Quantitative Biology 2021, 9(1): 61–72
https://doi.org/10.15302/J-QB-020-0231

RESEARCH ARTICLE

Computational simulations reveal the binding dynamics between human ACE2 and the receptor binding domain of SARS-CoV-2 spike protein

Cecylia S. Lupala[1], Xuanxuan Li[1,2], Jian Lei[3], Hong Chen[4], Jianxun Qi[5], Haiguang Liu[1,6,*], Xiao-Dong Su[4,*]

[1] Complex Systems Division, Beijing Computational Science Research Center, Beijing 100193, China
[2] Engingeering Physics Department, Tsinghua University, Beijing 100084, China
[3] State Key Laboratory of Biotherapy and Cancer Center, National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, Chengdu 610041, China
[4] School of Life Sciences, State Key Laboratory of Protein and Plant Gene Research and Biomedical Pioneering Innovation Center (BIOPIC), Peking University, Beijing 100871, China
[5] CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China
[6] Physics Department, Beijing Normal University, Beijing 100875, China
* Correspondence: hgliu@csrc.ac.cn, xdsu@pku.edu.cn

Received June 29, 2020; Revised September 16, 2020; Accepted October 27, 2020

***Background*: A novel coronavirus (the SARS-CoV-2) has been identified in January 2020 as the causal pathogen for COVID-19 , a pandemic started near the end of 2019. The Angiotensin converting enzyme 2 protein (ACE2) utilized by the SARS-CoV as a receptor was found to facilitate the infection of SARS-CoV-2, initiated by the binding of the spike protein to human ACE2.**

***Methods*: Using homology modeling and molecular dynamics (MD) simulation methods, we report here the detailed structure and dynamics of the ACE2 in complex with the receptor binding domain (RBD) of the SARS-CoV-2 spike protein.**

***Results*: The predicted model is highly consistent with the experimentally determined structures, validating the homology modeling results. Besides the binding interface reported in the crystal structures, novel binding poses are revealed from all-atom MD simulations. The simulation data are used to identify critical residues at the complex interface and provide more details about the interactions between the SARS-CoV-2 RBD and human ACE2.**

***Conclusion*: Simulations reveal that RBD binds to both open and closed state of ACE2. Two human ACE2 mutants and rat ACE2 are modeled to study the mutation effects on RBD binding to ACE2. The simulations show that the N-terminal helix and the K353 are very important for the tight binding of the complex, the mutants are found to alter the binding modes of the CoV2-RBD to ACE2.**

Keywords: SARS-CoV-2; COVID-19; ACE2; mutation; molecular dynamics simulations

Author summary: The first step of SARS-CoV-2 virus infection is the binding of spike protein to a receptor protein called ACE2 on the surface of host cell. The detailed interactions between virus spike protein at the receptor binding domain (RBD) and ACE2 were computationally modelled and the binding stability was assessed using molecular dynamics simulations. Using the same method, the study reveals that the mutations mimicking rat ACE2 protein lead to reduced binding to the RBD of spike protein. Simulations reveals persistent presence of water molecules at the interface, indicating their roles in the binding between ACE2 and spike protein.

INTRODUCTION

The outbreak of a new type of severe pneumonia COVID-19 started in December 2019 [1] has been going on worldwide, and caused over 920,000 fatalities, infected more than 28 million individuals globally. One urgent desire in coping with this global crisis is to develop drugs that can effectively treat the diseases caused by the novel coronavirus, the SARS-CoV-2 (also known as 2019-nCoV) [2]. According to the genome comparative studies, the SARS-CoV-2 belongs to the genus beta-coronavirus, with nucleotide sequence identity of about 96% compared to the closest bat coronavirus RaTG13, approximately 89% compared to two other bat SARS-like viruses (Bat-SL-CoVZC45 & Bat-SL-CoVZXC21), and 79% compared to the SARS-CoV [3,4]. Furthermore, the SARS-CoV-2 spike protein has a protein sequence identity of 73% for the receptor binding domain (RBD) with the SARS-CoV. The SARS-CoV and SARS-CoV-2 both utilize the human Angiotensin converting enzyme 2 protein (hACE2) to initiate the spike protein binding and facilitate the fusion to host cells [5–9]. The 193-residue RBD of the SARS-CoV spike protein has been found to be sufficient to bind hACE2 [6]. Based on these facts, the RBD of SARS-CoV-2 becomes an important protein target for drug development to treat the COVID-19. When this study was started, neither the crystal structure of the SARS-CoV-2 spike protein nor the RBD segment were determined, so the homology modeling approach was applied to construct the structure of the SARS-CoV-2 spike RBD in complex with the hACE2 binding domain (denoted as CoV2-RBD/hACE2 in the following). Similar approach has been reported to predict the complex structure and estimate the binding energies [3]. Because of the high sequence similarity between the CoV2-RBD and SARS-RBD, the predicted structure was found to be highly consistent with the resolved crystal structures afterwards [10,11]. Both predicted and crystal structures of CoV2-RBD/hACE2 were subjected to all-atom molecular dynamics (MD) simulations to study the binding interactions.

On the basis of crystal structure and the predicted structure of the CoV2-RBD/ACE2 complex that provide important information about the binding interactions at the molecular interfaces, MD simulations extends the knowledge to a dynamical regime in a fully solvated environment. In this study, key ACE2 residues is investigated by simulating the complexes with ACE2 mutants, in which partial dissociations from the ACE2 are observed within 500 ns simulations. The results showed that the wild type CoV2-RBD/ACE2 complex is very stable in multiple 500-ns simulation trajectories, with a well-defined binding interface. On the other hand, the mutations on the helix-1 or K353 of the ACE2 alters the complex binding, revealing new binding poses that lead to a reduction in residue contacts compared to those in the wild type system. The analysis of the interaction energy showed that the binding is enhanced by adjusting conformations to form more favorable interactions as simulations progress, consistent with the increased hydrogen bonding interactions. Furthermore, the binding free energy analysis shows that rat ACE2 has reduced binding affinity to the RBD, compared to hACE2 and mutants that partially mimic rACE2. The dynamics information obtained by this study is useful in understanding SARS-CoV-2 host interaction and for designing inhibitors to block CoV2-RBD binding.

RESULTS

The homology structure of the CoV2-RBD/ACE2 is compared to both the SARS-RBD/ACE2 crystal structure and the crystal structure of the CoV2-RBD/ACE2. The results indicate that the homology model is highly accurate, especially at the binding interface. The MD simulations further refine the side chain orientations to improve the model quality. The simulation data confirmed the stable binding between the CoV2-RBD and the ACE2, in spite of the conformational changes of the ACE2. The relative movement between the CoV2-RBD and the ACE2 mainly exhibits as a swinging motion pivoted at the binding interface. Simulations also reveal the roles of water molecules in the binding of the RBD to the ACE2 receptor. The MD simulations of complex with ACE2 mutants suggest that mutation to the ACE2 helix-1 and the K353 can alter the binding pose and binding affinity.

