the indicated methods seem to underestimate the difference in the binding energies with respect to the experimental data. The reason of this drawback seems to be ionic effects which can reduce the electrostatic contribution as well as entropic changes caused by relaxation of individual subunits. We note also that all the methods provide absolute values of the binding energy varying by several times. Moreover, detailed analysis of an influence of conformational changes [68] on the binding demonstrates that the energy and hence binding affinity depends strongly on a conformation state of the complexes, conformation angle between RDB and hACE2, etc. Therefore, additional analysis of these effects is to be required to prove the 3DRISM findings.

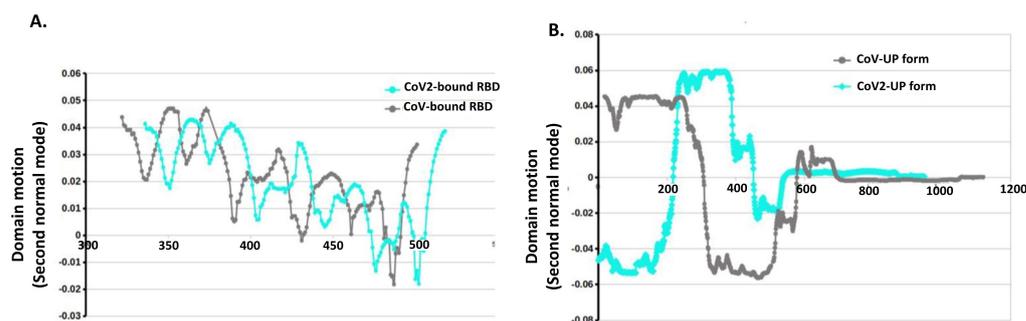
### 2.3. Posterior Bioinformatics Treatment by NMA and CGS

Significant conformational changes are commonly reflected in the differential domain motions. To indicate these changes we draw the Ramachandran plots for the complexes.

They display large changes in phi and psi dihedral angles for most of the RBD residues in favorable and allowed regions of the plot (Figure 5B). These analyses suggest characteristic conformational differences in hACE2-bound and non-bound RBD between SARS-CoV-1 and SARS-CoV-2 and, therefore, it necessitates further examination of its functional significance and rigidification properties of the all potential RBD conformations. To capture such differences, we have conducted NMA using the Gaussian network model and the anisotropic-network model, utilized in DynOmics webserver [72] and in iMODS [73]. We calculated the first twenty slowest modes for all the CoVs structures. The eigenvectors of these modes represent the global motions and the constrained residues help in identifying critical regions such as hinge-bending regions and thereby giving an idea of domain motions around these regions. This hybrid ENM has efficiently captured the dynamic differences between SARS-CoV-1 and SARS-CoV-2 with and without hACE2 bound. With the trimeric CoV-1 and CoV-2 macromolecular structure, we have calculated covariance matrix in iMODS which is computed using the Cartesian coordinates and the Karplus equation of collective motion in protein. This covariance matrix signifies coupling between pairs of residues. Overall, covariance patterns are very similar for CoV-1 and CoV-2; however, there are a very few sharp differences in few spots of the covariance matrices. As usual, such differences are prominent in the low-frequency normal modes. To examine the internal residual coupling and its effects, we have studied low-frequency normal modes for RBD-up forms of both the structures and noted that the second normal mode is capable to effectively capture these differences. It shows that RBD has relatively high amplitude motions for both the CoVs without hACE2 binding, while the S2-unit holds lower-amplitude motions (Figure 5A). When we compare this prefusion up-form RBD local motion along the residues of CoV-1 and CoV-2, it shows that the extended RBD region follows anti-phase like dynamics; CoV-2-RBD has positive eigenvector components in oppose to negative components of CoV-1-RBD (Figure 6B). In contrast, these anti-phase dynamics are mostly vanished due to hACE2 binding; hACE2 binding makes the RBD relatively stable, with little faster movements of CoV-2-RBD (Figures 5A and 6A).



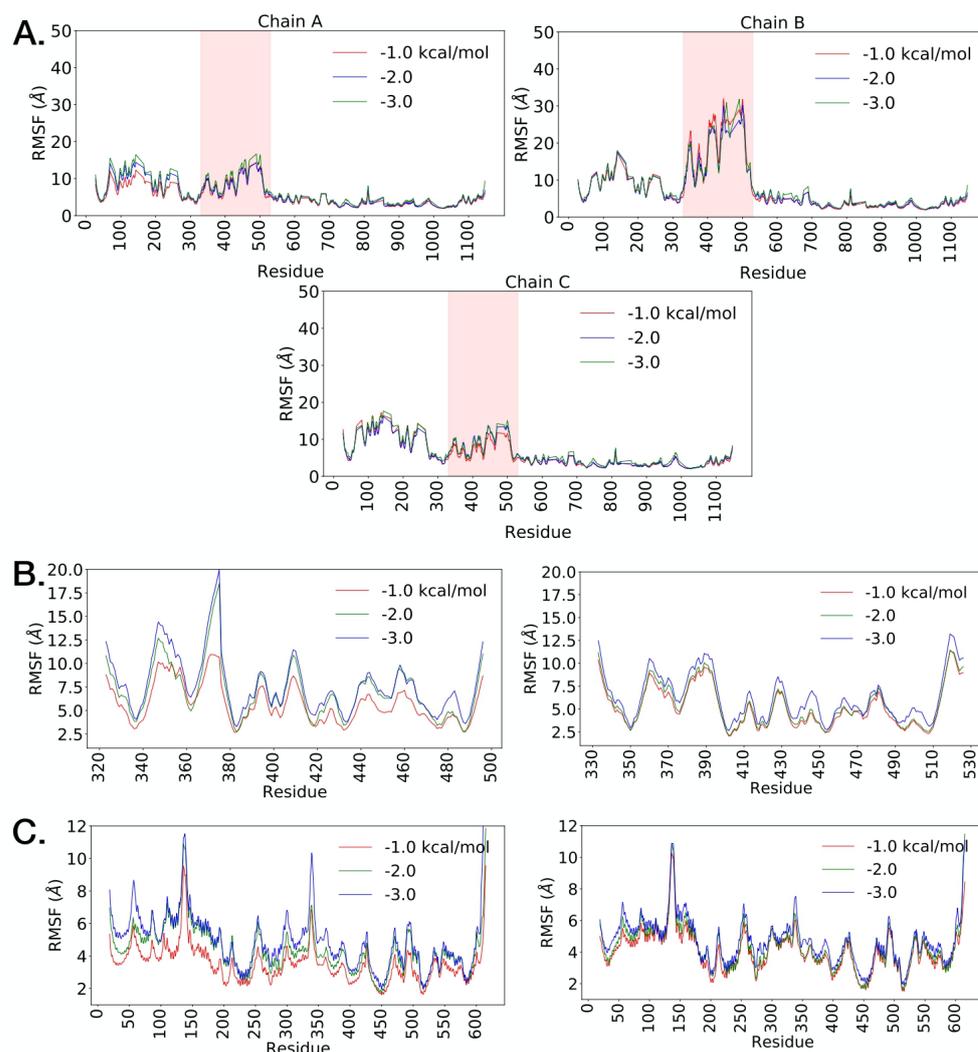
**Figure 5.** NMA (A) and Ramachandran plot (B) for RBD-hACE2 complexes. Cyan dots represent CoV-2-RBD AA residue phi-psi position, while black is for CoV-1 complex. We used low-frequency second normal mode to differentiate between SARS-CoV-1 and SARS-CoV-2 with and without hACE binding. (A) Its RBD-up conformation region in the S1-unit has higher mobility relative to other parts of the S-protein. In contrast, hACE2 binding make the RBD less mobile. (B) Ramachandran plot of AA residues of RBD bound with hACE2 shows significant conformational changes in the RBD.



