

Computational study of the strong binding mechanism of SARS-CoV-2 spike and ACE2

Cheng Peng^{1,2,#}, Zhengdan Zhu^{1,2,#}, Yulong Shi^{1,2}, Xiaoyu Wang^{1,3}, Kaijie Mu^{1,4}, Yanqing Yang^{1,2}, Xinben Zhang¹, Zhijian Xu^{1,2,*}, Weiliang Zhu^{1,2,*}

¹CAS Key Laboratory of Receptor Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China

²School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, PR China

³College of Mathematics and Physics, Shanghai University of Electric Power, Shanghai 200090, China.

⁴Nano Science and Technology Institute, University of Science and Technology of China, Suzhou, Jiangsu, 215123, China.

[#]These authors contributed equally to this work.

^{*}To whom correspondence should be addressed.

Phone: +86-21-50806600-1201 (Z.X.), +86-21-50805020 (W.Z.),

E-mail: zjxu@simm.ac.cn (Z.X.), wlzhu@simm.ac.cn (W.Z.).

ORCID:

Zhijian Xu: 0000-0002-3063-8473

Weiliang Zhu: 0000-0001-6699-5299

Abstract

The spike protein of SARS-CoV-2 (SARS-CoV-2-S) helps the virus attach to and infect human cells. With various computational methods applied in this work, the accessibility of its RBD to ACE2, its key residues for stronger binding to ACE2 than the SARS-CoV spike (SARS-CoV-S), the origin of the stronger binding, and its potential sites for drug and antibody design were explored. It was found that the SARS-CoV-2-S could bind ACE2 with an RBD-angle ranging from 52.2° to 84.8°, which demonstrated that the RBD does not need to fully open to bind ACE2. Free energy calculation by an MM/GBSA approach not only revealed much stronger binding of SARS-CoV-2-S to ACE2 ($\Delta G = -21.7 \sim -29.9$ kcal/mol) than SARS-CoV-S ($\Delta G = -10.2 \sim -16.4$ kcal/mol) at different RBD-angles but also demonstrated that the binding becomes increasingly stronger as the RBD-angle increases. In comparison with the experimental results, the free energy decomposition disclosed more key residues interacting strongly with ACE2 than with the SARS-CoV-S, among which the Q493 might be the decisive residue variation (-5.84 kcal/mol) to the strong binding. With the mutation of all 18 different residues of SARS-CoV-S on the spike-ACE2 interface to the corresponding residues of SARS-CoV-2-S, it was found that the mutated SARS-CoV-S has almost the same binding affinity as SARS-CoV-2-S to ACE2, demonstrating that the remaining mutations outside the spike-ACE2 interface have little effect on its binding affinity to ACE2. Simulation of the conformational change pathway from “down” to “up” states disclosed 5 potential ligand-binding pockets correlated to the conformational change. Taking together the key residues, accessible RBD-angle and pocket correlation, potential sites for drug and antibody design were proposed, which should be helpful for interpreting the high infectiousness of SARS-CoV-2 and for developing a cure.

1. Introduction

Very recently, a new coronavirus closely related to SARS-CoV (severe acute respiratory syndrome coronavirus),¹⁻³ temporarily named SARS-CoV-2 by the International Committee on the Taxonomy of Viruses (ICTV), has emerged and spread worldwide rapidly.^{4,5} Although the number of infections is still growing, no drug has thus far been approved to be effective. Therefore, it is very urgent to discover and develop safe and effective therapeutics.

Compared to SARS-CoV, SARS-CoV-2 is likely easier to transmit human to human.^{6,7} The spike (S) glycoprotein of SARS-CoV-2 (SARS-CoV-2-S hereinafter) is a class I viral fusion protein, which plays an essential role in viral infection engagement of human angiotensin-converting enzyme 2 (ACE2) as a receptor and mediating the fusion of SARS-CoV-2 and cellular membranes.^{8,9} The significant functions of the protein make it an ideal drug target against SARS-CoV-2. The protein consists of an amino (N)-terminal S1 subunit and a carboxyl (C)-terminal S2 subunit. To recognize ACE2, the receptor-binding domain (RBD) of the S1 subunit undergoes hinge-like conformational changes to expose enough space for receptor binding.¹⁰⁻¹² Therefore, there are two states of the protein that are referred to as “down” and “up” conformations, where the “down” conformation (small RBD-angle, defined as the angle composed of the three C α atoms of D405-V622-V991, Figure 1A) is a receptor-inaccessible state, and the “up” conformation (large RBD-angle, Figure 1B) is a receptor-accessible state.¹³⁻¹⁶ Once binding to ACE2, the SARS-CoV-2-S employs host proteases for S protein priming.^{17,18}

To increase our understanding of the binding mechanisms between the SARS-CoV-2-S and ACE2 at the atomic level, various computational techniques, including homology modeling, MD (molecular dynamics) simulation, binding free energy calculation with the MM/GBSA (Molecular Mechanics/Generalized Born Surface Area) method^{19,20}, binding free energy decomposition analysis, conformational transition pathway prediction with NUMD (a combination of normal-mode analysis and umbrella sampling MD) approach,^{21,22} and potential ligand-binding sites prediction were performed in this study. The results not only revealed that the SARS-CoV-2-S binds to ACE2 with higher affinity compared with the SARS-CoV spike glycoprotein (SARS-CoV-S hereinafter) but also identified key residues of the interactions between the SARS-CoV-2-S and ACE2. In addition, we found that 52.2° is the smallest accessible RBD-angle for binding ACE2. We also found that as the RBD-angle increases starting from 52.2°, the binding strength of the SARS-CoV-2-S and ACE2 becomes increasingly stronger. Based on the “down” to “up” conformational change pathway, we identified 5 potential ligand-binding pockets that are dynamically correlated with

the conformational change of SARS-CoV-2-S. We think that the knowledge of the interactional details at the atomic and energetic levels between the SARS-CoV-2-S and ACE2 should be useful for understanding the functional mechanism of SARS-CoV-2-S and for developing inhibitors and antibodies for coronavirus disease 2019 (COVID-19).