Homology modeling and comparison to the SARS-RBD/ACE2 complex

The predicted CoV2-RBD/ACE2 complex structure is very similar to the SARS-RBD/ACE2, as shown in Fig. 1. The RBD domain has a root-mean-square-deviation (RMSD) of 0.99 Å for the aligned residues (1.53 Å for all 174 residue pairs), indicating that the homology model of the CoV2-RBD is in good agreement with the SARS-RBD. For ACE2 residues near the binding interface (within 4.0 Å of the RBD), the RMSD is smaller than 0.43 Å compared to the SARS-RBD/ACE2 complex. The superposed structures reveal that the RBD/ACE2 interfaces are almost identical in two complexes (Fig. 1C).

In a retrospective comparison, the homology model is superimposed to a newly resolved crystal structure (PDB ID: 6LZG; see Fig. 2D for a detailed comparison at the interface). The comparisons show that the homology

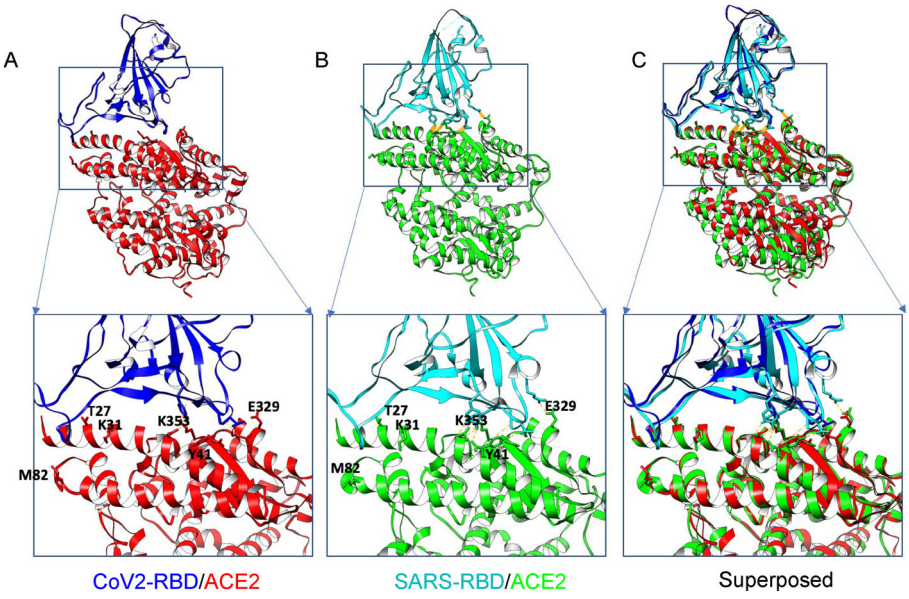

Figure 1. Predicted structure and the binding interface. (A) Predicted model of the CoV2-RBD/hACE2. (B) Crystal structure of the SARS-RBD/hACE2 (PDB: 2AJF). (C) Superposed models. The binding interface is indicated in the boxed region.

model is very accurate, especially at the binding interface. The interfacing residues (defined as the combined set of ACE2 residues within 4.0 Å of RBD and the RBD residues within 4.0 Å of the ACE2) exhibit a difference of 0.43 Å RMSD, which is comparable to the difference between the two independently reported crystal models (an RMSD of 0.25 Å for the comparison of the same residues). The RMSD is about 0.77 Å for residues in an extended region within 10.0 Å of the binding interface. The RBD of the spike protein showed an overall RMSD smaller than 1.5 Å, and the ACE2 with an RMSD of 2.0 Å between the predicted model and the crystal structure.

The conformational changes of the ACE2 and the CoV2-RBD

The comparison with respect to the starting structure revealed that MD simulations generated conformation ensembles near the starting model. The structures sampled in the MD simulations are distributed from 2.5 Å to 5.0 Å of the starting model (Fig. 3A) measured with RMSD. The largest contribution to the conformational changes in CoV2-RBD is originated from loop regions, indicated by the peaks shown in the fluctuation plots (Fig. 3B). The residues with well-defined secondary structures (mainly β-strands) exhibit very small conformational changes (the RMSD is about 1.0 Å including side chain atoms), while loop regions have RMSD values between 3.7 Å to 5.4 Å. This is consistent with the residue fluctuations measured with the RMSF (root-mean-square-fluctuation) for the RBD (Fig. 3B). For the ACE2, the major conformational change is the opening/closing of the enzymatic active site, which is remote from the RBD binding interface (see Supplementary video 1).

Based on the structural similarity, a clustering analysis is carried out to identify the most populated structures, from which the representative model for each trajectory is identified (Fig. 2). The CoV2-RBD/ACE2 interfaces are highlighted using blue/red colors in stick representations. The RMSD of structures in simulations are calculated with respect to these representative models, in order to assess the model convergence and coverage of the simulated conformations (Fig. 3A, right panels). The RMSD values indicate that a large portion of structures sampled by MD simulations are similar to the representative model (Fig. 3A, for trajectory 1, see Supplementary Fig. S1 for the other trajectories). Furthermore, as the simulations progress, the conformations converge to the representative model obtained from this trajectory. The clustering analysis and the RMSD with respective to the most populated structure indicate that MD simulation yielded a new stable structure. The representative structures for the other two trajectories show smaller deviations compared to the initial structure (Fig. 2B and C), suggesting the predicted binding pose is also quite stable. The state of hACE2 was carefully examined by calculating opening angle near the catalytic site of ACE2. In trajectories 1 and 2, the hACE2 remained in its native open state during most of simulation time while the conformation of trajectory 3 changed to closed state (Fig. 2E–G). Interestingly, in the 10 μs trajectory released by the DE Shaw group [12], the ACE2 also showed a