**Figure 6.** Domain motions in second normal mode (A) Residue-wise eigenvectors of bound RBD AA residues is representing low-frequency motions. There are no significant differences in bound RBD motions between both the SARS-CoV-1 complexes. (B) Without hACE binding, RBD up-conformation protomer is showing significant differences in the domain motion in the RBD regions. It is like anti-phase-type dynamics between the two CoVs structures. Cyan dots represent CoV-2-RBD AA residue phi-psi position, while black is for CoV-1 complex.

We applied the computational method FRODAN [74,75] (see Section 4) on the whole spike protein with the RBD in the up-state and the RBD-hACE2 complex of the single RBD domain bound to the hACE2. The method (see Figure 7) reveals RMSF fluctuations of the respective backbone atoms in both spike proteins at different hydrogen-bond energy cut-offs (temperatures). Consistent with the normal mode analysis, it is evident that for both spike proteins in SARS-CoV-1 and SARS-CoV-2, the RBD in the up conformation has higher conformational fluctuation relative to other domains. As we increase the hydrogen-bond energy cutoff (i.e., removing weak transient hydrogen bonds) conformational fluctuations in RBD tend to further increase. There is an overall increase in conformational dynamics in SARS-CoV-1 S structure in comparison to the SARS-CoV-2 (Table S4 in the Supporting Information). In particular, RBD has higher RMSF in SARS-CoV-1. This is in agreement with our rigidity analysis (see details below), where we have shown that SARS-CoV-2 retains its overall rigidity better than SARS-CoV-1. Furthermore, large-scale computing efforts via Folding@home project have carried out millisecond MD simulation [76], where it was shown the RBD domain in the up state of the SARS-CoV-1 exhibits higher deviation from the respective crystal structure in comparison to the SARS-CoV-2 [76]. Interestingly, one of the RBDs in SARS-CoV-1 that is initially in the down configuration transitions to the open configuration. This correlates with previous studies which have shown that for the SARS-CoV-1, the two-up and three-up RBD states are also populated in the unbound spike protein [61]. On the other hand, for SARS-CoV-2 previous studies indicated that the two-up and three-up states are rarely observed [77]. This trend is reflected also in our simulation results. For SARS-CoV-2, the RBDs that are initially in the down state remain in the closed state at wide range of energy cut-offs. Overall, it is evident that temperature increase affects the SARS-CoV-1 spike protein's conformational stability in the more pronounced way than the SARS-CoV-2. This difference may provide clues why RBD binds tighter with hACE2 in SARS-CoV-2, and a general trend that SARS-CoV-1 is more sensitive to environment conditions than SARS-CoV-2 [64,78,79].

Figure 7 shows RMSF profiles of RBD and hACE2 in the complex at different energy cut-offs. Similarly, as in the whole spike protein, the increase in the hydrogen-bond energy cutoff leads to the overall higher RMSF values in both complexes. The RBD domain fluctuates less in SARS-CoV-2. The average RMSF in SARS-CoV-2 is lower at all considered cut-offs (Table S4 in the Supporting Information). This is supported by rigidity analysis findings that are indicating excessive flexibility in the SARS-CoV-1-hACE2 complex. Overall, it is evident that the stability of SARS-CoV-2-hACE2 complex results from the stronger interface contacts [64]. On the other hand, with the increase in the hydrogen-bond energy cutoff, hACE2 fluctuation magnitude increases in the similar manner in both SARS-CoV-1 and SARS-CoV-2 (Table S4).



**Figure 7.** Conformational dynamics analysis using FRODAN method: (A) Backbone RMSF profiles SARS-CoV-2 Spike protein with RBD shown in pink. (B) RMSF of the RBD domain in the complex with hACE2 (left SARS-CoV-1, right SARS-CoV-2). (C) RMSF of the hACE2 domain in the complex (left SARS-CoV-1, right SARS-CoV-2).

To provide further insight to structural flexibility, we have carried out rigidity analysis using Floppy Inclusion and Rigid Substructure Topography (FIRST) program [80], which is based on the pebble game algorithm and techniques in mathematical rigidity theory [81]. We used FIRST to decompose the spike protein structures into rigid clusters with the input of H-bond energy cutoff (H-cut-off, kcal/mol). With H-cut-off of  $-0.1$ ,  $-0.5$ ,  $-1.0$ ,  $-1.5$ ,  $-2.0$ ,  $-2.5$ , and  $-3.0$ , CoV-1 and CoV-2 spike with- and without hACE2 bound have been decomposed. Number of rigid clusters and degree of freedoms consistently increase with increasing H-cut-off. Without hACE2 bound, CoV-1 S have been decomposed into higher number of rigid clusters with much higher degrees of freedom relative to CoV-2 S (see Table S5 and Figure S5 in the Supporting Information). At a low energy cutoff ( $-0.5$  kcal/mol) SARS-CoV-2 is significantly more rigid than SARS-CoV-1, as it is dominated by a few large rigid clusters. As we increase the energy cutoff, SARS-CoV-1 continues to gain more flexibility while SARS-CoV-2 retains most of its rigidity. At  $-1.5$  kcal/mol SARS-CoV-1, including its RBD, has lost almost all internal structural rigidity, while SARS-CoV-2 still maintains significant rigidity. Extending this analysis to the whole complex, it is evident that the SARS-CoV-2 RBD-hACE2 complex at  $-1$  kcal/mol is dominated by one large rigid cluster, whereas SARS-CoV-1 RBD-hACE2 complex is more flexible consisting of several rigid cluster. As hydrogen bond energy cutoff is increased, SARS-CoV-1 RBD-hACE2

complex is losing its rigidity more rapidly than SARS-CoV-2 RBD-hACE2 complex. This strongly indicates that CoV-1 S becomes more flexible, generating a large conformational ensemble with less potential for binding with hACE2. This result is consistent with the recent observations in an extensive 0.1 s MD simulation which noted a large opening of spike with presence multiple cryptic epitopes [76]. In contrast, when RBD binds with hACE2, CoV-2-RBD-hACE2 becomes relatively more rigid than CoV-1-RBD-hACE2, which favors proteolytic processing for membrane fusions.

