In silico study of intrinsic dynamics of full-length apo-ACE2 and RBD-ACE2 complex

Qing Liu, Yuwei Wang, Elaine Lai-Han Leung, Xiaojun Yao *

Dr. Neher's Biophysics Laboratory for Innovative Drug Discovery, State Key Laboratory of Quality Research in Chinese Medicine, Macau Institute for Applied Research in Medicine and Health, Macau University of Science and Technology, Taipa, Macau, China

A R T I C L E   I N F O

A B S T R A C T

The key step for SARS-CoV-2 to infect human cells is the membrane fusion triggered by the binding of the viral extracellular Spike protein to the human extracellular receptor, the angiotensin-converting enzyme 2 (ACE2). Although the Cryo-electron microscopy (Cryo-EM) uncovered the static atomic details of ACE2 homodimers, there is still a lack of research on the kinetic and thermodynamic properties of these full-length structures. This information is helpful to understand and interpret the role of ACE2 in the cell entry of SARS-CoV-2. In order to obtain this information, we performed microsecond-scale conventional and accelerated molecular dynamics (MD) simulations of full-length all-atomic systems of the RBD-ACE2 complex, the normal and torsional conformations of the apo-ACE2 homodimer. The comparative analysis of these systems showed that there were differences in their allosteric signal pathways and motion trends. These results may be helpful to further explore the cell entry mechanism of SARS-CoV-2. Moreover, the binding free energy and hydrogen bond distribution analysis of RBD-ACE2 binding interface provided the binding motifs that may be critical to allosteric signal transmission and RBD binding. These multi-conformational binding motifs can be used as targets or templates for the inhibitor design of the cell entry of SARS-CoV-2.

© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has led to millions of deaths worldwide. The cell entry of the virus is triggered by the binding of viral extracellular spike glycoprotein, to the angiotensin-converting enzyme 2 (ACE2), a human extracellular receptor [1]. Previous studies have shown that spike protein bound to ACE2 is proteolytic processed by the transmembrane protease serine 2 (TMPRSS2) from the host, and then undergo multiple metastable conformations [2,3]. Because of this, Cryo-electron microscopy (Cryo-EM) and other experimental methods are difficult to uncover these metastable three-dimensional structures. The model based on biochemical analysis suggests that the bound spike protein had approached the host cell membrane before being proteolytic processed, so as to prepare for the membrane fusion [4]. The mechanism of this motion is unclear. Therefore, the molecular mechanism of the viral cell entry mediated by the spike protein association with ACE2 needs to be further studied.

Illustrating the mechanism requires a clear understanding of the structural and functional properties of the spike protein and ACE2. Compared with a large number of studies on the spike protein, there is little discussion on the intrinsic dynamics of ACE2 based on its full-length structures. The Cryo-EM structures of apo-ACE2 and RBD-ACE2 complexes show that the full-length ACE2 is a homodimer; both protomers consist of a large claw-like extracellular head domain and the transmembrane domain (TMD); the head domain is divided into the zinc-binding peptidase domain (PD) and the smaller neck domain (Neck); the spike protein binds to the PD with its receptor binding domain (RBD); the apo-ACE2 manifests in both normal and torsional conformations; the spacing of the two PDs in ACE2 is sufficient to allow simultaneous binding of the two spike proteins (Fig. 1) [5]. However, these static structures cannot display kinetic and thermodynamic properties depending on the two conformations of apo-ACE2 and RBD binding. So, it is necessary to study the intrinsic dynamics related to different conformations of ACE2 homodimers.

In this study, the full-length all-atomic systems of ACE2 were constructed for the RBD-ACE2 complex (PDB ID: 6M17), the normal (PDB ID: 6M18) and torsional (PDB ID: 6M1D) conformation of apo-ACE2 homodimers [5]. The conventional and accelerated