2. Materials and methods

2.1 Homology modeling. The structures of 4 ACE2-bound SARS-CoV-S (PDB IDs: 6ACG, 6ACK, 6ACJ, and 6CS2) and 2 ACE2-free SARS-CoV-S (5X58, and 5X5B) proteins were downloaded from Protein Data Bank (PDB).²³ They served as templates to build the three-dimensional (3D) models of the “up” and “down” conformations of the SARS-CoV-2-S with the SWISS-MODEL server using the “user template” mode.²⁴

2.2 MD simulation. For optimizing the constructed 3D model and calculating the binding free energy between the spikes and ACE2, each system was solvated in a cubic box of TIP3P water extended by 9 Å from the solute, with a rational number of counter ions of Na⁺ or Cl⁻ to neutralize the system. Amber ff03 force field²⁵ was used to parameterize the protein. To remove bad contacts formed during the system preparation, 10,000 steps of minimization with constraints (10 kcal/mol/Å²) on heavy atoms, including 5,000 steps of steepest descent minimization and 5,000 steps of conjugate gradient minimization, were performed. Then, each system was heated to 300 K within 0.2 ns followed by 0.1 ns of equilibration in NPT ensemble. The heating and equilibrium simulations were performed with constraints (5 kcal/mol/Å²) on heavy atoms. Finally, a 5-ns MD simulation on each system at 300 K was performed. Langevin dynamics was used to control temperature, and the SHAKE algorithm²⁶ was applied to fix bonds involving hydrogen atoms. A cutoff distance of 12 Å was applied for van der Waals interactions. The particle mesh Ewald method was applied to address long-range electrostatic interactions²⁷. Minimization, heating and equilibrium were performed with *sander* program in Amber16. The 5 ns production run was performed with *pmemd.cuda*.

2.3 Binding free energy calculation. To evaluate the binding affinity between the SARS-CoV-2-S or SARS-CoV-S and ACE2, the MM/GBSA method was used to calculate the binding free energy (ΔG) based on 5-ns MD trajectories with equation (1):

$$\Delta G = \Delta H - T\Delta S = \Delta E_{ele} + \Delta E_{VDW} + \Delta G_{gb} + \Delta G_{np} - T\Delta S \quad (1)$$

where ΔE_{ele} and ΔE_{VDW} are the electrostatic and van der Waals energy terms, and ΔG_{gb} and ΔG_{np} are the polar and nonpolar solvation free energies, respectively. The *Nmode* module in Amber16 was used to calculate the conformational entropy ($T\Delta S$). In this study, the dielectric constants for solvent and solute were set to 80.0 and 1.0, respectively, and OBC solvation model ($igb = 5$, $mbondi2$)²⁸ was applied. The ΔG_{np} was calculated by equation (2):

$$\Delta G_{np} = \gamma \text{SASA} + \beta \quad (2)$$

where the solvent-accessible surface area (SASA) was calculated with $\gamma=0.0072$ kcal/mol/Å² and $\beta= 0.0$ kcal/mol, respectively.²⁹ To find out crucial residues in the binding of S protein with ACE2, the binding affinity was further decomposed into energy contribution of each residue using the *MMPBSA.py* script in Amber16 (idecomp = 1).

2.4 Conformational change pathway prediction. The “down” to “up” conformational change pathway of the SARS-CoV-2-S was generated by the NUMD method, of which the details have been described in our previous study.^{21, 22} Briefly, many iterations of normal modes analysis (NMA) were run to predict the conformational changes from the initial structure to the final target structure gradually. For example, the intermediate structure $R^{(k)}$ in iteration k was generated by the following equation based on the intermediate structure $R^{(k-1)}$ in iteration $(k-1)$:

$$R^{(k)} = R^{(k-1)} + v^{(k)} = R^{(k-1)} + S^{(k)} \sum_i^{m^{(k)}} (d^{(k-1)} \cdot u_i^{(k)}) u_i^{(k)} \quad (3)$$

where $v^{(k)}$ is the displacement combined with $m^{(k)}$ low-frequency eigenmodes that are calculated by NMA. For the i_{th} eigenmode, its displacement is proportional to the projection $d^{(k-1)} u_i^{(k)}$, where $d^{(k-1)}$ is the instantaneous distance vector on eigenvector $u_i^{(k)}$, and scaled by the step size $S^{(k)}$. In this study, the step size was set to be 10.0, consistent with our previous study²¹. The starting and final structures were obtained from homology modeling based on the 3D structures of 5X58 (“down” conformation with RBD-angle of 31.6°) and 5X5B (“up” conformation with RBD-angle of 84.8°), respectively, chosen from Table 1 with the highest resolution. To evaluate potential ligand-binding sites for molecular docking in SARS-CoV-2-S, we applied the D3Pockets³⁰ to analyzed the dynamic properties of potential binding sites (<https://www.d3pharma.com/D3Pocket/index.php>), where the predicted conformational change pathway was prepared in format of ‘*mdcrd*’ as input.

3. Results and Discussion

3.1 Overview of the trimer structures of the SARS-CoV-S in PDB. Amino acid sequence alignment revealed that the SARS-CoV-2-S shares 76% identity with that of the SARS-CoV-S (Supporting Information Figure S1). The SARS-CoV-S adopts a homotrimer architecture, of which the RBD undergoes a hinge-like conformational switch from a prefusion state to a postfusion state. As shown in Table 1, 5 trimer structures of ACE2-free SARS-CoV-S (PDB ID: 5WRG, 5X58, 5XLR, 6ACC and

6ACD) were found in PDB, with three “down” RBDs (small RBD-angles ranging from 30.1° to 33.4°). Four ACE2-bound SARS-CoV-S trimers (6ACG, 6ACJ, 6ACK, and 6CS2) were found in PDB, each of which has a single RBD in the “up” conformation with different RBD-angles ranging from 54.8° to 84.6°, revealing the flexibility of the “up” RBD-angle. In addition, an “up” conformation was observed in an ACE2-free structure (5X5B), demonstrating that the trimer of the spike protein itself may adopt different conformations, with the RBD-angle ranging from 30.1 to 84.8°.