### 3. Discussion and Conclusions

Using recently reported cryo-EM structures of SARS-CoV-1 and CoV-2 S protein we have investigated binding and conformational properties of the CoV-hACE2 by utilizing a combination of SDFT and bioinformatics methods. The SDFT has been applied within the framework of 3DRISM. The method is used to reveal the role of water in interfacial interactions between RBD and hACE2 and to evaluate the difference in the binding affinities for the complexes under the consideration. The normal mode analysis, constrained geometric simulations, and the rigidity analysis have been used to analyze a role of flexibility and conformational motion on the stability of the CoV-hACE2 complexes. The 3DRISM analysis reveals the essential role of interfacial water for the complex stability. It indicates the CoV-2-hACE2 is more stable than CoV-1-hACE2 complex and has an order of magnitude lower dissociation constant than the CoV-1-hACE2 complex. The latter is in a quantitative and qualitative agreement with the experimental data and other calculations based on the CG and the MM/GBSA methods.

We note that the considered SDFT may be very useful to treat stability of coronavirus binding to hACE2, especially when the method is accompanied by further bioinformatics analysis for treating flexibility and conformational motions. To provide such analysis we have used the NMA, CGS, and FIRST tools. Using these treatments we indicate that the RBD-up conformation in the SARS-CoV-1 is less stable than in the SARS-CoV-2. When bound with higher surface area of CoV-2 RBD and large local conformational changes in the AA residues, it establishes higher interactions with the human receptor hACE2. These underlying conformational differences are illustrated with the higher RMSD of  $C\alpha$  atoms, changes in phi and psi dihedral angles and relatively large amplitude anti-phase-like RBD motion. In particular, we note that hACE2.GLU35–RBD.GLN493 and hACE2.THR27–RBD.TYR489 come closer in the CoV-2-RBD-hACE2 complex and forms important H-bonding interactions. However, it remains to be explored how efficiently CoV-2 regulate and open-up RBD for hACE2-binding. The RBD of CoV-2 forms stronger water bridges with the hACE2, and it plays a significant role in the total stabilization of the CoV-2/hACE2 complex. Inclusion of the two noted additional polar interactions affects structural flexibilities. We find such interaction changes critically affect structural flexibility. In particular, hACE2 binding makes the CoV-2 structure more rigid relative to CoV-1 complex. These hACE2 induced changes in structural flexibility favours subsequent proteolytic processing which is essential for membrane fusion. Understanding prefusion conformational dynamics as well as its binding mechanism to the receptor among closely related species is critical for designing vaccine and inhibitors to stop viral entry. Recent major therapeutic efforts are targeting the interactions between the SARS-CoV-2-S and the hACE2 [82,83]. This understanding certainly provides a clue for designing novel vaccine and antiviral drugs.

### 4. Methods

#### 4.1. Classical Site Density Functional Theory

We investigate hydrated protein complexes. Although there are numerous data about an essential role of ions in stabilization of protein complexes [84,85], we ignore this effect at the current level of the consideration and treat the solvent as a pure water. Then, the problem reduces to evaluations of site density profiles  $\rho_i(\mathbf{r})$  ( $i = O, H$ ) of inhomogeneous water subjected to external potential  $v(\mathbf{r})$  caused by solvent-protein interactions. As we indicated earlier [22,23,86], the rigorous classical SDFT formulation is to be based on construction of a

generating functional depending on two coupled variables site densities  $\rho_i(\mathbf{r})$  and site field  $J_i(\mathbf{r})$  and further evaluations of the functional. The densities and fields are to be obtained by the relevant minimization of the functional. Omitting all the technical details which have been published earlier [22,23,86], we provide here only the final relations for the fields and densities. In the vector form, they can be written as

$$\begin{cases} \rho(\mathbf{r}) = \rho_0[1 + \xi([\mathbf{J}], \mathbf{r})]e^{-\beta J(\mathbf{r})} \\ J(\mathbf{r}) = v(\mathbf{r}) + \phi(\mathbf{r}) \end{cases} \quad (3)$$

The meaning of the relations is rather simple. The first equation represents the density of an inhomogeneous molecular liquid subjected to the field  $J(\mathbf{r})$ , while the second one closes the self-consistent loop and indicates the field to be composed from the solvent-protein potential  $v(\mathbf{r})$  and intermolecular potential  $\phi(\mathbf{r})$  caused by interactions between solvent molecules. The first relation looks like as an barometric expression, but it contains an additional term  $\xi([\mathbf{J}], \mathbf{r})$  referred to as the correlation hole functional, which accounts for intra-molecular correlation effects. All variety of the SDFT models lies in expressions for the correlation hole functional and intermolecular potential. In the general case the latter can be presented as a density functional:

$$-\beta\phi(\mathbf{r}) = \int (\mathbb{S}_m^{-1}(|\mathbf{r} - \mathbf{r}'|) - \mathbb{S}^{-1}(|\mathbf{r} - \mathbf{r}'|)) \Delta\rho(\mathbf{r}') d\mathbf{r}' + b(\rho(\mathbf{r})) \quad (4)$$

where  $\mathbb{S}_m(\mathbf{r})$  and  $\mathbb{S}(\mathbf{r})$  is a structure factor of single water molecule and uniform water, respectively, while  $b(\rho(\mathbf{r}))$  is the bridge functional accounting contributions beyond the linear response. Various approximations can be used for this functional and its distance dependence can be extracted from molecular simulations [87–89].

The presence of the correlation hole functional reveals the main difference between treatment of molecular and simple liquids. The functional expresses the fact that the external field imposed on a given site will be also propagated to all other sites through the molecular bonds. The correlation hole functional can be expressed as a cluster expansion in terms of the Mayer function  $f(\mathbf{r}) = e^{-\beta J(\mathbf{r})} - 1$  and intramolecular correlation functions,  $\mathbb{D}^{(s)}$ :

$$\xi([\mathbf{J}], \mathbf{r}) = \sum_{s=2}^M \frac{\text{Tr}[\mathbb{D}^{(s)} f^{s-1}]}{(s-1)!} \quad (5)$$

where  $M$  is the total number of water sites. We note that the second order intramolecular correlation function  $\mathbb{D}^{(2)}(\mathbf{r})$  depends only on bond lengths  $l_{ij}$ , whereas the third order one involves angles between bonds, the forth one depend on dihedral angles, etc. As a result, the expansion for liquids whose molecules containing more than two sites requires integration over angles. Therefore, the SDFT application to water models like as SPC (simple point charge) model is a time consuming procedure.