* Corresponding author.
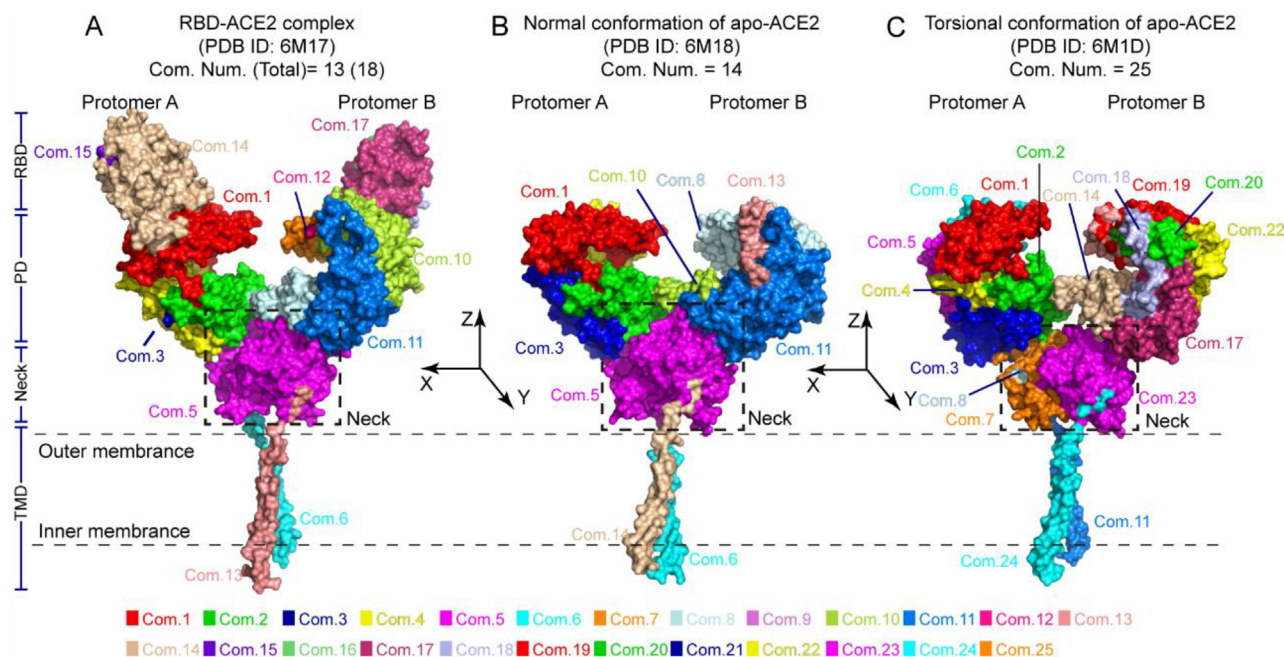   *E-mail address:* xjyao@must.edu.mo (X. Yao).

Fig. 1. Communities (com.) distribution on the RBD-ACE2 complex (normal conformation), normal and torsional conformations of apo-ACE2. ACE2 and RBD are displayed in the surface model. Communities are labeled and show up in different colors. TMD and Neck are marked by dotted lines.

molecular dynamics (MD) simulations were performed for these systems to generate microsecond-scale trajectories (Table S1) [6]. Based on these trajectories, dynamical network analysis and Markov state model (MSM) were implied to analyze and compare the conformational dynamics of ACE2 in terms of RBD binding and different apo-ACE2 conformations [7]. These analyses displayed the differences in the allosteric signal pathways and motion trends of these systems, which may be related to the cell entry of SARS-CoV-2. Combined with the molecular mechanics/generalized Born surface area (MM/GBSA), above MD trajectories were also used to calculate the binding free energy per residue and hydrogen bonds (H-bonds) distribution of the RBD-ACE2 interface [8,9]. This analysis found out the binding motifs related to the allosteric signal transmission and the RBD binding affinity. These multi-conformational binding motifs can be used as targets or templates for the blocker design of SARS-Cov-2 cell entry.

2. Materials and methods

2.1. Construction of ACE2 homodimer systems

In this study, the aqueous solution systems of the RBD-ACE2 complex (PDB ID: 6M17), and two apo-ACE2 homodimers (PDB ID: 6M18, 6M1D) were constructed from the corresponding Cryo-EM structures, respectively [5]. In these structures, the ACE2 homodimer of the RBD-ACE2 complex showed the same conformation as one of two apo-ACE2 dimers (RMSD < 2 Å; PDB ID: 6M18). For the convenience of later description, this conformation is called normal conformation of apo-ACE2 in this study. However, in the other apo-ACE2 homodimer (PDB ID: 6M1D), the PD has a counterclockwise rotation of about 12.1 degrees relative to the normal one, and there was no zinc ions in the zinc-binding pocket. So, this conformation is called torsional conformation in this study. In all above structures, ACE2 acts as the chaperone of the amino acid transporter, B^0AT1. With the assistance of B^0AT1, ACE2 homodimers form stably to be analyzed by Cryo-EM [5]. Therefore, in order to maintain the structure of ACE2 homodimers in subsequent

molecular dynamics simulations, B^0AT1 was retained in the three systems constructed in this study.