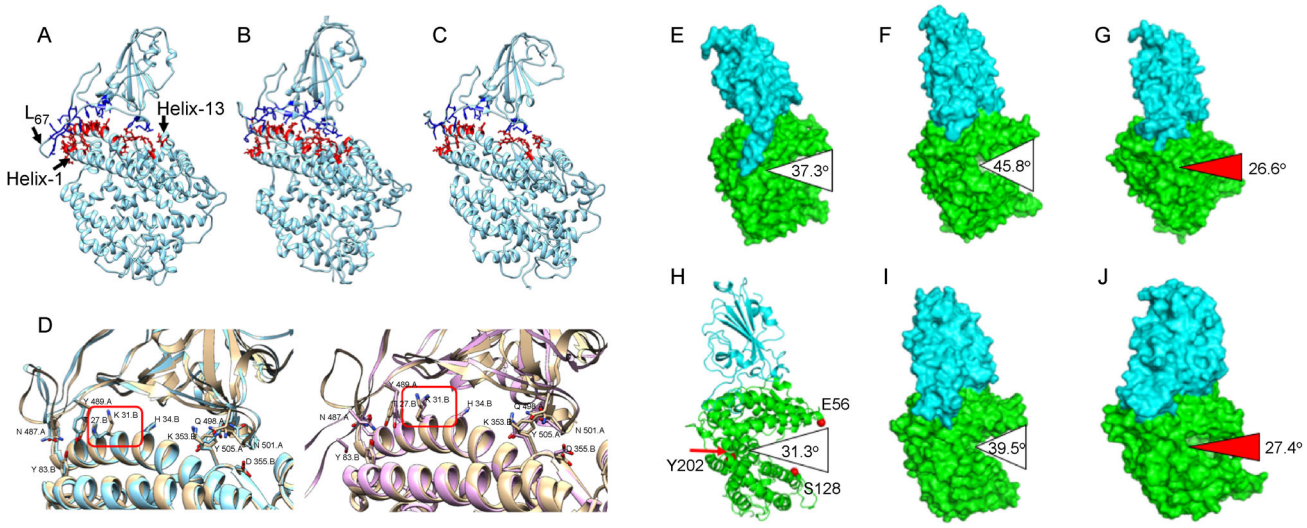

Figure 2. Representative structures extracted from simulation trajectories using clustering analysis. (A) The most populated structure from simulation trajectory # 1, with the loop (L_{67}, S477-G485) between β6 and β7 moved towards the hACE2 helix-1, forming additional contacts. (B, C) The most populated structures in the simulations #2 and #3. The complex structures and the binding interfaces are similar to the predicted model and the crystal structure. (D) The comparison of initial (cyan) and representative (pink) structures to the crystal structure (gold) of CoV2-RBD/hACE2. The side chain of K31 (encircled in the red boxes) was refined to the correct position during the MD simulations (right panel). (E–G) Representative structures from three simulations of CoV2-RBC/hACE2 complex. (H) Crystal structure of the CoV2-RBC/hACE2 complex, with three residues (E56, S128, Y202) defining an opening angle highlighted with spheres colored in red. (I,J) Structures from the DE Shaw trajectories at (I) 500 ns and (J) 10 μs. The opening of the enzymatic sites is indicated by the angles in (E–J).

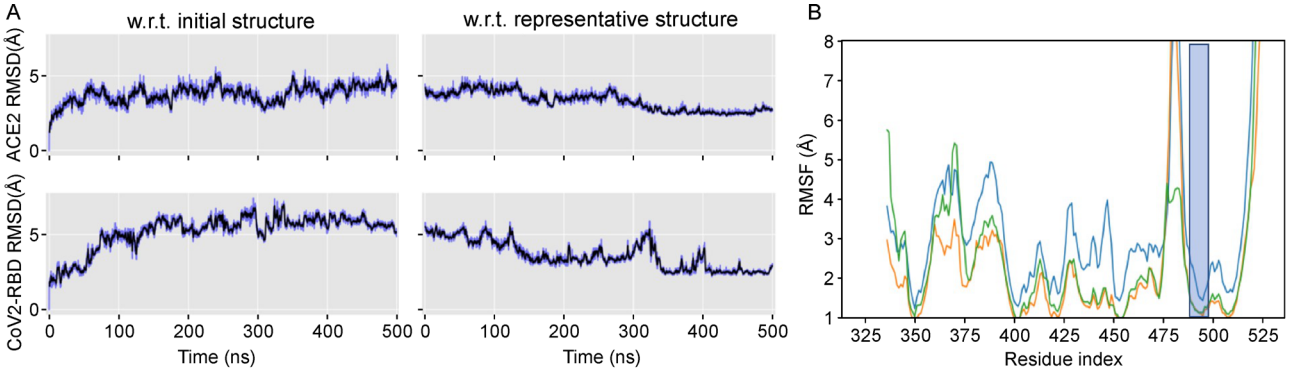

Figure 3. Dynamics analysis of the simulation trajectory. (A) The root-mean-square-deviation (RMSD) of the simulated structures compared to the initial and representative models. (B) The residue level fluctuations of CoV2-RBD in three simulations. The shaded area corresponds to the CoV2-RBD binding region to the hACE2.

shrinkage of open cavity near the catalytic site (Fig. 2J). In all simulated systems, the CoV2-RBD remains bound to hACE2, even though substantial conformational changes occur to the hACE2 and the loop region of the RBD (see Supplementary video 2). This observation implies that the binding of CoV2-RBD to the ACE2 is very stable, similar to the case of the SARS-RBD that binds to both open and closed forms of hACE2 [7,13].

Regardless of the differences in three simulations of CoV2-RBD/hACE2 complex, the binding interface is highly stable, exhibiting very small conformational changes, especially for the interfacing residues of the ACE2 protein. The RMSD for the residues at the RBD binding interface is 0.85 Å (±0.13 Å) on average. Side chain atom positions were refined to form more favorable interactions through simulations (see Fig. 2D). One outstanding example is the K31 side-chain of the ACE2, which points in the wrong orientation in the

predicted structure, was quickly refined to the correct orientation, consistent with the crystal structure (right panel of Fig. 2D).

In terms of collective conformational changes, the CoV2-RBD/hACE2 complex showed two interesting movements: (1) the loop (L_{67}) between β_6 and β_7 (residues between S477 and G485) of CoV2-RBD enhances its interactions with the N-terminal helix (the helix-1) of the hACE2 (Fig. 2A), while it has nearly no direct contacts with the helix-1 of the hACE2 in the predicted and the crystal structures (Fig. 2D, left); (2) the RBD undergoes a tilting motion relative to the hACE2, which can be depicted as a swinging motion with the binding interface as the pivot (see Fig. 4 for an illustration). In both predicted and crystal structures of the CoV2-RBD/hACE2 complex, the L_{67} does not form close contacts with the hACE2. The analysis of the crystal packing reveals that this loop participates in the interaction with another asymmetric unit (see Supplementary Fig. S2) in the unit cell. Interestingly, the simulation data suggest that the L_{67} could form contacts with the hACE2 helix-1. These contacts can potentially enhance the binding, as reflected in the change of interaction energies. In the crystal structure, the C480 and C488 of the RBD are cross-linked via a disulfide bond that reduces the flexibility of the L_{67} region, limiting its access to the hACE2. On the other hand, it has been reported that the binding of SARS-RBD to hACE2 is insensitive to the redox states of the cysteines to a high extend [14]. Based on the simulation results, we hypothesize that the reduced form of C480 and C488 can also exist during the virus invasion to host cells, and the reduced cysteines can potentially enhance the binding to hACE2. In the other two simulation trajectories, the L_{67} remains in conformations similar to that in the crystal structure and the cysteines (C480 and C488) are very close.

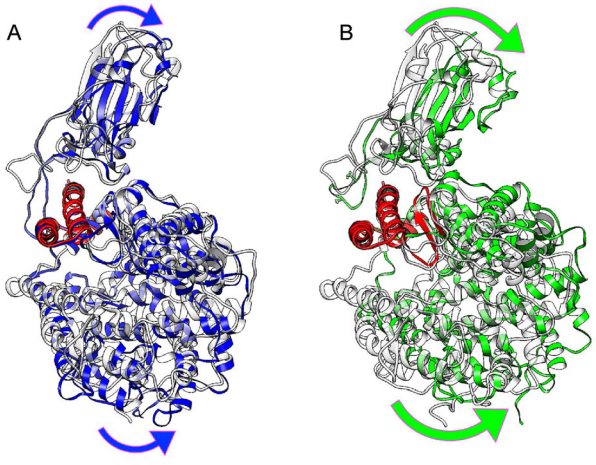

Figure 4. Collective motion of the CoV2-RBD/ACE2 complex. (A) The representative model (blue) from wild type CoV2-RBD/hACE2 simulations superposed to the initial model (silver). The arrows indicate the directions of the collective motion, and the hinge region is colored in red. (B) The same representation for the CoV2-RBD complexed with the hACE2 h1-mut mutant (green), and comparison between the initial model (silver) and the representative model (green) indicates a large tilting movement compared to the wild type system.