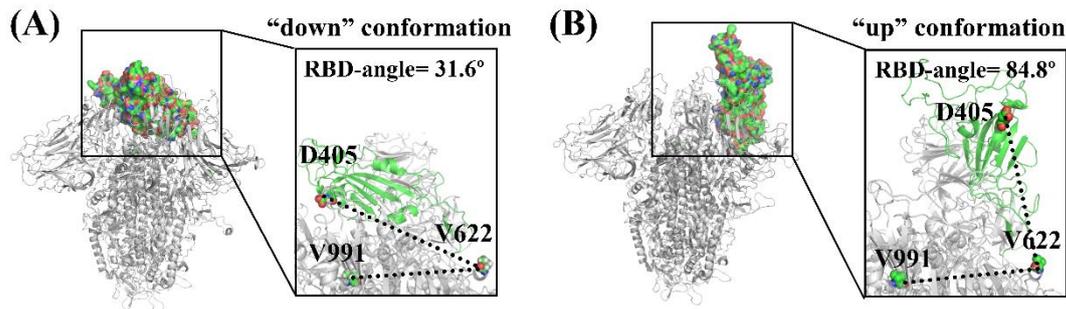


Figure 1. The three-dimensional models of SARS-CoV-2 spike glycoprotein (SARS-CoV-2-S) and the schematic diagram of RBD-angle. (A), The “down” conformation of the RBD modeled using the chain A of 5X58 as a template. (B), The “up” conformation of the RBD modeled using the chain A of 5X5B as a template. The RBD-angle is defined as the angle composed of the three C α atoms of the residues D405-V622-V991, corresponding to the angle of D392-T608-V973 in SARS-CoV-S.

Table 1. Summary of the SARS-CoV-S trimers in PDB.

PDB ID	Resolution (Å)	Chain	ACE2	RBD state ^a	RBD-angle (°) ^b
5WRG ¹⁴	4.3	A	-	down	30.1
		B	-	down	30.1
		C	-	down	30.1
5X58 ¹⁶	3.2	A	-	down	31.6
		B	-	down	31.6
		C	-	down	30.7
5X5B ¹⁶	3.7	A	-	up	84.8
		B	-	down	30.9
		C	-	down	30.9
5XLR ¹⁴	3.8	A	-	down	32.1
		B	-	down	32.1
		C	-	down	32.1
6ACC ⁸	3.6	A	-	down	33.4
		B	-	down	33.4
		C	-	down	33.4

6ACD ⁸	3.9	A	-	down	32.8
		B	-	down	32.8
		C	-	down	32.9
6ACG ⁸	5.4	A	-	down	32.6
		B	-	down	32.7
		C	ACE2	up	54.8
6ACJ ⁸	4.2	A	-	down	33.0
		B	-	down	33.3
		C	ACE2	up	68.3
6ACK ⁸	4.5	A	-	down	33.1
		B	-	down	33.8
		C	ACE2	up	84.6
6CRV ³¹	3.2	A	-	-	-
		B	-	-	-
		C	-	-	-
6CRW ³¹	3.9	A	-	down	34.3
		B	-	up	68.8
		C	-	down	34.2
6CRX ³¹	3.9	A	-	up	71.6
		B	-	up	70.6
		C	-	down	38.1
6CRZ ³¹	3.3	A	-	down	34.1
		B	-	up	68.8
		C	-	down	34.1
6CS0 ³¹	3.8	A	-	down	34.2
		B	-	up	68.8
		C	-	down	34.1
6CS1 ³¹	4.6	A	-	up	71.6
		B	-	up	70.7
		C	-	down	38.1
6CS2 ³¹	4.4	A	-	-	-
		B	ACE2	up	74.0
		C	-	-	-
6NB6 ¹⁵	4.2	A	-	down	30.7
		B	-	up	77.9
		C	-	up	55.2
6NB7 ¹⁵	4.5	A	-	up	75.3
		B	-	up	70.6
		C	-	up	78.5

^a: The RBD state is assigned according to the corresponding references.

^b: The RBD-angle is determined by the residues D392-T608-V973 in the SARS-CoV-S.

3.2 Higher affinity of the SARS-CoV-2-S binding to ACE2 than the SARS-CoV-S.

To compare the difference in the binding affinities of the two S proteins and ACE2, the MM/GBSA method was used to predict the ΔG , which has been recommended and applied in quite a number of protein-protein systems.³² Two initial structures of ACE2-bound SARS-CoV-2-S with the smallest (54.8°) and largest (84.6°) RBD-angles, respectively, were obtained from homology modeling using 6ACG and 6ACK as templates (Table 1).

As shown in Table 2, the calculated ΔG of the SARS-CoV-2-S binding to ACE2 with the RBD-angle of 54.8° is -21.74 ± 0.65 kcal/mol, which is obviously stronger than that of the SARS-CoV-S binding to ACE2 (-10.17 ± 0.63 kcal/mol), consistent with recent experimental results³³. The calculated ΔG (-29.90 ± 0.80 kcal/mol) of the SARS-CoV-2-S binding to ACE2 with the RBD-angle of 84.6° is also much stronger than that of the SARS-CoV-S (-15.46 ± 0.68 kcal/mol), and similar results could be found for 6ACJ with the RBD-angle of 68.3° and 6CS2 with the RBD-angle of 74.0° (Table S1), revealing that the SARS-CoV-2-S could maintain higher affinity binding to ACE2 than the SARS-CoV-S, regardless of the small or large RBD-angles. These findings provide a theoretical validation that the SARS-CoV-2-S binds to ACE2 with higher affinity than to the SARS-CoV-S, in good agreement with the observation that SARS-CoV-2 might be more readily transmitted human to human than SARS-CoV. The results also indicated that the SARS-CoV-2-S might have higher affinity, with more “up” RBD domain, according to the results of the two systems, with RBD-angles of 54.8° and 84.6° , respectively, which deserves further study.

Table 2. Components of the binding free energy (kcal/mol) calculated by the MM/GBSA method*

Energy term	6ACG (RBD-angle = 54.8°)		6ACK (RBD-angle = 84.6°)	
	SARS-CoV	SARS-CoV-2	SARS-CoV	SARS-CoV-2
E_{VDW}	-80.57 ± 0.46	-87.07 ± 0.49	-96.89 ± 0.59	-105.05 ± 0.36
E_{ele}	65.07 ± 0.58	-673.99 ± 3.96	-7.57 ± 0.32	-641.25 ± 4.07
E_{gb}	0.90 ± 0.02	737.98 ± 3.86	83.60 ± 0.26	714.56 ± 3.65
E_{np}	-10.31 ± 0.06	-12.21 ± 0.06	-12.93 ± 0.08	-15.03 ± 0.07
ΔH	-24.91 ± 0.50	-35.30 ± 0.60	-33.80 ± 0.74	-46.77 ± 0.61
$-T\Delta S$	-14.74 ± 0.76	-13.56 ± 0.70	-18.34 ± 0.62	-16.87 ± 0.98
ΔG	-10.17 ± 0.63	-21.74 ± 0.65	-15.46 ± 0.68	-29.90 ± 0.80

*: The statistical error was estimated based on 0.5-5 ns MD simulation trajectory. A total of 500 snapshots evenly extracted from the 0.5-5 ns MD trajectory of each complex were used for MM/GBSA calculations, with 10 snapshots for the entropy term calculations.