The 3DRISM is an alternative to the rigorous SDFT treatment. The model assumes that correlation hole depends only on bond lengths and can be expressed as [22]:

$$\xi(\mathbf{r}) \approx \xi^{RISM}(\mathbf{r}) \equiv \exp \left[ \int (\mathbf{1} - \mathbb{S}_m^{-1}(|\mathbf{r} - \mathbf{r}'|)) \Delta\rho(\mathbf{r}') d\mathbf{r}' \right] - 1 \quad (6)$$

As a result, the relation between the site fields and densities are rather simplified. If we introduce new variable referred to as site direct correlation function  $c(\mathbf{r})$ :

$$c(\mathbf{r}) \equiv \int \mathbb{S}^{-1}(|\mathbf{r} - \mathbf{r}'|) \Delta\rho(\mathbf{r}') d\mathbf{r}' \quad (7)$$

Then the 3DRISM relations can be rewritten in the form similar to those in simple fluids:

$$\begin{aligned}\rho(\mathbf{r}) &= \rho_0 e^{-\beta v(\mathbf{r}) + \frac{\Delta\rho(\mathbf{r})}{\rho_0} + c(\mathbf{r}) + b(\mathbf{r})} \\ \Delta\rho(\mathbf{r}) &= \rho_0 \int \mathbb{S}(|\mathbf{r} - \mathbf{r}'|) c(\mathbf{r}') d\mathbf{r}'\end{aligned}\quad (8)$$

The first one of these relations is referred to as a closure, while the second one is the site Ornstein-Zernike integral equations. Due to neglecting angular dependencies in intramolecular correlations, the computational 3DRISM costs are by two orders less expensive than the SDFT calculations by (3)–(5). The 3DRISM input is protein-solvent potential  $v(\mathbf{r})$ , susceptibility of pure water  $\mathbb{S}(r)$ , bridge function (or functional)  $b(\mathbf{r})$ , and parameters determining thermodynamic state of water (i.e.,  $T$  and  $\rho_0$ ). It is important that calculating site densities by Equation (8) we fix a configuration of protein complex. In general, we can also calculate not only site densities but also changes in free-energy of the complex caused by hydration. Various approximation can be used for the free energy [33], we will use the so-called the pressure correction approximation which has been shown [48] to provide rather accurate evaluations for the hydration free energy.

We note the complexes under our consideration are too large to be studied by the 3DRISM in the fully atomistic format. For example, the size of 6M17 complex exceeds 300 Å, while the number of heavy atoms in this complex is more than 24 thousands. It is obvious that only residues near the interface region yield an essential contribution to the binding, whereas others located at large distances provide only external electrostatic field affected the interface region. To select the regions of with the essential contribution, we apply an iterative scheme. First, using the data obtained from the bioinformatics treatment we consider interfacial region which includes all contacts indicated in Table S2 in the Supporting Information. Then we extend the size of calculation box and recalculate the water distribution. The iterations are repeated until the changes in the calculated water distribution, caused by the extension of the calculation box, become less than  $10^{-3}$  in absolute values. The final clipped regions are shown in Figures S1 and S2 in Supporting Information. Details of the iteration scheme are given in Table S3 in the Supporting Information. The similar consideration has been carried out for SARS-CoV-1 (pdb id: 6CRZ) and SARS-CoV-2 (pdb id: 6VYB) complexes (see details in the Supporting Information). The 3DRISM with the 3D-Kovalenko-Hirata closure was used to calculate these complexes. The calculations were carried out for the studied complexes hydrated in water at ambient conditions. For water the modified version of the SPC/E model (MSPC/E) was used [90]. The corresponding LJ parameters of the solute atoms were taken from the ff14SB force fields [91]. The 3DRISM equations were solved on a 3D-grid of  $350 \times 320 \times 350$  points with a spacing of 0.025 nm. A residual tolerance of  $10^{-6}$  was chosen. These parameters are enough to accommodate the complex together with sufficient solvation space around it so that the obtained results are without significant numerical errors.

## 4.2. Bioinformatics Tools

### 4.2.1. Root-Mean-Squared Deviation (RMSD) of C- $\alpha$ Atoms

We used MatchMaker in UCSF-Chimera to superimpose structures. It performs a best fit after automatically identifying pair of residues. It considers similar structures to be superimposed while there are large sequence dissimilarities by producing residue pairing uses with the inputs of AA sequences and secondary structural elements. After superposition, it uses the usual Euclidian distance formula for calculating RMSD in angstrom unit with the use of PDB atomic coordinates files.

### 4.2.2. Normal Mode Analysis

DynOmics ENM Server [72] was used to perform normal mode analysis (NMA). NMA is a well-explored technique for exploring functional motions of proteins. It combines two elastic network models (ENMs)—the Gaussian Network Model (GNM) and the Anisotropic

Network Model (ANM) to evaluate the dynamics of structurally resolved systems, from individual molecules to large complexes and assemblies, in the context of their physiological environment. In the GNM model, network nodes are the C-alpha atoms and the elastic springs represented the interactions. We used GNM with interaction cut-off distance of 7.3 Å and spring constant scaling factor cut-off of 1 Å for the calculation of the elastic network model. We calculated the first twenty slowest modes for all the CoVs structures. The eigenvectors of these modes represent the global motions and the constrained residues help in identifying critical regions such as hinge-bending regions and thereby giving an idea of domain motions around these regions. We plotted the second slowest mode in different conditions which showed a significant difference in motions.

#### 4.2.3. Rigidity Analysis

We use the program Floppy Inclusion and Rigid Substructure Topography (FIRST) [80]. In this approach, we first create a constraint network, where the spike protein is modeled in terms of nodes (atoms) and edges (covalent bonds, hydrogen bonds, etc.). A hydrogen bond cutoff energy value is selected where all bonds weaker than this cutoff are ignored. The strength of hydrogen bonds is calculated using a Mayo energy potential. The spike protein network is next decomposed into rigid clusters and flexible regions. The value of energy strength was selected in such way that the bonds strength below this cut off were ignored. First then applies the rigorous mathematical theory of rigid and flexible molecular structure and pebble game algorithm calculates the conformational degrees of freedom rapidly decompose a protein into rigid clusters and flexible region.