Simulations reveal detailed binding interface interactions between the CoV2-RBD and the ACE2

By examining the binding interface of CoV2-RBD and the hACE2, we found the polar and charged residues account for a large fraction, therefore the electrostatic interactions play critical roles for the complex formation. Based on the distances between the two proteins, the key residues at the binding interface are identified and summarized in Table 1 for the three representative structures (see Fig. 2). Majority of these residues are conserved for these three models, except that the model #1 (see Fig. 2A) has additional contacts to the ACE2 from residues in the L_{67} region (highlighted with green color in Table 1). As shown in Fig. 2, the L_{67} remains in the starting position for the other two representative models (Fig. 2B, C).

The hydrogen bonds between the CoV2-RBD and ACE2 are extracted using VMD program with default criteria (D-A distance cutoff = 3.0 Å and D-H-A angle cutoff = 20 degrees, where D,A,H are Donor atom, Acceptor atom, and the Hydrogen atom linked to the Donor atom). The numbers of hydrogen bonds are summarized in Fig. 5 for the simulation trajectory #1 of CoV2-RBD/hACE2 complex. Because of the stringent criteria for hydrogen bonds, the numbers are smaller compared to the reported values in other studies. For a larger distance cutoff of 3.9 Å, there are about 8.3 hydrogen bonds on average (Fig. 5B). The number of hydrogen bonds fluctuates over time, and exhibits an overall increasing trend indicated by the fitted lines (Fig. 5). Similar trends are observed in the other simulations, suggesting that stronger binding were establishing as the simulation progresses (see Supplementary Fig. S3). Moreover, using distance cutoff = 3.0 Å our results are similar with the DE Shaw group simulation trajectory (of 10 μs), showing an increased number of hydrogen bonds (Supplementary Fig. S4A).

It is also noteworthy to point out the important roles of water molecules at the complex interface for CoV2-RBD/ACE2 complex. At any instant time, there are approximately 15 water molecules at the binding interface (Fig. 6). These water molecules can function as bridges by forming hydrogen bonds with the residues from the

Table 1 Contact residues between the CoV2-RBD and the ACE2

Model #1		Model #2		Model #3	
ACE2	CoV2-RBD	ACE2	CoV2-RBD	ACE2	CoV2-RBD
Q24	K417	S19	K417	Q24	K417
T27	Y453	Q24	G446	T27	G446
F28	L455	T27	Y449	F28	Y449
D30	F456	F28	Y453	D30	Y453
K31	Q474	D30	L455	K31	L455
H34	A475	K31	F456	H34	F456
E35	G476	H34	A475	E35	A475
E37	S477	E35	F486	E37	F486
D38	T478	D38	N487	D38	N487
Y41	G485	Y41	Y489	Y41	Y489
Q42	F486	Q42	Q493	Q42	Q493
M82	N487	L45	Y495	M82	Y495
Y83	Y489	M82	G496	Y83	G496
N330	Q493	Y83	Q498	N330	Q498
K353	Y495	N330	T500	K353	T500
G354	G496	K353	N501	G354	N501
D355	Q498	G354	G502	D355	G502
R357	T500	D355	Y505	R357	Y505
	N501	R357			
	G502				
	Y505				

Green color denotes new interaction not observed in crystal structure.

RBD or the ACE2. The dwelling time of water molecules at the interface can be up to a few nanoseconds, revealed by the MD simulations. This observation is also consistent with the crystal structure, which reveals 12 water molecules at the interface (Fig. 6C). Similar number of water molecules (15±3) were observed in the extensive long MD simulation by the DE Shaw Research (Supplementary Figs. S4B and S5). These results emphasize the role of the water molecules, which deserve detailed quantification to understand the interactions between the RBD and the ACE2.

Simulations of the hACE2 mutants to analyze the impacts of Helix-1 and K353 residues

It has been demonstrated that the ACE2 proteins from several mammalian species possess high sequence similarities, yet their binding to the RBD of the SARS-CoV differs significantly. In particular, the binding of SARS-RBD to the rat ACE2 is much weaker as discovered in experiments [15]. Inspired by these information, two mutants of the ACE2 were constructed (see Table 2): (1) *ACE2-mut-h1* by mutating N-terminal helix-1 to that of the rat ACE2; (2) *ACE2-K353H* by mutating K353 to Histidine (the amino acid in wild type rat ACE2). Two 500-ns MD simulations were carried out for each complex system with mutant ACE2. The simulations show that the mutations in *ACE2-mut-h1* reduce the interaction between the CoV2-RBD and the helix-1, and the *ACE2-K353H* shows weaker binding between the CoV2-RBD and the β-hairpin centered at the H353. Using structure clustering analysis, the representative structures are identified for each simulation trajectory (Fig. 7). Although the overall topology is very similar to the wild type complex structure, there are pronounced differences. For the *ACE2-mut-h1*, the CoV2-RBD tilts further away from the ACE12 helix-1 in one simulation trajectory (Fig. 7A); the CoV2-RBD loses its contact with helix-13 (G326 to N330) in one simulation trajectory for the *ACE2-K353H* (Fig. 7A). In the case of wild type ACE2, the K353 is a hydrogen donor, and its mutant H353 cannot form the hydrogen bond with the CoV2-RBD as in the wild type CoV2-RBD/ACE2 complex. The number of contacting residues is significantly smaller in the *ACE2-K353H* mutant system. This is in line with the report that K353 is a critical residue, as its hydrophobic

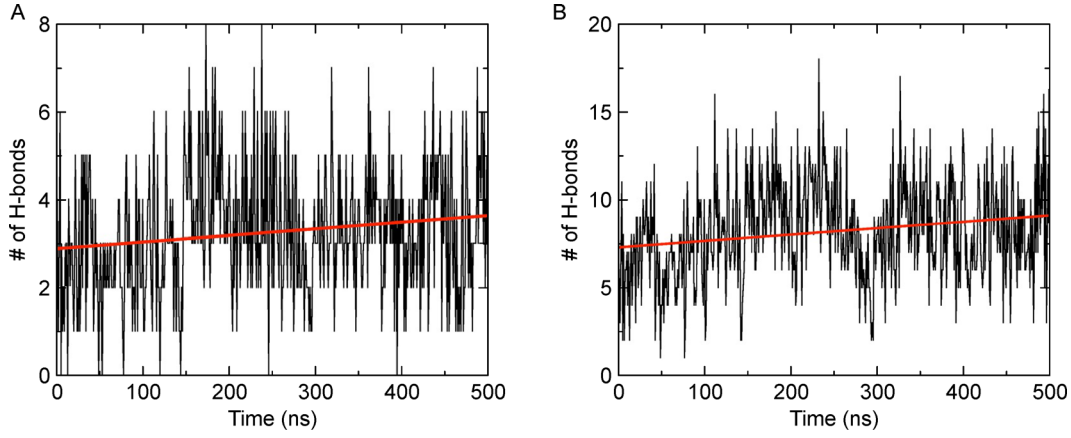

Figure 5. Hydrogen bonds between the CoV2-RBD and the ACE2. (A) The number of hydrogen bonds extracted from the simulation data using Donor-Receptor distance cutoff of 3.0 Å. (B) Same as (A) except that the cutoff distance was relaxed to 3.9 Å. The trend lines were obtained by least square fitting.