Metal ions in metalloproteins often play an important role in the function of proteins. However, the zinc ions in the normal ACE2 conformation is missing in the zinc-binding pocket of the torsional one. In order to ensure the comparability between the normal and the torsional conformations, zinc ions were added into the PDs of the torsional conformation by superimposing the normal PD (PDB ID: 6M0J) onto the torsional one in this study [10].

In the three structures, missing residues were made up with Modeller9.11 [11]. The addition of hydrogen atoms and the minimization of energy in the vacuum were performed by the Protein Prepare Wizard module in the Maestro [12]. These three systems were placed in a cubic water box containing 0.15 M NaCl aqueous solution. The minimum distance between the solute and the box boundary is 12 Å. The axis of symmetry of the ACE2 homodimer is parallel to the longest edge of the water box, namely the Z axis. The X-Y plane is the plane in which the cell membrane lies. The ACE2 homodimers were approximately perpendicular to the X-Y plane, and the two protomers were abreast of each other along the Y-axis. Protein and aqueous solutions were described by the AMBER-FF14SB all-atomic force field and the TIP3P water model, respectively [13,14]. The construction of all three ACE2 homodimer systems were done using the LEaP program in the AmberTools18 package [15].

2.2. Molecular dynamics simulation

In this study, microsecond-scale conventional and accelerated MD simulations of the above three systems were performed using the Amber18 software package (Table S1) [15]. Above three systems were minimized and equilibrated 2 ns locally under NPT conditions. After the equilibrium with the coordinate constraint, the production MD simulations were performed. The temperature and pressure of the systems were maintained at 310 K and 1 atmosphere, respectively. The coordinates, energy, temperature and other parameters of the trajectory were output once every 1 ns.

The covalent bonds containing hydrogen atoms were constrained by Shake algorithm. The simulated step size was 2 fs. Periodic boundary conditions were used in the system.

Compared with the conventional MD simulations, accelerated MD (aMD) simulations sample the conformational space more efficiently by adding the robust boosting potential energy to the system [6]. The aMD parameters were calculated from results of conventional MD simulations. Because the biased potential energy surface in aMD simulation can still fit the normal potential energy surface well, the aMD trajectories can be directly used for subsequent analysis without the weighted processing [16].

2.3. Dynamical network analysis

Dynamical network analysis was implied to investigate the allosteric signal pathways of above three systems. The conventional and accelerated MD trajectories of these systems were dried and aligned to the corresponding Cryo-EM structures. These aligned trajectories were used to perform the dynamical network analysis by the NetworkView plug-in in VMD. In this analysis, only heavy atoms were considered, and each residue or metal ion acted as a node. Each node was located by the coordinates of its alpha-carbon atom. If the distance between nodes was less than 4.5 Å during the 75% of the simulation time, then there was an edge connecting the two nodes. The end-to-end edges between two nodes formed the pathways. For every node pair, the shortest one was the optimal pathway [17]. The weight of each edges was given by the correlation coefficient matrix. Based on this correlation matrix, the structures of the ACE2 homodimer were divided into many sub networks, namely communities. The correlation coefficients of residual motion within communities are positive and higher than those of inter-community residues [18]. Therefore, communities can act as motion units of the ACE2 homodimer.

2.4. Construction and verification of Markov state models

Markov state models (MSMs) were constructed to predict the long time scale dynamics of the ACE2 homodimer [7]. In this study, the metric of MSMs was the center of mass (COM) of communities given by the dynamic network analysis. After dimensionality reduction, the conformational space of each ACE2 homodimer system was divided into $70 \sim 150$ microstates. Then, the microstates were merged into metastates. This partition scheme was verified. Each metastate had 10 representative conformations. The average COM coordinates of communities of each metastates were used to calculate the motion correlation coefficients by the R-3.3.4 software package [19]. The construction and verification of MSMs were carried out by the PyEMMA-2.5.7 software packages [20].

2.5. Calculation of the bind free energy of RBD

In order to evaluate the binding affinity of RBD in all metastates, MM/GBSA was used to calculate the binding free energy between RBD and ACE2 [9]. All representative conformations were used to construct the aqueous solution systems of RBD-ACE2 complex. After the equilibrium under the NPT conditions, the last six frames were used for the MM/GBSA calculation. For each metastate, the top 10 residues with high binding-energy values were found and put into a union set. The binding-energy values of residues belonging to the set were used to draw the histogram of binding free energy per residue. The binding free energy calculation of MM/GBSA was carried out using the Amber18 software package [15].