3.3 Key residues of the SARS-CoV-2-S and SARS-CoV-S interacting with ACE2.

To identify key residues contributing to the strong binding between the S protein and ACE2, the binding free energies were decomposed into residues. Although both spikes share similar interaction profiles with ACE2^{34, 35}, the SARS-CoV-2-S has always more residues interacting with ACE2 than that of the SARS-CoV-S (Figure 2). For the two complexes with the RBD-angle of 54.8° (Figure 2A), 9 residues of the SARS-CoV-S, e.g., Y442, L443, P462, L472, N473, Y475, Y484, T487, and Y491 (blue colored), are favorable for its binding to ACE2 with $\Delta G \leq -1.0$ kcal/mol per residue, while the corresponding 9 residues of the SARS-CoV-2-S are L455, F456, A475, F486, N487, Y489, Q498, N501, and Y505 (black colored), assigned by sequence alignment, which were all included in recent experiments.^{34, 35} In particular, the residue Y491 contributes -4.03 ± 0.60 kcal/mol to the binding between the SARS-CoV-S and ACE2, while the corresponding residue Y505 contributes -4.23 ± 0.56 kcal/mol to that of the SARS-CoV-2-S. The residues Y442, L472, and T487 of the SARS-CoV-S have been previously reported to be important for ACE2 binding.³⁶

We identified 5 more residues of the SARS-CoV-2-S interacting with ACE2 than that of the SARS-CoV-S, which are Y449, Q493, G496, T500, and G502 in the SARS-CoV-2-S. In particular, the residue Q493 contributes -3.49 ± 0.48 kcal/mol to the binding of the SARS-CoV-2-S and ACE2, which has also been noted as a particularly critical residue to provide favorable interactions with ACE2³⁶. Apparently, the higher number of interactions and residue variations on the protein-protein interface should be the main reason for the higher binding affinity of the SARS-CoV-2-S than that of the SARS-CoV-S to ACE2, which could be used to at least partially interpret the stronger binding observed from experimental results.^{34, 36}

For the two complexes with RBD-angles of 84.6° (Figure 2B), the simulation results revealed 3 more residues of the SARS-CoV-2-S interacting with ACE2 than that of the SARS-CoV-S, with $\Delta G \leq -1.0$ kcal/mol per residue (Figure 2C & D). In particular, the residue Q493 contributes -5.84 kcal/mol (Figure 2B), which is approximately one-fifth of the overall binding free energy (-29.9 kcal/mol). Therefore, the residue variation of N479 of SARS-CoV-S to Q493 of SARS-CoV-2-S might play a decisive role in its strong binding to ACE2. It was also found that the length of the hydrogen bond formed by T486 in SARS-CoV-S and K353 in ACE2 (2.65 ± 0.04 Å) of 6ACK was obviously shorter than that of 6ACG (3.37 ± 0.04 Å) during the MD simulations, and a similar shift was observed for the length of hydrogen bond formed by G488 in SARS-CoV-S and Y41 in ACE2 (3.44 ± 0.05 Å in 6ACK and 4.56 ± 0.04 Å in 6ACG, respectively),

resulting in the stronger binding of T486 and G488 in 6ACK (Figure S2). A site composed of residues with more binding affinity contributions should be a potential site for drug discovery and antibody design (Figure S3). More impressively, most of the key residues of SARS-CoV-2-S showed stronger binding affinity per residue to ACE2 in the large “up” conformation (RBD-angle=84.6°, Figure 2B) than in the small “up” conformation (RBD-angle=54.8°, Figure 2A), leading to the much stronger binding affinity of SARS-CoV-2-S to ACE2 (-29.9 kcal/mol) than that of SARS-CoV-S (-15.5 kcal/mol) at the RBD-angle of 84.6°. This result demonstrated again that a potential correlation between RBD-angle and binding affinity deserves further study.