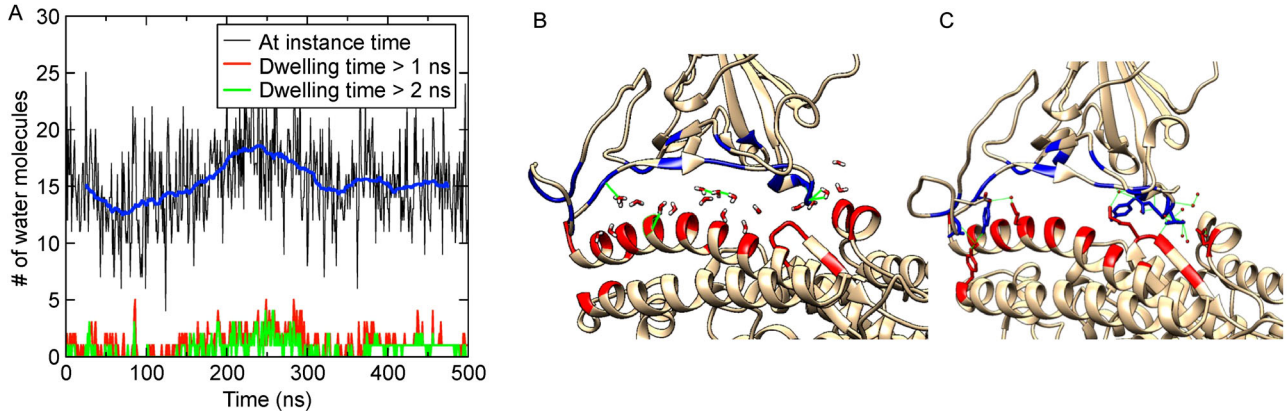

Figure 6. Water molecules at the binding interface. (A) The number of water molecules that bridge the CoV2-RBD and hACE2 found in a simulation trajectory. (B) Water molecules at the interface for the initial structure after equilibration. (C) Water molecules at the interface observed in the crystal structure.

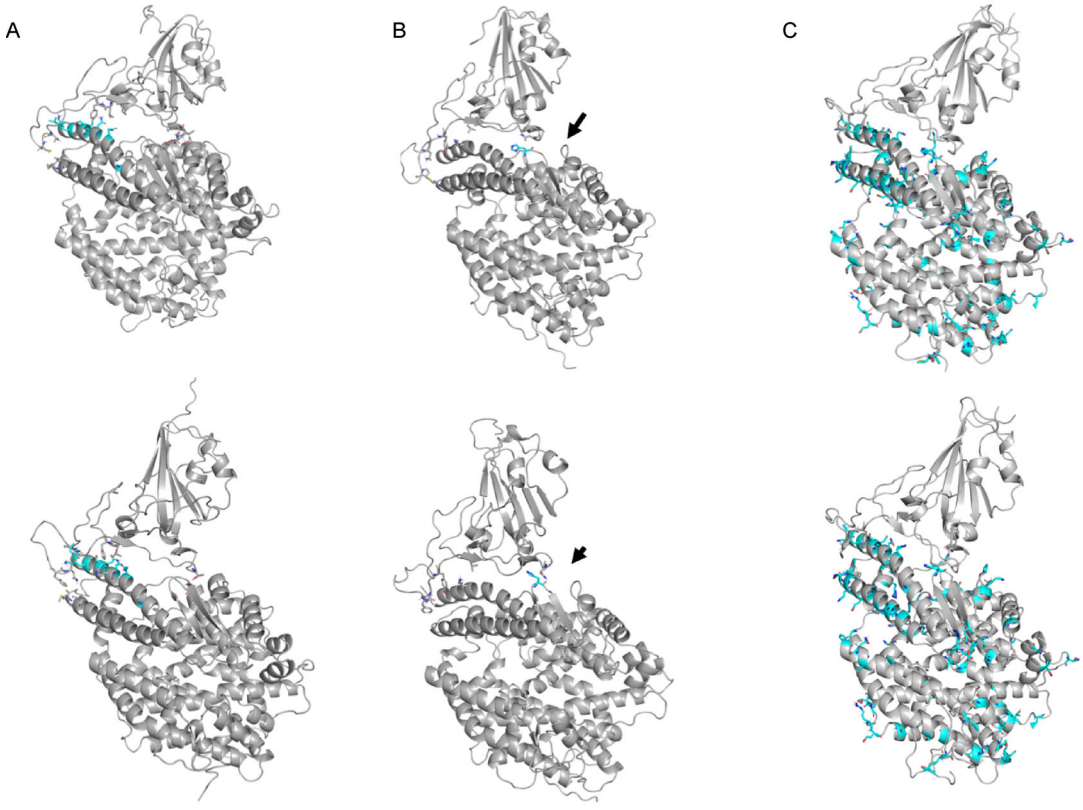

Figure 7. Representative structures for the hACE2 mutants and rACE2 in complex with the CoV2-RBD. (A) *ACE2-mut-h1* from two simulation trajectories. The structure in upper panel of (A) is the same as the green model shown in Fig. 4B, which illustrates a pronounced tilting movement. (B) h*ACE2-K353H,* with hACE2 K353 mutated to histidine. In both structures, the CoV2-RBD lost its contact with the Helix-13 of the ACE2 (residues 498-503, indicated by the black arrows). (C) Representative structures of rACE2 in complex with CoV2-RBD show separations near the Helix-13 of ACE2. The mutation sites are highlighted with cyan sticks for the hACE2-mutants in (A,B). In rACE2, the residues different from hACE2 are highlighted with cyan sticks in (C). Each structure represents the largest cluster from structure clustering analysis of 500 ns MD simulation trajectories.

Table 2 The ACE2 mutants to mimic rat ACE2

Mutant name	Mutations on ACE2	Remark
mut_h1	T20L, Q24K, K26E, T27S, D30N, H34Q, F40S	L20, N30, Q34, and S40 are conserved between rat and mouse
K353H	K353H	H353 is conserved between rat and mouse

neighborhood enables this positively charged residue high selectivity to the RBD [16,17].