#### 4.2.4. Constrained Geometric Simulations for Conformational Sampling

To further probe the dynamical features of the whole spike protein and the impact of hACE2 binding, we have applied the CGS by the Framework Rigidity Optimized Dynamics Algorithm New (FRODAN) [74,75]. This method utilizes a coarse-grained molecular mechanics potential based on rigidity theory to explore receptor conformations well outside the starting structure. FRODAN can be regarded as a low computational complexity alternative to MD simulations which can sample wide regions of high dimensional conformational space. We run FRODAN in the non-targeted mode, generating 30,000 candidate structures for each case. The temperature impact was evaluated by running simulations at different hydrogen bond energy cut-offs. The considered range of the energy cut-offs was from  $-1.0$  to  $-3.0$  kcal/mol, where the more negative cut-off values correspond to higher temperature. During each individual run the cut-off value was kept constant. Before running simulations, hydrogen bonds to each considered structure were added using MolProbity server [92].

**Supplementary Materials:** The following supporting information can be downloaded online, Figure S1: F 6m17 complex. Purple ribbons—B(0)AT1 neutral amino acid transporter, orange ribbons—hACE2, cyan ribbons—surface spike of SARS-CoV-2, blue and red structures are parts of the spike and hACE2 used for 3DRISM calculations, Figure S2: 2ajf complex. Orange ribbons—hACE2, cyan ribbons—spike of SARS-CoV, blue and red structures are parts of the spike and hACE2 used for 3DRISM calculations, Figure S3: 6vyb complex. Cyan ribbons—spike of SARS-CoV-2, blue structure is a part of spike used for 3DRISM calculations, Figure S4: Water distribution around RBD of SARS-CoV-2. (A) Differences in distributions of water oxygen (blue) and water hydrogens (red) between the CoV-2/hACE2 and CoV-2. (B) The pmf of RBD-O and RBD-H for the CoV-2 and CoV-2/hACE2 complex. The RBD is indicated by blue ribbons, the RBD of CoV-2 is shown as background for the differences, Figure S5: Rigidity analysis of protein structures. Individual colours indicate presence of large rigid clusters (blue being the largest rigid cluster), while black colour represents flexible regions, Table S1: Interface residual interactions between CoV-RBD and hACE2, Table S2: Interface interactions between RBD and hACE2, Table S3: 3DRISM parameters of calculated complexes, Table S4: Average RMSF values of the whole spike protein SARS-CoV-1 and SARSCoV-2 (top table). The corresponding values for the hACE2-RBD complex (bottom table). Highest value among SARS-CoV-1 and SARS-CoV-2 is indicated in bold, Table S5: Summary of rigidity analysis.

**Author Contributions:** A.C. and A.S. have conceived the study. N.K., S.B. and A.C. have conducted computational experiments and analytical analysis. A.S., K.T. and A.T. carried out rigidity theory and Monte Carlo simulation analysis based on PDB files and have written the corresponding part. G.N.C., M.V.F. and S.E.K. have formulated the problem of hydration for COVID-hACE2 complex, performed the initial calculations and written the corresponding section of the paper. M.V. helped with theoretical SDFT analysis. S.E.K. has performed all 3D-RISM calculations and a visualization of the results. M.V.F. has conducted a formal analysis and discussion of 3D-RISM results and written the corresponding section of the paper. G.N.C., A.S. and A.C. have reviewed and written the article. All authors have read and agreed to the published version of the manuscript.

**Funding:** The results of G.N.C., M.V.F. and S.E.K. presented in Sections 2.2 and 4.1 were obtained within the Russian Science Foundation (project No. 22-23-00184, <https://rscf.ru/en/project/22-23-00184/>, accessed on 15 December 2021).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** Our thanks to our lab members, Sikkim University for their support while completing a part of computational study in this work.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yao, H.; Song, Y.; Chen, Y.; Wu, N.; Xu, J.; Sun, C.; Zhang, J.; Weng, T.; Zhang, Z.; Wu, Z.; et al. Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **2020**, *183*, 730–738.e13. [[CrossRef](#)] [[PubMed](#)]
2. Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 11727–11734. [[CrossRef](#)] [[PubMed](#)]
3. Qing, E.; Gallagher, T. SARS Coronavirus Redux. *Trends Immunol.* **2020**, *41*, 271–273. [[CrossRef](#)] [[PubMed](#)]
4. Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444–1448. [[CrossRef](#)] [[PubMed](#)]
5. Li, F. Structure, function, and evolution of coronavirus spike proteins. *Annu. Rev. Virol.* **2016**, *3*, 237–261. [[CrossRef](#)] [[PubMed](#)]
6. Cohen, A.J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289–320. [[CrossRef](#)]
7. Jones, R.O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **2015**, *87*, 897–923. [[CrossRef](#)]
8. Hansen, J.P.; McDonald, I.R. *Theory of Simple Liquids: With Applications to Soft Matter*; Elsevier Academic Press: Amsterdam, The Netherlands, 2013.
9. Ding, L.; Levesque, M.; Borgis, D.; Belloni, L. Efficient molecular density functional theory using generalized spherical harmonics expansions. *J. Chem. Phys.* **2017**, *147*, 094107. [[CrossRef](#)]
10. Ishizuka, R.; Yoshida, N. Application of efficient algorithm for solving six-dimensional molecular Ornstein-Zernike equation. *J. Chem. Phys.* **2012**, *136*, 114106. [[CrossRef](#)]
11. Chandler, D.; Andersen, H.C. Optimized Cluster Expansions for Classical Fluids. II. Theory of Molecular Liquids. *J. Chem. Phys.* **1972**, *57*, 1930–1937. [[CrossRef](#)]
12. Chandler, D.; McCoy, J.D.; Singer, S.J. Density Functional Theory of Nonuniform Polyatomic Systems. 1. General Formulation. *J. Chem. Phys.* **1986**, *85*, 5971–5976. [[CrossRef](#)]
13. Chandler, D.; McCoy, J.D.; Singer, S.J. Density Functional Theory of Nonuniform Polyatomic Systems. 2. Rational Closures for Integral-Equations. *J. Chem. Phys.* **1986**, *85*, 5977–5982. [[CrossRef](#)]
14. Seok, C.; Oxtoby, D.W. Nucleation in n-alkanes: A density-functional approach. *J. Chem. Phys.* **1998**, *109*, 7982–7990. [[CrossRef](#)]
15. Sumi, T.; Sekino, H. An interaction site model integral equation study of molecular fluids explicitly considering the molecular orientation. *J. Chem. Phys.* **2006**, *125*, 034509. [[CrossRef](#)] [[PubMed](#)]
16. Lischner, J.; Arias, T.A. Kohn-Sham-Like Approach toward a Classical Density-Functional Theory of Inhomogeneous Polar Molecular Liquids: An Application to Liquid Hydrogen Chloride. *Phys. Rev. Lett.* **2008**, *101*, 216401. [[CrossRef](#)]
17. Lischner, J.; Arias, T.A. Classical Density-Functional Theory of Inhomogeneous Water Including Explicit Molecular Structure and Nonlinear Dielectric Response. *J. Phys. Chem. B* **2010**, *114*, 1946–1953. [[CrossRef](#)]
18. Chong, S.H.; Ham, S. Aqueous interaction site integral-equation theory that exactly takes into account intramolecular correlations. *J. Chem. Phys.* **2012**, *137*, 154101. [[CrossRef](#)]
19. Liu, Y.; Zhao, S.; Wu, J. A Site Density Functional Theory for Water: Application to Solvation of Amino Acid Side Chains. *J. Chem. Theory Comput.* **2013**, *9*, 1896–1908. [[CrossRef](#)]
20. Sundararaman, R.; Arias, T. Efficient classical density-functional theories of rigid-molecular fluids and a simplified free energy functional for liquid water. *Comput. Phys. Commun.* **2014**, *185*, 818–825. [[CrossRef](#)]