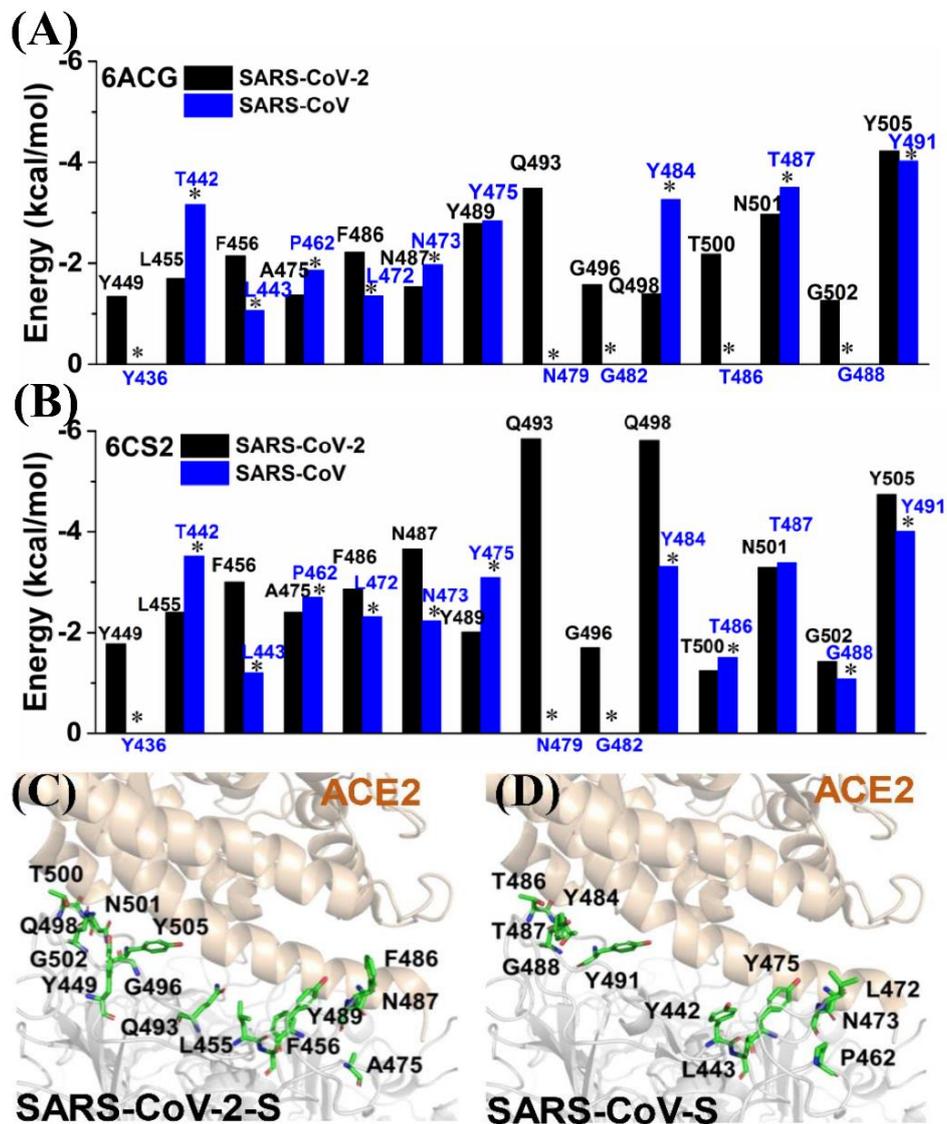


Figure 2. The spike-ACE2 interaction spectrum of SARS-CoV-2 and SARS-CoV. The initial structures of MD simulations were modeled based on the 3D structures of 6ACG (A) and 6ACK (B). The difference that is statistically significant at the 1% level ($p <$

0.01) was labeled with “*”. Key residues of the SARS-CoV-2-S (C) and SARS-CoV-S (D) interacting with ACE2 with the RBD-angle of 84.6°, shown in sticks colored green. Each residue contributed ≤ -1.00 kcal/mol to the overall binding free energy.

3.4 Origin of the stronger binding affinity of SARS-CoV-2-S with ACE2. To further investigate the influence of residue variation on the RBD-ACE2 interface on the significant difference of the binding affinity between SARS-CoV-S and SARS-CoV-2-S to ACE2, we compared all the residues on the spike-ACE2 interfaces of the two spikes and found that there are 32 residues from the spikes in the interface, among which 18 residues are different between SARS-CoV-S and SARS-CoV-2-S. To investigate how the 18 residues’ variations affect the binding affinity to ACE2, we mutated all 18 residues of SARS-CoV-S to the corresponding residues of SARS-CoV-2-S (Table S2), *viz.*, K390R, R426N, S432V, T433G, Y442L, L443F, F460Y, P462A, D463G, K465T, L472F, Y476F, N479Q, D480S, Y484Q, T485P, T487N and I489V. Then, we calculated the ΔG of the mutated SARS-CoV-S binding to ACE2 in 6ACG with the MM/GBSA method. The predicted free energy of the mutated spike of SARS-CoV is -21.02 ± 0.57 kcal/mol, which is almost the same as the strength of the SARS-CoV-2-S binding to ACE2 (-21.74 ± 0.65 kcal/mol, Table 2), but is much stronger than that of the wild SARS-CoV-S to ACE2 in 6ACG (-10.17 ± 0.63 kcal/mol). This result demonstrated that the stronger binding affinity of SARS-CoV-2-S to ACE2 could be quantitatively attributed to these residue variations. The remaining residue mutations of the RBD have little effect on its binding strength to ACE2. These results suggested that the high infectiousness of SARS-CoV-2 should result from the variation of those residues on the ACE2 interface.

3.5 ACE2-accessible RBD-angle with application for classifying “up” and “down” conformations. The two states, RBD “down” and “up” conformations, correspond to the receptor-inaccessible and receptor-accessible states, respectively. As shown in Table 1, ACE2-bound SARS-CoV-S has a flexible RBD-angle ranging from 54.8° to 84.6°. However, it remains unknown that how large the RBD-angle becomes enables the SARS-COV-2-S to start binding ACE2. To identify the smallest ACE2-accessible RBD-angle, we predicted the conformational change pathway by NUMD (Figures 3A & S4) and found that the structure of the RBD is insensitive to changes in the RBD-angle (Figure S5). Then, we superimposed the RBD-ACE2 structures of 6ACG on 240 conformations of SARS-CoV-2-S with different RBD-angles (Figure 3A) and found that there is no atomic collision between ACE2 and the S protein if the RBD-angle is

larger than 52.2° , suggesting that 52.2° is the starting accessible RBD-angle for the SARS-CoV-2-S to bind ACE2 (Figure 3B). Accordingly, we defined 52.2° as the smallest ACE2-accessible RBD-angle of the SARS-CoV-2-S, which could be used as the criterion for classifying “up” or “down” states. For example, as shown in Table 1, all the “down” conformations of SARS-CoV-S have RBD-angles less than 52.2° , and the smallest RBD “up” angle 54.8° in 6ACG is larger than 52.2° . Accordingly, the SARS-CoV-2-S trimer in the most recent cryo-EM structure (PDB ID: 6VSB³³) could be classified as an “up” conformation, as it has a RBD-angle of 57.5° , which is able to bind the ACE2 molecules of human cells.