3. Results

3.1. Conformational dynamics comparison between normal and torsional conformations of apo-ACE2

The structural differences between the torsional and normal conformations of apo-ACE2 come from their PDs. As shown in the Fig. 2B, the PD of the torsional conformation has a counter-clockwise rotation of about 12.1 degrees relative to the normal one. Through the structural alignment, zinc ions were artificially added into the zinc-binding pockets of the PD in the torsional conformation (see Section 2.1). In zinc-binding pockets of the normal conformation, the distances between zinc ions and heavy atoms of the five nearby residues are less than 4.5 Å, the distances at which two heavy atoms contact each other [21]. However, in the torsional conformation, the corresponding distances are greater than 4.5 Å. This suggests that the interactions between zinc ions and the zinc-binding pockets in the torsional conformation are weaker than those in the normal conformation. Actually, for most of the simulation time, zinc ions were outside the zinc-binding pockets of the torsional conformation. These results indicate that the zinc-binding pockets in the torsional conformation do not capture zinc ions as well as the normal ones.

Dynamical network analysis was implied to investigate the influences of above two conformations on the allosteric signal pathways of ACE2 homodimers [21]. In this analysis, the RBD-ACE2 complex, the normal and torsional conformations of apo-ACE2 are divided into 18, 14 and 25 communities, respectively (Fig. 1). The correlation coefficients of residual motion within communities are positive and higher than those of inter-community residues.

The more residues contained in a community are, the more residues in the high synergy motion are. As shown in the Fig. 1, the torsion conformation has nearly twice (25 vs. 14) as many communities as the normal one. So, compared with the normal conformation, the average size of the communities in the torsional conformation system is smaller. That is, the residues of the torsional conformation have a weak tendency of the highly coordinated movement.

According to the concepts of dynamical network analysis, there is an edge connecting two residues if the heavy atomic distance between these residues is less than 4.5 Å (the distance required to form the atomic contact) over 75% of the simulation time [21]. Naturally, the allosteric signals should travel along these edges. The strength of the allosteric signal within community is positively correlated with the absolute value of the corresponding correlation coefficients. Compared with the inter-community residue pairs, residue pairs within the same communities are more likely to form atomic contacts and transmit allosteric signals due to their high positive correlation coefficients. This leads to more edges and easier signal transmitting within communities than between them. As shown in Figs. 1B and 6A, there are only one community and 23 cross-protomer edges in the Neck of the normal conformation. In the torsional conformation, the Neck domain splits into two communities (Fig. 1C), and the number of cross-protomer edges decreases to 7 (Fig. 6A). Therefore, the community can be regarded as the motion unit of the ACE2 homodimer, and the signal transmission within the community is easier than that between the communities.