21. Sheng, S.; Miller, M.; Wu, J. Molecular Theory of Hydration at Different Temperatures. *J. Phys. Chem. B* **2017**, *121*, 6898–6908. [[CrossRef](#)]
22. Chuev, G.N.; Fedotova, M.V.; Valiev, M. Chemical bond effects in classical site density functional theory of inhomogeneous molecular liquids. *J. Chem. Phys.* **2020**, *152*, 041101. [[CrossRef](#)] [[PubMed](#)]
23. Chuev, G.N.; Fedotova, M.V.; Valiev, M. Renormalized site density functional theory. *J. Stat. Mech.* **2021**, *2021*, 033205. [[CrossRef](#)]
24. Chuev, G.N.; Fedotova, M.V.; Valiev, M. Renormalized site density functional theory for models of ion hydration. *J. Chem. Phys.* **2021**, *155*, 064501. [[CrossRef](#)] [[PubMed](#)]
25. Beglov, D.; Roux, B. An integral equation to describe the solvation of polar molecules in liquid water. *J. Phys. Chem. B* **1997**, *101*, 7821–7826. [[CrossRef](#)]
26. Du, Q.H.; Beglov, D.; Roux, B. Solvation free energy of polar and nonpolar molecules in water: An extended interaction site integral equation theory in three dimensions. *J. Phys. Chem. B* **2000**, *104*, 796–805. [[CrossRef](#)]
27. Howard, J.J.; Perkyins, J.S.; Choudhury, N.; Pettitt, B.M. Integral Equation Study of the Hydrophobic Interaction between Graphene Plates. *J. Chem. Theory Comput.* **2008**, *4*, 1928–1939. [[CrossRef](#)]
28. Kast, S.M. Combinations of simulation and integral equation theory. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5087–5092. [[CrossRef](#)]
29. Chuev, G.N.; Fedorov, M.V. Reference interaction site model study of self-aggregating cyanine dyes. *J. Chem. Phys.* **2009**, *131*, 074503. [[CrossRef](#)]
30. Genheden, S.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Ryde, U. An MM/3D-RISM Approach for Ligand Binding Affinities. *J. Phys. Chem. B* **2010**, *114*, 8505–8516. [[CrossRef](#)]
31. Luchko, T.; Gusarov, S.; Roe, D.R.; Simmerling, C.; Case, D.A.; Tuszynski, J.; Kovalenko, A. Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber. *J. Chem. Theory Comput.* **2010**, *6*, 607–624. [[CrossRef](#)]
32. Perkyins, J.S.; Lynch, G.C.; Howard, J.J.; Pettitt, B.M. Protein solvation from theory and simulation: Exact treatment of Coulomb interactions in three-dimensional theories. *J. Chem. Phys.* **2010**, *132*, 064106. [[CrossRef](#)] [[PubMed](#)]
33. Palmer, D.S.; Chuev, G.N.; Ratkova, E.L.; Fedorov, M.V. In Silico Screening of Bioactive and Biomimetic Solutes Using Molecular Integral Equation Theory. *Curr. Pharm. Des.* **2011**, *17*, 1695–1708. [[CrossRef](#)] [[PubMed](#)]
34. Sindhikara, D.J.; Yoshida, N.; Hirata, F. Placevent: An algorithm for prediction of explicit solvent atom distribution—Application to HIV-1 protease and F-ATP synthase. *J. Comput. Chem.* **2012**, *33*, 1536–1543. [[CrossRef](#)] [[PubMed](#)]
35. Gusarov, S.; Pujari, B.S.; Kovalenko, A. Efficient treatment of solvation shells in 3D molecular theory of solvation. *J. Comput. Chem.* **2012**, *33*, 1478–1494. [[CrossRef](#)] [[PubMed](#)]
36. Fedotova, M.V.; Kruchinin, S.E. 1D-RISM study of glycine zwitterion hydration and ion-molecular complex formation in aqueous NaCl solutions. *J. Mol. Liq.* **2012**, *169*, 1–7. [[CrossRef](#)]
37. Fedotova, M.V.; Kruchinin, S.E. Hydration of para-aminobenzoic acid (PABA) and its anion—The view from statistical mechanics. *J. Mol. Liq.* **2013**, *186*, 90–97. [[CrossRef](#)]
38. Fedotova, M.V.; Kruchinin, S.E. Ion-binding of glycine zwitterion with inorganic ions in biologically relevant aqueous electrolyte solutions. *Biophys. Chem.* **2014**, *190*, 25–31. [[CrossRef](#)]
39. Fedotova, M.V.; Dmitrieva, O. Ion-selective interactions of biologically relevant inorganic ions with alanine zwitterion: A 3D-RISM study. *Amino Acids* **2015**, *47*, 1015–1023. [[CrossRef](#)]
40. Fedotova, M.V.; Dmitrieva, O.A. Characterization of selective binding of biologically relevant inorganic ions with the proline zwitterion by 3D-RISM theory. *New J. Chem.* **2015**, *39*, 8594–8601. [[CrossRef](#)]
41. Eiberweiser, A.; Nazet, A.; Kruchinin, S.E.; Fedotova, M.V.; Buchner, R. Hydration and ion binding of the osmolyte ectoine. *J. Phys. Chem. B* **2015**, *119*, 15203–15211. [[CrossRef](#)]
42. Fedotova, M.V.; Dmitrieva, O. Proline hydration at low temperatures: Its role in the protection of cell from freeze-induced stress. *Amino Acids* **2016**, *48*, 1685–1694. [[CrossRef](#)] [[PubMed](#)]
43. Fedotova, M.V.; Kruchinin, S.E.; Chuev, G.N. Local ion hydration structure in aqueous imidazolium-based ionic liquids: The effects of concentration and anion nature. *J. Mol. Liq.* **2017**, *247*, 100–108. [[CrossRef](#)]
44. Fedotova, M.V.; Kruchinin, S.E.; Chuev, G.N. Hydration structure of osmolyte TMAO: Concentration/pressure-induced response. *New J. Chem.* **2017**, *41*, 1219–1228. [[CrossRef](#)]
45. Fedotova, M.V.; Kruchinin, S.E. Hydration and ion-binding of glycine betaine: How they may be involved into protection of proteins under abiotic stresses. *J. Mol. Liq.* **2017**, *244*, 489–498. [[CrossRef](#)]
46. Güssregen, S.; Matter, H.; Hessler, G.; Lionta, E.; Heil, J.; Kast, S.M. Thermodynamic Characterization of Hydration Sites from Integral Equation-Derived Free Energy Densities: Application to Protein Binding Sites and Ligand Series. *J. Chem. Inf. Model.* **2017**, *57*, 1652–1666. [[CrossRef](#)]
47. Dmitrieva, O.A.; Fedotova, M.V.; Buchner, R. Evidence for cooperative Na<sup>+</sup> and Cl<sup>−</sup> binding by strongly hydrated L-proline. *Phys. Chem. Chem. Phys.* **2017**, *19*, 20474–20483. [[CrossRef](#)]
48. Yoshida, N. Role of Solvation in Drug Design as Revealed by the Statistical Mechanics Integral Equation Theory of Liquids. *J. Chem. Inf. Model.* **2017**, *57*, 2646–2656. [[CrossRef](#)]
49. Fedotova, M.V. Compatible osmolytes—Bioprotectants: Is there a common link between their hydration and their protective action under abiotic stresses? *J. Mol. Liq.* **2019**, *292*, 111339. [[CrossRef](#)]
50. Fedotova, M.V.; Kruchinin, S.E.; Chuev, G.N. Features of local ordering of biocompatible ionic liquids: The case of choline-based amino acid ionic liquids. *J. Mol. Liq.* **2019**, *296*, 112081. [[CrossRef](#)]