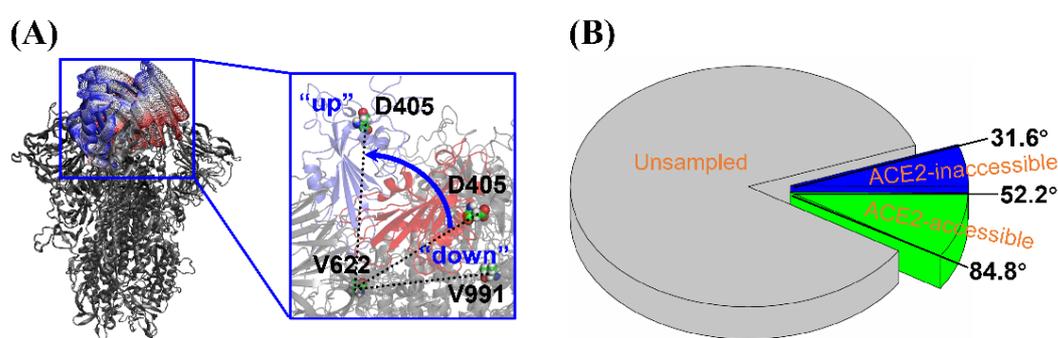


Figure 3. (A), Twenty aligned conformations extracted from the predicted conformational change pathway of SARS-CoV-2-S. (B), The ACE2-inaccessible RBD-angle (blue), ACE2-accessible RBD-angle (green), and unsampled RBD-angle (gray) of SARS-CoV-2-S protein.

3.6 Correlation between the RBD-angle and binding affinity.

To further investigate the correlation between the RBD-angle and binding affinity, we calculated the ΔG against different RBD-angles along the conformational change pathway. As shown in Figure 4 and Table S3, one can conclude that the SARS-CoV-2-S has higher affinity with a larger RBD-angle. This result demonstrated that the virus could start to interact with host cells at an RBD-angle as small as 52.2° and steadily strengthens its binding interaction with the host cell as the RBD-angle increases in size. When the RBD becomes almost fully opened (RBD-angle= 84.6°), the binding strength in terms of MM/GBSA could be as strong as -29.9 kcal/mol, which is much stronger than that of SARS-CoV-S at the same angle (-15.5 kcal/mol). A similar correlation between the RBD-angle and binding affinity could also be found for SARS-CoV-S and ACE2 with correlation coefficient of $R^2=0.64$ (Figure S6).

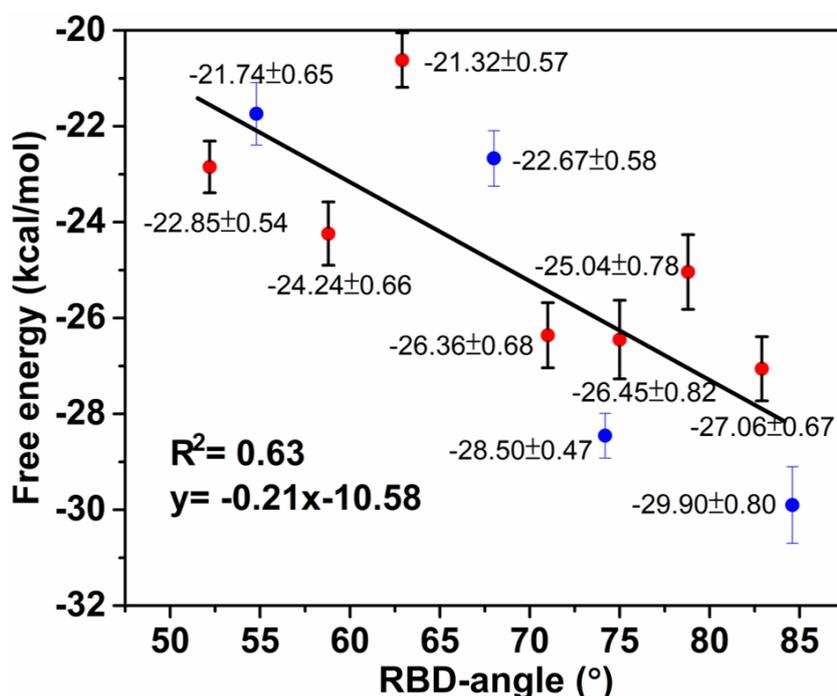


Figure 4. The calculated binding free energy of SARS-CoV-2-S to ACE2 against the RBD-angle. The correlation coefficient ($R^2 = 0.63$) is calculated based on 11 systems, including 7 conformations predicted with NUMD (colored red) and 4 homology models (colored blue).

3.7 Potential ligand-binding sites for molecular docking.

Based on the “down” to “up” conformational change pathway predicted by NUMD, twenty conformations were then extracted for exploring potential ligand-binding pockets correlated to the conformational transition, which could be used as potential sites for molecular docking. Using the method D3Pockets³⁰, we identified 5 potential druggable ligand-binding pockets, which are dynamically correlated with the state change from “down” to “up” (Figure 5A). Meanwhile, it was also observed that there are positive correlations among the 5 pockets except that between Pocket 1 and Pocket 5. For instance, Pocket 2 has a positive volume correlation with Pocket 4 (Figure 5B) along the conformation transition pathway. In detail, as the conformation of the SARS-CoV-2-S changes from “down” to “up” states, both Pocket 2 and Pocket 4 also become larger (Figure 5C). As discussed above, the RBD is inaccessible to ACE2 when the RBD-angle is less than 52°. Taking into account the angle and the pocket correlation, all 5 predicted pockets on the conformations with RBD-angles less than 52° should be promising conformations and sites that should be useful for discovering drugs to block

the conformational change of spike from “down” to “up” states, which may cause the spike to freeze in conformations unable to bind ACE2.

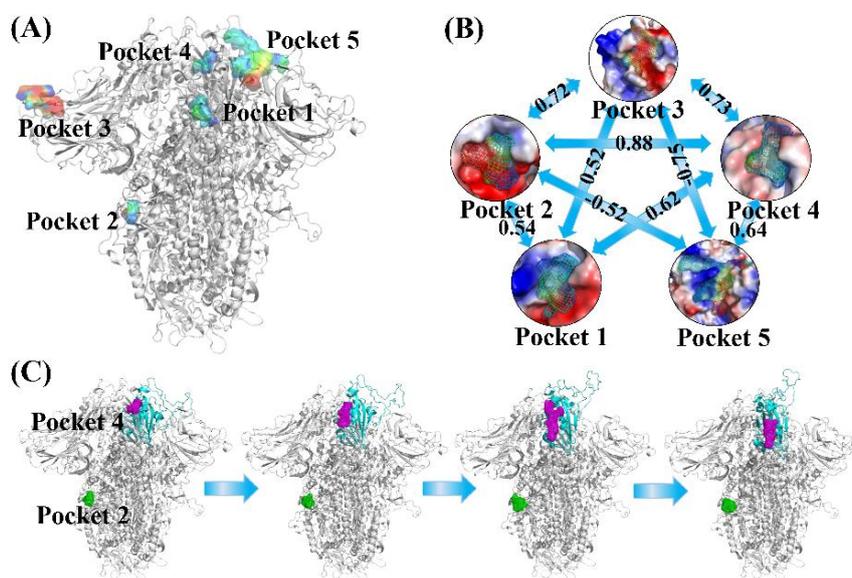


Figure 5. The predicted potential ligand binding sites of SARS-CoV-2-S. A, Five potential ligand-binding pockets. The redder the pocket grids are, the more stable the subpocket regions throughout the pathway. B, Pocket correlation between the 5 predicted binding pockets. The numbers are the correlation coefficients between two pockets during conformational changes from “down” to “up” states. C, The pocket correlation between Pocket 2 and Pocket 4.