The physical interactions between the RBD and the ACE2 are quantified for the simulated structures. We calculated the molecular mechanics energy E_{MM} (van der Waals and electrostatic interactions), binding free energy, and the number of residue contacts (*NC*) between RBD and ACE2 for the structures in the last 100 ns of simulations (Table 3). A contact is defined if two residues from each subunit have at least one pair of non-hydrogen atoms within 4 Å. The average values for these three quantities calculated from structure ensembles sampled in the last 100 ns simulations are reported. The mutation impacts are reflected in these quantitative analyses (Fig. 8A): in one simulation of the CoV2-RBD/h*ACE2-mut_h1* complex, the physical interactions is significantly reduced, likely due to the tilting movement of CoV2-RBD, making it further from the hACE2 helix-1. In the other simulation trajectory for the CoV2-RBD/ACE2-mut_h1 complex, the interaction becomes stronger, reflected on larger *NC* and E_{MM}. However, the binding free energy calculated from these two trajectories reveals reduced binding interactions between hACE2 and RBD (Fig. 8B). It is noted that the physical interactions as described using E_{MM} or *NC* are not equivalent to the binding energies, because the solvation effects are not incorporated. For simulations of the RBD with *ACE2-K353H* mutant (green diamonds), the number of contacts are reduced compared to the wild type system in both trajectories. In one simulation, the contacts between the CoV2-RBD and the Helix-13 of the ACE2 are completely lost (see Fig. 7B), consistent with the least favorable interactions (green diamond at lower right in Fig. 8A). On the other hand, the binding free energy of K353H mutant and RBD are comparable to the case of wild type hACE2. One outstanding information revealed from the quantitative comparison analysis is that the rACE2 exhibits weaker binding to the RBD, based on the assessment of both E_{MM} and binding free energies. These results suggest that mutations on helix-1 or K353 alone may not be sufficient to significantly reduce the RBD binding to ACE2.

DISCUSSIONS AND CONCLUSIONS

The homology modeling of the CoV2-RBD/ACE2 complex yielded highly consistent models compared to crystal structures. All-atom molecular dynamics simulations were carried out to study the dynamic interactions of CoV2-RBD with ACE2. The hACE2 mutants were also constructed to mimic rACE2 to investigate the roles of critical residues. According to the structural analysis, MD simulations improve the structure at the binding interface and strengthen the interactions between RBD and ACE2. The structure of complex interface is highly stable for all simulations of wild type CoV2-RBD/hACE2 complex. In these simulations, the glycosylation modifications were not included. While their significance in both physiological activities of proteins and impact in therapeutic design are well known, their effects on protein conformation and dynamics have been previous explored [18], where Lee *et al.* show that the N-glycosylation does not induce significant changes in protein structures. Furthermore, they concluded glycosylation decreases protein

Table 3 The interactions between ACE2 and CoV2-RBD

System	Trajectory id	VDW[#]	Electrostatic interaction[#]	Solvation[#]	Binding free energy[#]	E_{MM}[#]	Number of contacts
WT (prediction)	1	–339.5	–703.5	483.6	–601.8	–1043.0	33.7
	2	–351.8	–680.6	481.2	–596.6	–1032.4	35.9
	3	–302.2	–673.6	435.2	–578.2	–975.8	29.3
WT (crystal)	1	–326.3	–684.3	577.3	–475.5	–1010.6	33.4
	2	–316.4	–688.8	545.1	–500.2	–1005.2	30.8
K353H	1	–299.2	–783.7	514.1	–608.0	–1082.8	27.8
	2	–271.6	–679.2	426.7	–561.2	–950.8	26.1
mut_h1	1	–396.6	–815.5	748.4	–515.2	–1212.2	40.3
	2	–307.4	–762.6	642.3	–469.4	–1070.0	29.2
rat ACE2	1	–333.9	–641.6	707.2	–313.8	–975.5	33.0
	2	–336.1	–596.1	619.2	–356.0	–932.3	31.7

[#]All energy terms are measured in the unit of kJ mol^{-1}.

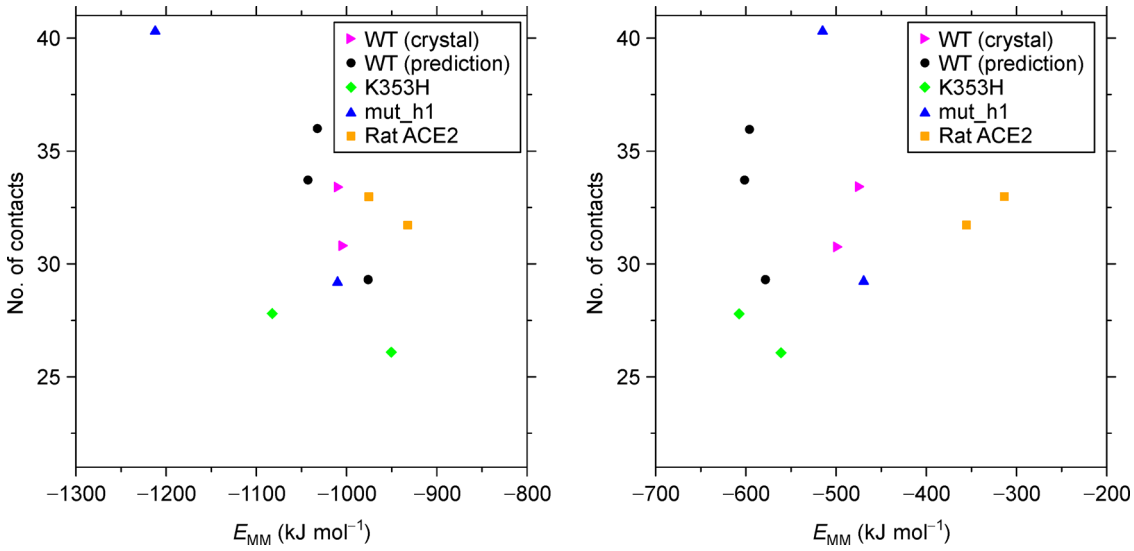

Figure 8. Quantitative comparison of RBD and ACE2 interactions. Three quantities were computed based on the structures obtained from molecular dynamics simulations. (A) The number of contacts and the molecular mechanics interaction energies composed of van der Waals and electrostatic interactions are compared for all systems. (B) The number of contacts and the binding free energies are compared. In the legend, the WT, K353H, mut_h1 are the abbreviations for different constructs of the human ACE2; the rat ACE2 is marked separately.

dynamics which may lead to enhanced protein stability. Our simulation data showed that the structure of COV2-RBD/ACE2 complex is highly stable even in the absence of glycosylation. The representative structures obtained from this work for the wild type COV2-RBD/hACE2 are compared to the 10 μs MD simulation results and the crystal structures. These models possess similar conformations, exhibiting stable binding between RBD and hACE2 (see Supplementary Table S1).

The loop region between β6 and β7 can potentially form more contacts with the ACE2 as observed in multiple simulation trajectories. The simulations results also reveal that the interactions between CoV2-RBD and hACE2 are mediated by water molecules at the interface, stressing the necessity of accounting for the explicit water molecules when quantifying the binding affinity. The detailed analysis shows that the molecular mechanics energy E_{MM} is dominated by the electrostatic interactions (Table 3), making the interfacing water molecules even more important in accurately characterizing the complex binding interactions. This is in line with the finding of a recent study, where Wang *et al.* showed that the hydrogen bonding and hydrophobic interactions enhanced the receptor binding [19]. The amino acids located at the binding interface comprise both hydrophobic and polar residues, revealing a complex interaction pattern (see Table 1). Here, we take a different approach by decoupling the energies between ACE2 and CoV2-RBD into van der Waals and coulomb interactions. As shown in Table 3, the electrostatic interactions constitute a large component to the total molecular interactions.