51. Fedotova, M.V.; Kruchinin, S.E.; Chuev, G.N. Molecular insight on ion hydration and association in aqueous choline chloride solutions. *J. Mol. Liq.* **2020**, *313*, 113563. [[CrossRef](#)]
52. Fedotova, M.V.; Kruchinin, S.E.; Chuev, G.N. Hydration features of the neurotransmitter acetylcholine. *J. Mol. Liq.* **2020**, *304*, 112757. [[CrossRef](#)]
53. Friesen, S.; Fedotova, M.V.; Kruchinin, S.E.; Buchner, R. Hydration and dynamics of l-glutamate ion in aqueous solution. *Phys. Chem. Chem. Phys.* **2021**, *23*, 1590–1600. [[CrossRef](#)] [[PubMed](#)]
54. Kruchinin, S.E.; Fedotova, M.V. Ion Pairing of the Neurotransmitters Acetylcholine and Glutamate in Aqueous Solutions. *J. Phys. Chem. B* **2021**, *125*, 11219–11231. [[CrossRef](#)]
55. Olson, B.; Cruz, A.; Chen, L.; Ghattas, M.; Ji, Y.; Huang, K.; Ayoub, S.; Luchko, T.; McKay, D.J.; Kurtzman, T. An online repository of solvation thermodynamic and structural maps of SARS-CoV-2 targets. *J. Comput. Aid. Mol. Des.* **2020**, *34*, 1219–1228. [[CrossRef](#)] [[PubMed](#)]
56. Kobryn, A.E.; Maruyama, Y.; Velázquez-Martínez, C.A.; Yoshida, N.; Gusarov, S. Modeling the interaction of SARS-CoV-2 binding to the ACE2 receptor via molecular theory of solvation. *New J. Chem.* **2021**, *45*, 15448–15457. [[CrossRef](#)]
57. Kinoshita, M.; Okamoto, Y.; Hirata, F. First-principle determination of peptide conformations in solvents: Combination of Monte Carlo simulated annealing and RISM theory. *J. Am. Chem. Soc.* **1998**, *120*, 1855–1863. [[CrossRef](#)]
58. Lavecchia, A.; Di Giovanni, C. Virtual Screening Strategies in Drug Discovery: A Critical Review. *Curr. Med. Chem.* **2013**, *20*, 2839–2860. [[CrossRef](#)]
59. Wang, E.; Weng, G.; Sun, H.; Du, H.; Zhu, F.; Chen, F.; Wang, Z.; Hou, T. Assessing the performance of the MM/PBSA and MM/GBSA methods. 10. Impacts of enhanced sampling and variable dielectric model on protein–protein Interactions. *Phys. Chem. Chem. Phys.* **2019**, *21*, 18958–18969. [[CrossRef](#)]
60. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.L.; Abiona, O.; Graham, B.S.; McLellan, J.S. Cryo-EM structure of the 2019-nCoV-1 spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263. [[CrossRef](#)]
61. Kirchdoerfer, R.; Wang, N.; Pallesen, J.; Wrapp, D.; Turner, H.; Cottrell, C.; Corbett, K.; Graham, B.; McLellan, J.; Ward, A. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci. Rep.* **2018**, *8*, 15701–15701. [[CrossRef](#)]
62. Walls, A.C.; Park, Y.J.; Tortorici, M.A.; Wall, A.; McGuire, A.T.; Veesler, D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **2020**, *181*, 281–292.e6. [[CrossRef](#)] [[PubMed](#)]
63. Li, F.; Li, W.; Farzan, M.; Harrison, S. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **2005**, *309*, 1864–1868. [[CrossRef](#)] [[PubMed](#)]
64. Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221–224. [[CrossRef](#)] [[PubMed](#)]
65. Peters, M.H.; Bastidas, O.; Kokron, D.S.; Henze, C. Static all-atom energetic mappings of the SARS-Cov-2 spike protein and dynamic stability analysis of “Up” versus “Down” protomer states. *PLoS ONE* **2020**, *15*, e0241168. [[CrossRef](#)] [[PubMed](#)]
66. Xue, L.; Rodrigues, J.; Kastiris, P.; Amjj, B.; Vangone, A. PRODIGY: A web-server for predicting the binding affinity in protein-protein complexes. *Bioinformatics* **2016**, *32*, 3676–3678. [[CrossRef](#)]
67. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215–220. [[CrossRef](#)]
68. Peng, C.; Zhu, Z.; Shi, Y.; Wang, X.; Mu, K.; Yang, Y.; Zhang, X.; Xu, Z.; Zhu, W. Computational Insights into the Conformational Accessibility and Binding Strength of SARS-CoV-2 Spike Protein to Human Angiotensin-Converting Enzyme 2. *J. Phys. Chem. Lett.* **2020**, *11*, 10482–10488. [[CrossRef](#)]
69. Bai, C.; Warshel, A. Critical Differences between the Binding Features of the Spike Proteins of SARS-CoV-2 and SARS-CoV. *J. Phys. Chem. B* **2020**, *124*, 5907–5912. [[CrossRef](#)]
70. Han, B.; Liu, Y.; Ginzinger, S.W.; Wishart, D. SHIFX2: Significantly improved protein chemical shift prediction. *J. Biomol. NMR* **2011**, *50*, 43–57. [[CrossRef](#)]
71. Malik, A.; Prahlad, D.; Kulkarni, N.; Kayal, A. Interfacial Water Molecules Make RBD of SPIKE Protein and Human ACE2 to Stick Together. *bioRxiv*. **2020**. [[CrossRef](#)]
72. Atilgan, A.R.; Durell, S.R.; Jernigan, R.L.; Demirel, M.C.; Keskin, O.; Bahar, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **2001**, *80*, 505–515. [[CrossRef](#)]
73. López-Blanco, J.R.; Aliaga, J.I.; Quintana-Ortí, E.S.; Chacón, P. iMODS: Internal coordinates normal mode analysis server. *Nucleic Acids Res.* **2014**, *42*, W271–W276. [[CrossRef](#)] [[PubMed](#)]
74. Farrell, D.W.; Speranskiy, K.; Thorpe, M.F. Generating stereochemically acceptable protein pathways. *Proteins* **2010**, *78*, 2908–2921. [[CrossRef](#)] [[PubMed](#)]
75. Zhu, S.; Shala, A.; Bezginov, A.; Sljoka, A.; Audette, G.; Wilson, D. Hyperphosphorylation of intrinsically disordered tau protein induces an amyloidogenic shift in its conformational ensemble. *PLoS ONE* **2015**, *10*, e0120416. [[CrossRef](#)]
76. Zimmerman, M.I.; Porter, J.R.; Ward, M.D.; Singh, S.; Vithani, N.; Meller, A.; Mallimadugula, U.L.; Kuhn, C.E.; Borowsky, J.H.; Bowman, G.R. SARS-CoV-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome. *Nat. Chem.* **2021**, *13*, 651–659. [[CrossRef](#)]
77. Yurkovetskiy, L.; Wang, X.; Pascal, K.E.; Tomkins-Tinch, C.; Nyalile, T.P.; Wang, Y.; Baum, A.; Diehl, W.E.; Dauphin, A.; Carbone, C.; et al. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **2020**, *183*, 739–751.e8. [[CrossRef](#)]