4. Conclusions

There is no approved effective drug for COVID-19 to date. The SARS-CoV-2 spike glycoprotein (SARS-CoV-2-S) is an ideal drug target because of its indispensable function for viral infection and fusion, engaging ACE2 as an entry receptor. To facilitate drug discovery and development with the SARS-CoV-2-S as the drug target, various computational techniques were used in this study to investigate the strong binding mechanism between the protein and its receptor, ACE2. Compared with SARS-CoV-S, the SARS-CoV-2-S has obviously higher binding affinity to ACE2, as predicted by MM/GBSA, which may partially account for the stronger infectiousness of SARS-CoV-2. The binding free energy decomposition analysis further showed that more interactions formed in SARS-CoV-2-S binding to ACE2, resulting in the higher binding affinity. In particular, the residue variation of N479 to Q493 might be the decisive variation of SARS-CoV-2-S to gain its strong binding to ACE2. Based on the calculated binding free energies of SARS-CoV-2-S at different RBD-angles, it was noted that the

binding affinity of SARS-CoV-2-S to ACE2 becomes increasingly stronger as the RBD-angle becomes increasingly larger. In addition, we found that 52.2° is a starting ACE2-accessible RBD-angle, consistent with experimental results. Therefore, conformations with RBD-angles smaller than 52.2° are ideal target structures for designing drugs and antibodies to block the binding between the spike and ACE2. Based on the “down” to “up” conformational change pathway, we identified 5 potential ligand-binding pockets that are dynamically correlated with the conformational change of SARS-CoV-2-S with the method D3Pockets. We hope that this work will provide useful information for understanding the interaction mechanism of the SARS-CoV-2-S and ACE2 and for developing inhibitors and antibodies to address the ongoing public health crisis.

Associated contents

Supporting Information: Binding free energy calculations of 6ACJ and 6CS2 (Table S1 & S3) and residues on the RBD-ACE2 interface (Table S2). All figures were discussed in this manuscript.

Acknowledgments

This work was supported by the National Key R&D Program of China (2016YFA0502301 & 2017YFB0202601). The simulations were partially run at the TianHe 1 supercomputer in Tianjin.

References

1. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New. Engl. J. Med.* **2020**, 382,727-733.
2. Zhou, P.; Yang, X.; Wang, X.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.; Zhu, Y.; Li, B.; Huang, C.; Chen, H.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R.; Liu, M.; Chen, Y.; Shen, X.; Wang, X.; Zheng, X.; Zhao, K.; Chen, Q.; Deng, F.; Liu, L.; Yan, B.; Zhan, F.; Wang, Y.; Xiao, G.; Shi, Z. A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin. *Nature* **2020**, 579, 270-273.
3. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.; Wang, W.; Song, Z.; Hu, Y.; Tao, Z.; Tian, J.;

- Pei, Y.; Yuan, M.; Zhang, Y.; Dai, F.; Liu, Y.; Wang, Q.; Zheng, J.; Xu, L.; Holmes, E. C.; Zhang, Y. A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* **2020**, *579*, 265-269.
4. Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F. A Novel Coronavirus Outbreak of Global Health Concern. *Lancet* **2020**, *395*, 470-473.
 5. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; Cheng, Z.; Yu, T.; Xia, J.; Wei, Y.; Wu, W.; Xie, X.; Yin, W.; Li, H.; Liu, M.; Xiao, Y.; Gao, H.; Guo, L.; Xie, J.; Wang, G.; Jiang, R.; Gao, Z.; Jin, Q.; Wang, J.; Cao, B. Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497-506.
 6. Nanshan Chen; Min Zhou; Xuan Dong; Jieming Qu; Fengyun Gong; Yang Han; Yang Qiu; Jingli Wang; Ying Liu; Yuan Wei; Jia'an Xia; Ting Yu; Xinxin Zhang; Zhang, L. Epidemiological and Clinical Characteristics of 99 Cases of 2019 Novel Coronavirus Pneumonia in Wuhan, China: a Descriptive Study. *Lancet* **2020**, *395*, 507-513.
 7. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K. S. M.; Lau, E. H. Y.; Wong, J. Y.; Xing, X.; Xiang, N.; Wu, Y.; Li, C.; Chen, Q.; Li, D.; Liu, T.; Zhao, J.; Li, M.; Tu, W.; Chen, C.; Jin, L.; Yang, R.; Wang, Q.; Zhou, S.; Wang, R.; Liu, H.; Luo, Y.; Liu, Y.; Shao, G.; Li, H.; Tao, Z.; Yang, Y.; Deng, Z.; Liu, B.; Ma, Z.; Zhang, Y.; Shi, G.; Lam, T. T. Y.; Wu, J. T. K.; Gao, G. F.; Cowling, B. J.; Yang, B.; Leung, G. M.; Feng, Z. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **2020**.
 8. Song, W.; Gui, M.; Wang, X.; Xiang, Y. Cryo-EM Structure of the SARS Coronavirus Spike Glycoprotein in Complex with Its Host Cell Receptor ACE2. *PLoS Pathog.* **2018**, *14*, e1007236.
 9. Li F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* **2016**, *3*, 237-261.
 10. Bosch, B. J.; van der Zee, R.; de Haan, C. A.; Rottier, P. J. The Coronavirus Spike Protein is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *J. Virol.* **2003**, *77*, 8801-8811.
 11. Walls, A. C.; Tortorici, M. A.; Snijder, J.; Xiong, X.; Bosch, B. J.; Rey, F. A.; Velesler, D. Tectonic Conformational Changes of a Coronavirus Spike Glycoprotein Promote Membrane Fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 11157-11162.
 12. Belouzard, S.; Millet, J. K.; Licitra, B. N.; Whittaker, G. R. Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. *Viruses* **2012**, *4*, 1011-1033.
 13. Jesper, P.; Nianshuang, W.; Kizzmekia, S. C.; Daniel, W.; Robert, N. K.; Hannah, L. T. Immunogenicity and Structures of a Rationally Designed Prefusion MERS-CoV Spike Antigen. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 7348-7357.
 14. Gui, M.; Song, W.; Zhou, H.; Xu, J.; Chen, S.; Xiang, Y.; Wang, X. Cryo-electron Microscopy Structures of the SARS-CoV Spike Glycoprotein Reveal a