The motion correlation coefficients of residue pairs can also be considered as the ability of one residue to influence the motion of the other. The influence is exerted in a way that one residue transmits its motion signal to the other. The edges along which the motion signal passes constitute the motion signal pathways. The shortest of them is called the optimal motion signal pathway.
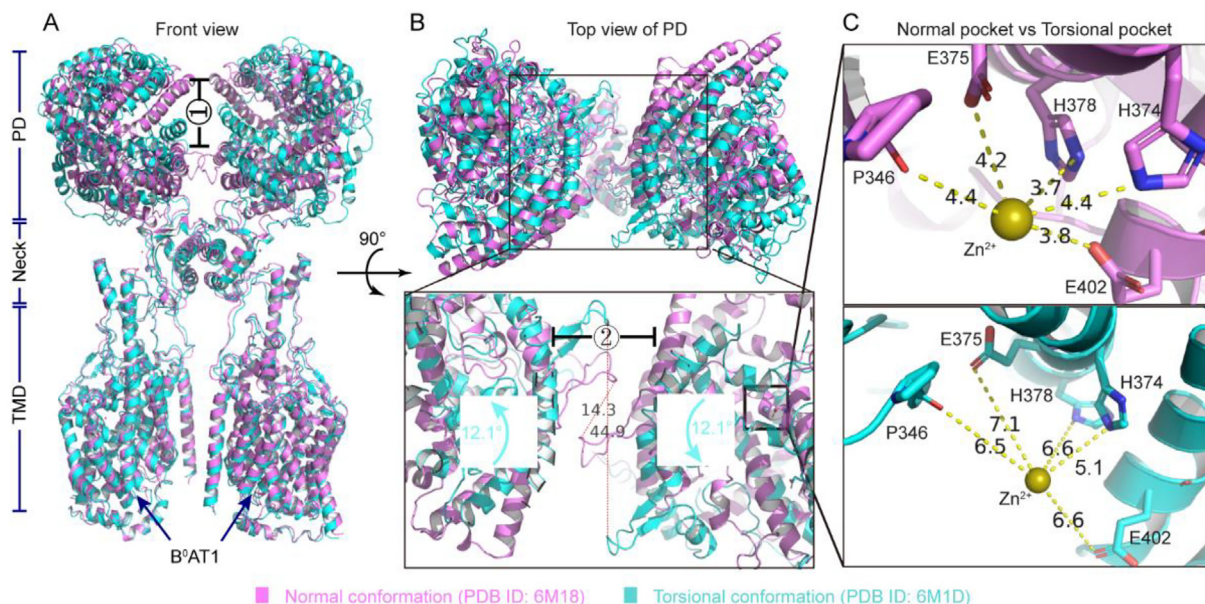
Fig. 2. Comparison of normal and torsional ACE2 conformations. (A) Front view of the two conformations, and the distance between the upper and lower boundaries of the zinc-binding pocket is marked as ①. (B) Top view of PDs and the torsional angles; The distance between the upper halves of the two PD's is marked as ②. (C) Zinc-binding pocket. Cyan: normal conformation of apo-ACE2. Violet: torsional conformation of apo-ACE2. Yellow balls: zinc ions. Yellow dotted lines: distances between zinc ions and heavy atoms of adjacent residues. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
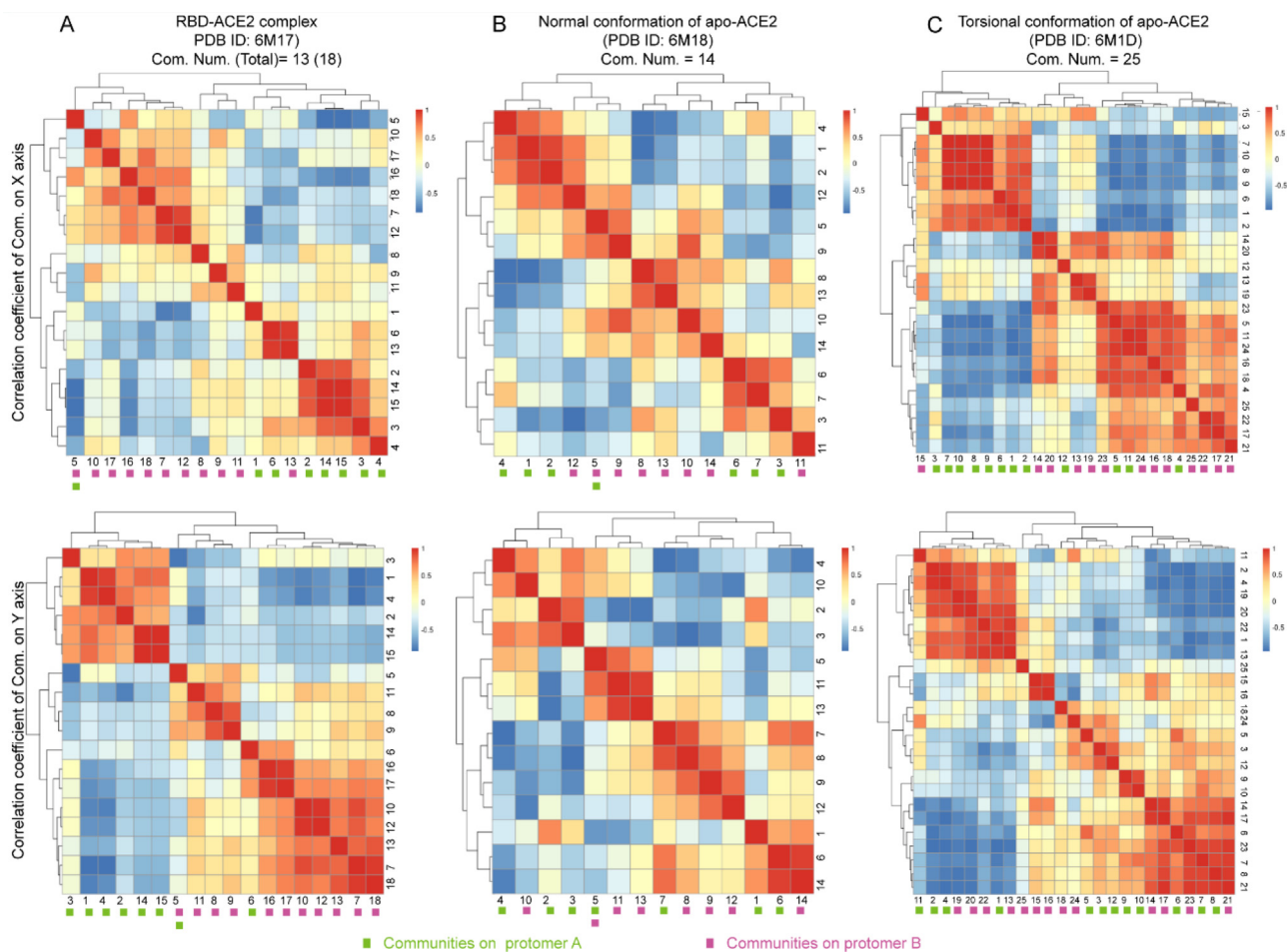