This study was started with a structure predicted using homology modeling method, which later found to be highly consistent with the crystal structure, demonstrating the potentiality of structure prediction and dynamics simulation in revealing molecular details before the availability of high resolution experimental information. Furthermore, the interactions between CoV2-RBD and the ACE2 mutants mimicking rat ACE2 protein were investigated. The results provide valuable information at the atomic level for the reduced binding affinity in mutant systems. The recent report on the SARS-CoV-2 infection to a dog [20] raise the concerns about the COVID-19 transmissions to pet animals. The approach discussed in this study can be applied to analyze the CoV2-RBD interaction with a broader range of ACE2 variants, including the genetic polymorphism of the ACE2 in human and the ACE2 of other mammals. Preliminary analysis of several ACE2 proteins from mammalian animals using homology modeling and active site prediction methods shows that the binding sites exhibit different features compared to human ACE2. More thorough and deeper analysis can be carried out using the presented approach.

The high quality homology model allowed us to start the simulation analysis before the high resolution experimental data became available. The protein structure prediction community are making efforts to provide high quality models using advanced modeling and prediction methods [21–24]. The presented approach, combining

homology modeling and large scale molecular dynamics simulations can also guide the development of peptide drugs. For example, the peptides derived from human ACE2 proteins have been predicted to bind to RBD of the spike protein, providing potential neutralizing agents [25,26]. At least one peptide has been experimentally verified for tight binding to RBD [25,27]. The binding interactions can be improved by designing peptides with more favorable interactions, such as hydrogen bonds, salt bridges or hydrophobic packing. Based on predicted or experimentally determined structures, molecular dynamics simulations will add more insights to the molecular mechanism and allow better investigation of binding interactions due to mutation or modification to proteins.

MATERIALS AND METHODS

The computer model of the SARS-CoV-2 spike RBD in complex with hACE2

The spike RBD of SARS-CoV-2 (GenBank: MN908947 [2]) comprises Cys336-Gly526 residues according to the sequence homology analysis with SARS-CoV spike RBD (see Supplementary Fig. S6). The predicted three-dimensional structure model of these residues was obtained with the SWISS-MODEL server [28]. This predicted SARS-CoV-2 RBD model was subsequently superimposed into the X-ray structure of SARS-CoV RBD in complex with hACE2 (PDB code:2AJF, Chain D [7]). Finally, the computer model of SARS-CoV-2 RBD with hACE2 (CoV2-RBD/ACE2) was obtained for further simulations and analysis.

Mutant preparation

Based on the analysis of the predicted model, sequence alignment (see Supplementary Fig. S7 for full sequence alignments of ACE2 proteins), and literature survey, two other systems containing mutations in the hACE2 were prepared and subject to MD simulation studies. The mutant construct is based on the fact that rat ACE2 (rACE2) markedly diminishes interactions with SARS spike protein [15], and it was proposed that the rat ACE2 likely has reduced binding affinity to the CoV2-RBD [29]. To investigate the roles of critical residues on the ACE2, we created two mutants of the hACE2 (see Table 2): (1) mutant *mut_h1*, with the hACE2 N-terminal (residue 19–40) mutated to the residues of rACE2; and (2) mutant *K353H*, in which the highly conserved K353 was mutated to histidine (the corresponding amino acid in rat and mouse ACE2 proteins). To focus on the impact of these two binding sites, the rest of the ACE2 were kept to be the same as hACE2. Using the same approach, rACE2 in complex with RBD was constructed for simulations to compare with the cases of hACE2 and mutants.

Molecular dynamics simulation and analysis

The predicted model of CoV2-RBD/hACE2 complex was used as starting models for MD simulations. The spike protein RBD domain is composed of 180 residues (323–502), while the ACE2 protein contains 597 residues (19–615) from the N-terminal domain. The protonation states of histidine were determined at pH 7.0 and in accordance of local hydrogen bonds. In particular, the histidine at position 353 of K353H mutant is in neutral state. The pKa values are estimated using H++ webserver [30], and hydrogens were added using chimera addH program [31]. The simulation parameterization and equilibration were prepared for complex structures including the mutant systems, using the CHARMM-GUI webserver [32]. Each system was solvated in TIP3P water model. Sodium chloride ions were added to neutralize the systems to a salt concentration of 150 mM. Approximately, each system was composed of about 220,000 atoms that were parametrized with the CHARMM36 force field [33].

After energy minimization using the steepest descent algorithm, each system was equilibrated at human body temperature 310.15 K, which was maintained by Nose-Hoover scheme with 1.0 ps coupling constant in the NVT ensemble (constant volume and temperature) for 125.0 ps under periodic boundary conditions with harmonic restraint forces applied to the complex molecules (400 kJ mol^{-1} nm^{-2} on backbone and 40 kJ mol^{-1} nm^{-2} on the side chains) [34,35]. In the subsequent step, the harmonic restraints were removed and the NPT ensembles were simulated at one atmosphere pressure (10^5 Pa) and body temperature. The pressure was maintained by isotropic Parrinello-Rahman barostat [36], with a compressibility of 4.5×10^{-5} bar^{-1} and coupling time constant of 5.0 ps. The simulation trajectories were propagated to 500 nanoseconds using the GROMACS 5.1.2 package [37].

Independent trajectories starting from random velocities based on Maxwell distributions were simulated for wild type CoV2-RBD/hACE2 complex. In all simulations, a time step of 2.0 fs was used and the PME (particle mesh Ewald) [38] was applied for long-range electrostatic interactions beyond 12.0 Å. The van der Waals interactions were evaluated within the distance cutoff of 12.0 Å. Hydrogen atoms were constrained using the LINCS algorithm [39].

The hACE2 mutants and rACE2 in complex with CoV2-RBD were constructed as described previously. Each complex model was simulated in two independent trajectories. Furthermore, based on the crystal structure of the CoV2-RBD/ACE2 complex, two additional simulations were carried out to cross-validate the simulation

results based on the homology model. Each trajectory was propagated to 500 ns by following the same protocol as the wild type CoV2-RBD/hACE2 complex simulations.

Analyses were carried out with tools in GROMACS (rmsd, rmsf, and pairdist) to examine the system properties, including the overall stability, local residue and general structure fluctuations through the simulations. The *g_mmpbsa* program [40] was applied to extract the binding free energy and molecular mechanics energy E_{MM} (Lennard Jones and electrostatic interactions) between ACE2 and the RBD of the spike protein. VMD and Chimera were applied to analyze the hydrogen bonds, molecular binding interface, water distributions, visualization, and rending model images [31,41].

ABBREVIATIONS

ACE2	angiotensin converting enzyme 2 protein
RBD	receptor binding domain
MD	molecular dynamics

SUPPLEMENTARY MATERIALS

The supplementary materials can be found online with this article at https://doi.org/10.15302/J-QB-020-0231.

ACKNOWLEDGEMENTS

The work is supported by Beijing Computational Science Research Center (CSRC) via a director discretionary grant. The authors acknowledge the Beijing Super Cloud Computing Center (BSCC) for providing HPC resources that have contributed to the research. The research is supported by the National Natural Science Foundation (No. U1930402).

COMPLIANCE WITH ETHICS GUIDELINES

The authors Cecylia S. Lupala, Xuanxuan Li, Jian Lei, Hong Chen, Jianxun Qi, Haiguang Liu and Xiao-Dong Su declare that they have no competing interests.