- Prerequisite Conformational State for Receptor Binding. *Cell. Res.* **2017**, *27*, 119-129.
15. Walls, A. C.; Xiong, X.; Park, Y. J.; Tortorici, M. A.; Snijder, J.; Quispe, J.; Cameroni, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.; Zambon, M.; Rey, F. A.; Corti, D.; Veesler, D. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **2019**, *176*, 1026-1039.
 16. Yuan, Y.; Cao, D.; Zhang, Y.; Ma, J.; Qi, J.; Wang, Q.; Lu, G.; Wu, Y.; Yan, J.; Shi, Y.; Zhang, X.; Gao, G. F. Cryo-EM Structures of MERS-CoV and SARS-CoV Spike Glycoproteins Reveal the Dynamic Receptor Binding Domains. *Nat. Commun.* **2017**, *8*, 15092.
 17. Shulla, A.; Heald-Sargent, T.; Subramanya, G.; Zhao, J.; Perlman, S.; Gallagher, T. A Transmembrane Serine Protease is Linked to the Severe Acute Respiratory Syndrome Coronavirus Receptor and Activates Virus Entry. *J. Virol.* **2011**, *85*, 873-882.
 18. Matsuyama, S.; Nagata, N.; Shirato, K.; Kawase, M.; Takeda, M.; Taguchi, F. Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2. *J. Virol.* **2010**, *84*, 12658-12664.
 19. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889-897.
 20. Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate–DNA Helices. *J. Am. Chem. Soc.* **1998**, *120*, 9401-9409.
 21. Wang, J.; Shao, Q.; Xu, Z.; Liu, Y.; Yang, Z.; Cossins, B. P.; Jiang, H.; Chen, K.; Shi, J.; Zhu, W. Exploring Transition Pathway and Free-Energy Profile of Large-Scale Protein Conformational Change by Combining Normal Mode Analysis and Umbrella Sampling Molecular Dynamics. *J. Phys. Chem. B* **2013**, *118*, 134-143.
 22. Wang, J.; Peng, S.; Cossins, B. P.; Liao, X.; Chen, K.; Shao, Q.; Zhu, X.; Shi, J.; Zhu, W. Mapping Central Alpha-Helix Linker Mediated Conformational Transition Pathway of Calmodulin via Simple Computational Approach. *J. Phys. Chem. B* **2014**, *118*, 9677-9685.
 23. Burley, S. K.; Berman, H. M.; Bhikadiya, C.; Bi, C.; Chen, L.; Di Costanzo, L.; Christie, C.; Dalenberg, K.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ghosh, S.; Goodsell, D. S.; Green, R. K.; Guranović, V.; Guzenko, D.; Hudson, B. P.; Kalro, T.; Liang, Y.; Lowe, R.; Namkoong, H.; Peisach, E.; Periskova, I.; Prlić, A.; Randle, C.; Rose, A.; Rose, P.; Sala, R.; Sekharan, M.; Shao, C.; Tan, L.; Tao, Y.; Valasatava, Y.; Voigt, M.; Westbrook, J.; Woo, J.; Yang, H.; Young, J.; Zhuravleva, M.; Zardecki, C. RCSB Protein Data Bank: Biological Macromolecular Structures Enabling Research and Education in Fundamental Biology, Biomedicine, Biotechnology and Energy. *Nucleic Acids Res.* **2018**, *47*,

D464-D474.

24. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic. Acids Res.* **2018**, 46, W296-W303.
25. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, 24, 1999-2012.
26. Ryckaert, J-P.; Ciccotti, G. Numerical Integration of the Cartesian Equations of Motion of a System With Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, 23, 327-341.
27. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, 103, 8577-8593.
28. Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins* **2004**, 55, 383-394.
29. Sanner, M. F.; Olson, A. J. Reduced Surface: an Efficient Way to Compute Molecular Surfaces. *Biopolymers* **1996**, 38, 305-320.
30. Chen, Z.; Zhang, X.; Peng, C.; Wang, J.; Xu, Z.; Chen, K.; Zhu, W. D3Pockets: A Method and Web Server for Systematic Analysis of Protein Pocket Dynamics. *J. Chem. Inf. Model.* **2019**, 59, 3353-3358.
31. Kirchdoerfer, R. N.; Wang, N.; Pallesen, J.; Wrapp, D.; Turner, H. L.; Cottrell, C. A.; Corbett, K. S.; Graham, B. S.; McLellan, J. S.; Ward, A. B. Stabilized Coronavirus Spikes are Resistant to Conformational Changes Induced by Receptor Recognition or Proteolysis. *Sci. Rep.* **2018**, 8, 15701.
32. Chen, F.; Liu, H.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 6. Capability to Predict Protein-Protein Binding Free Energies and Re-Rank Binding Poses Generated by Protein-Protein Docking. *Phys. Chem. Chem. Phys.* **2016**, 18, 22129-22139.
33. 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 Spike in the Prefusion Conformation. *Sciences* **2020**, 367, 1260-1263.
34. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Crystal Structure of the 2019-ncov Spike Receptor-Binding Domain Bound with the ACE2 Receptor. *BioRxiv* **2020**.
35. Yan, R.; Zhang, Y.; Guo, Y.; Xia, L.; Zhou, Q. Structural Basis for the Recognition of the 2019-ncov by Human ACE2. *BioRxiv* **2020**.
36. Wan, Y.; Shang, J.; Graham, R.; Baric, R. S.; Li, F. Receptor Recognition by Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS. *J. Virol.* **2020**, 94, e00127-20.

For Table of Contents Use Only