78. Chan, K.H.; Peiris, J.S.; Lam, S.Y.; Poon, L.L.; Yuen, K.Y.; Seto, W.H. The effects of temperature and relative humidity on the viability of the sars coronavirus. *Adv. Virol.* **2011**, *2011*, 734690. [[CrossRef](#)]
79. Kumar, S.; Paul, A.; Chatterjee, S.; Putz, S.; Nehra, N.; Wang, D.S.; Nisar, A.; Jennings, C.M.; Parekh, S.H. Effect of ambient temperature on respiratory tract cells exposed to sars-cov-2 viral mimicking nanospheres-an experimental study. *Biointerphases* **2021**, *16*, 011006. [[CrossRef](#)]
80. Jacobs, D.J.; Rader, A.J.; Kuhn, L.A.; Thorpe, M.F. Protein flexibility predictions using graph theory. *Proteins* **2001**, *44*, 150–165. [[CrossRef](#)]
81. Sljoka, A. Structural and Functional Analysis of Proteins Using Rigidity Theory. In *Sublinear Computation Paradigm*; Katoh, N., Higashikawa, Y., Ito, H., Nagao, A., Shibuya, T., Sljoka, A., Tanaka, K., Uno, Y., Eds.; Springer: Singapore, 2022; pp. 337–367.
82. Ou, X.; Liu, Y.; Lei, X.; Li, P.; Mi, D.; Ren, L.; Guo, L.; Guo, R.; Chen, T.; Hu, J.; et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **2020**, *11*, 1620. [[CrossRef](#)]
83. Shang, W.; Yang, Y.; Rao, Y.; Rao, X. The outbreak of SARS-CoV-2 pneumonia calls for viral vaccines. *Npj Vaccines* **2020**, *5*, 18. [[CrossRef](#)] [[PubMed](#)]
84. Pérez-Fuentes, L.; Drummond, C.; Faraudo, J.; Bastos-González, D. Interaction of organic ions with proteins. *Soft Matter* **2017**, *13*, 1120–1131. [[CrossRef](#)] [[PubMed](#)]
85. Oh, M.I.; Consta, S. What factors determine the stability of a weak protein–protein interaction in a charged aqueous droplet? *Phys. Chem. Chem. Phys.* **2017**, *19*, 31965–31981. [[CrossRef](#)] [[PubMed](#)]
86. Valiev, M.; Chuev, G.N. Site density models of inhomogeneous classical molecular liquids. *J. Stat. Mech.* **2018**, *2018*, 093201. [[CrossRef](#)]
87. Kezic, B.; Perera, A. Towards a more accurate reference interaction site model integral equation theory for molecular liquids. *J. Chem. Phys.* **2011**, *135*, 234104. [[CrossRef](#)]
88. Chuev, G.N.; Vyalov, I.; Georgi, N. Extraction of atom-atom bridge and direct correlation functions from molecular simulations: A test for ambient water. *Chem. Phys. Lett.* **2013**, *561–562*, 175–178. [[CrossRef](#)]
89. Chuev, G.N.; Vyalov, I.; Georgi, N. Extraction of site–site bridge functions and effective pair potentials from simulations of polar molecular liquids. *J. Comput. Chem.* **2014**, *35*, 1010–1023. [[CrossRef](#)]
90. Lue, L.; Blankschtein, D. Liquid-state theory of hydrocarbon-water systems: Application to methane, ethane, and propane. *J. Phys. Chem.* **1992**, *96*, 8582–8594. [[CrossRef](#)]
91. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [[CrossRef](#)]
92. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B.; et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **2018**, *27*, 293–315. [[CrossRef](#)]