All procedures performed in studies were in accordance with the ethical standards of the institution or practice at which the studies were conducted, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

REFERENCES

1. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., *et al.* (2020) A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med., NEJMoa2001017
2. Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., *et al.* (2020) A new coronavirus associated with human respiratory disease in China. Nature, 579, 265–269
3. Othman, H., Bouslama, Z., Brandenburg, J.T., da Rocha, J., Hamdi, Y., Ghedira, K., Srairi-Abid, N. and Hazelhurst, S. (2020) Interaction of the spike protein RBD from SARS-CoV-2 with ACE2: Similarity with SARS-CoV, hot-spot analysis and effect of the receptor polymorphism. Biochem. Biophys. Res. Commun., 527, 702–708
4. Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., *et al.* (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet, 395, 565–574
5. Li, W., Moore, M. J., Vasilieva, N., Sui, J., Wong, S. K., Berne, M. A., Somasundaran, M., Sullivan, J. L., Luzuriaga, K., Greenough, T. C., *et al.* (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature, 426, 450–454
6. Wong, S. K., Li, W., Moore, M. J., Choe, H. and Farzan, M. (2004) A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. J. Biol. Chem., 279, 3197–3201
7. Li, F., Li, W., Farzan, M. and Harrison, S. C. (2005) Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science, 309, 1864–1868
8. Kuba, K., Imai, Y., Ohto-Nakanishi, T. and Penninger, J. M. (2010) Trilogy of ACE2: a peptidase in the renin-angiotensin system, a SARS receptor, and a partner for amino acid transporters. Pharmacol. Ther., 128, 119–128
9. Bao, L., Deng, W., Huang, B., Gao, H., Ren, L., Wei, Q., Yu, P., Xu, Y., Liu, J., Qi, F., Qu, Y., *et al.* (2020)The pathogenicity of 2019 novel coronavirus in hACE2 transgenic mice. BioRxiv, 2020.02.07.939389
10. Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., *et al.* (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature, 581, 215–220
11. Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., Lu, G., Qiao, C., Hu, Y., Yuen, K. Y., *et al.* (2020) Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell, 181, 894–904.e9
12. D. E. Shaw Research, Molecular Dynamics Simulations Related to SARS-CoV-2, 2020. http://www.deshawresearch.com/resources_-sarscov2.html
13. Li, W., Zhang, C., Sui, J., Kuhn, J. H., Moore, M. J., Luo, S., Wong, S. K., Huang, I. C., Xu, K., Vasilieva, N., *et al.* (2005) Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. EMBO J., 24, 1634–1643
14. Lavillette, D., Barbouche, R., Yao, Y., Boson, B., Cosset, F. L., Jones, I. M. and Fenouillet, E. (2006) Significant redox insensitivity of the functions of the SARS-CoV spike glycoprotein: comparison with HIV envelope. J. Biol. Chem., 281, 9200–9204
15. Li, W., Greenough, T. C., Moore, M. J., Vasilieva, N., Somasundaran, M., Sullivan, J. L., Farzan, M. and Choe, H. (2004) Efficient replication of severe acute respiratory syndrome coronavirus in mouse cells is limited by murine angiotensin-converting enzyme 2. J. Virol., 78, 11429–11433
16. Li, F. (2008) Structural analysis of major species barriers between humans and palm civets for severe acute respiratory syndrome

coronavirus infections. J. Virol., 82, 6984–6991
17. Li, F. (2015) Receptor recognition mechanisms of coronaviruses: a decade of structural studies. J. Virol., 89, 1954–1964
18. Lee, H. S., Qi, Y. and Im, W. (2015) Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. Sci. Rep., 5, 8926
19. Wang, Y., Liu, M. and Gao, J. (2020) Enhanced receptor binding of SARS-CoV-2 through networks of hydrogen-bonding and hydrophobic interactions. Proc. Natl. Acad. Sci. USA, 117, 13967–13974
20. SCMP (2020) Coronavirus: Hong Kong confirms a second dog is infected. https://www.scmp.com/news/hong-kong/health-environment/article/3075993/coronavirus-hong-kong-confirms-second-dog (Accessed: March 15, 2020)
21. CASP (2020) CASP and Covid-19. http://predictioncenter.org/caspcommons/index.cgi (Accessed: March 25, 2020)
22. Institute for Protein Design (2020) Coronavirus (SARS-CoV-2/COVID-19/nCoV-2019). https://www.ipd.uw.edu/coronavirus/ (Accessed: March 20, 2020)
23. Zhang, C., Zheng, W., Huang, X., Bell, E. W., Zhou, X. and Zhang, Y. (2020) Structure models of all mature peptides in COVID-19 genome by C-I-TASSER, Zhang Lab Univ. Michigan
24. AlphaFold Team. (2020) Computational predictions of protein structures associated with COVID-19. https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19 (Accessed: March 10, 2020)
25. Zhang, G., ,Pomplun S., Loftis, A.R., Tan, X., Loas, A. and Pentelute, B.L. (2020) Investigation of ACE2 N-terminal fragments binding to SARS-CoV-2 Spike RBD. BioRxiv. 2020.03.19.999318
26. Lupala, C.S., Kumar, V., Li, X., Su, X. and Liu, H. (2020) Computational analysis on the ACE2-derived peptides for neutralizing the ACE2 binding to the spike protein of SARS-CoV-2. BioRxiv. 2020.05.03.075473
27. Han, Y. and Král, P. (2020) Computational design of ACE2-based peptide inhibitors of SARS-CoV-2. ACS Nano., 14, 5143–5147
28. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res., 46, W296–W303
29. Wu, L., Chen, Q., Liu, K., Wang, J., Han, P., Zhang, Y., Hu, Y., Meng, Y., Pan, X., Qiao, C., *et al.* (2020) Broad host range of SARS-CoV-2 and the molecular basis for SARS-CoV-2 binding to cat ACE2. Cell Discov., 6, 68
30. Anandakrishnan, R., Aguilar, B. and Onufriev, A. V. (2012) H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic Acids Res., 40, W537–W541
31. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem., 25, 1605–1612
32. Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., *et al.* (2016) CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. J. Chem. Theory Comput., 12, 405–413
33. Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M. and Mackerell, A. D. Jr. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles. J. Chem. Theory Comput., 8, 3257–3273
34. Nosé, S. (1984) A unified formulation of the constant temperature molecular dynamics methods. J. Chem. Phys., 81, 511–519
35. Hoover, W. G. (1985) Canonical dynamics: Equilibrium phase-space distributions. Phys. Rev. A Gen. Phys., 31, 1695–1697
36. Parrinello, M. and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. J. Appl. Phys., 52, 7182–7190
37. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. and Lindah, E. (2015) Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX, 1–2, 19–25
38. Darden, T., York, D. and Pedersen, L. (1993) Particle mesh Ewald: An $N\cdot\log(N)$ method for Ewald sums in large systems. J. Chem. Phys., 98, 10089–10092
39. Hess, B., Bekker, H., Berendsen, H. J. C. and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. J. Comput. Chem., 18, 1463–1472
40. Kumari, R., Kumar, R. and Lynn, A., and the Open Source Drug Discovery Consortium. (2014) *g_mmpbsa*–a GROMACS tool for high-throughput MM-PBSA calculations. J. Chem. Inf. Model., 54, 1951–1962
41. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. J. Mol. Graph., 14, 33–38