Fig. 3. Motion correlation coefficients of the COM of communities for RBD-ACE2 complex (A), normal (B) and torsional (C) ACE2 conformations. Top and bottom panels: motions correlation coefficients along X and Y axes, respectively. Green squares: communities on the protomer A. Violet squares: communities on the protomer B. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The allosteric effect can be regarded as the transmission of motion signals from the ligand-bound allosteric site to some distant sites. As described below, the allosteric effects also have optimal motion signal pathways. Therefore, signal pathways refer to the optimal allosteric signal pathways in the following.

The study of Emilia Barros et al. showed that the bending of the head domain of ACE2 towards the cell membrane was accompanied by the tilted TMD [22]. Since the RBD of spike protein binds to PD, the bending head domain of ACE2 also brings the RBD closer to the cell membrane. This may facilitate the direct contact between the latter two. So, it seems that the bending of ACE2 may be one of the pulling forces of the spike protein towards the cell membrane. Since this bending is accomplished cooperatively by PD and TMD, there should be signal pathways of the allosteric effect from PD to TMD. Finding these signal pathways may help us to understand the influences of different ACE2 conformations on the allosteric signal transmission and the allosteric effect. As described in Section 3.2, the RBD-ACE2 interface contains critical residues which are responsible for transmitting inter-molecular allosteric signals from RBD to PD (Fig. 4A and C). So, if signal pathways connecting PD and TMD exist, one terminal of these paths should be these critical residues. Dynamical network analysis shows that there are signal pathways from critical residues to two TMDs in the normal conformation (Fig. 6B, C and S1). In the torsional conformation, such signal pathways do not exist (Fig. 6F). The reason for this may be that the number of communities in the torsional conformation is almost as twice many as that of the normal conformation (Fig. 1B and C). As stated previously, the signal transmission between communities is more difficult than that between the communities. Therefore, compared with

the normal conformation, the transmission of allosteric signal is more difficult, which may result in motion differences between the two ACE2 conformations.

The MSM construction was carried out to study the overall motion trend of the apo-ACE2 homodimer [7]. Since the communities can be regarded as the motion unit, the motion of the apo-ACE2 homodimer can be decomposed into the motion of the communities on the three Cartesian axes. In this study, the communities are represented as their center of mass (COM). So, the COM coordinates of the communities are selected as the metric of MSMs. As shown in the Figs. S2 and S3, both the systems of the normal and torsional conformation systems are divided into 12 metastates with a variety of proportions. In the MSM of the torsional conformation, because zinc ions are outside pocket during most of the simulation time, each zinc ion acts as a community alone. If zinc ions were ignored during the MSM construction, the metastates number of the torsional conformation system is smaller than the normal one (10 vs 12). Therefore, the torsional conformation has less conformational diversity than the normal one.

Since the metric of MSMs in this study are the COM coordinates of communities, the motion trends of the communities can be described by the motion correlation coefficients of these COMs. These coefficients are calculated based on the COM coordinates of the communities in all metastates. As shown in the Fig. 3, at the bottom of each correlation coefficient panel, the communities belonging to the protomer A and B are marked with violet and green squares, respectively. The positions of all communities in these panels were rearranged according to the motion correlation coefficients. The higher the correlation coefficients are, the closer
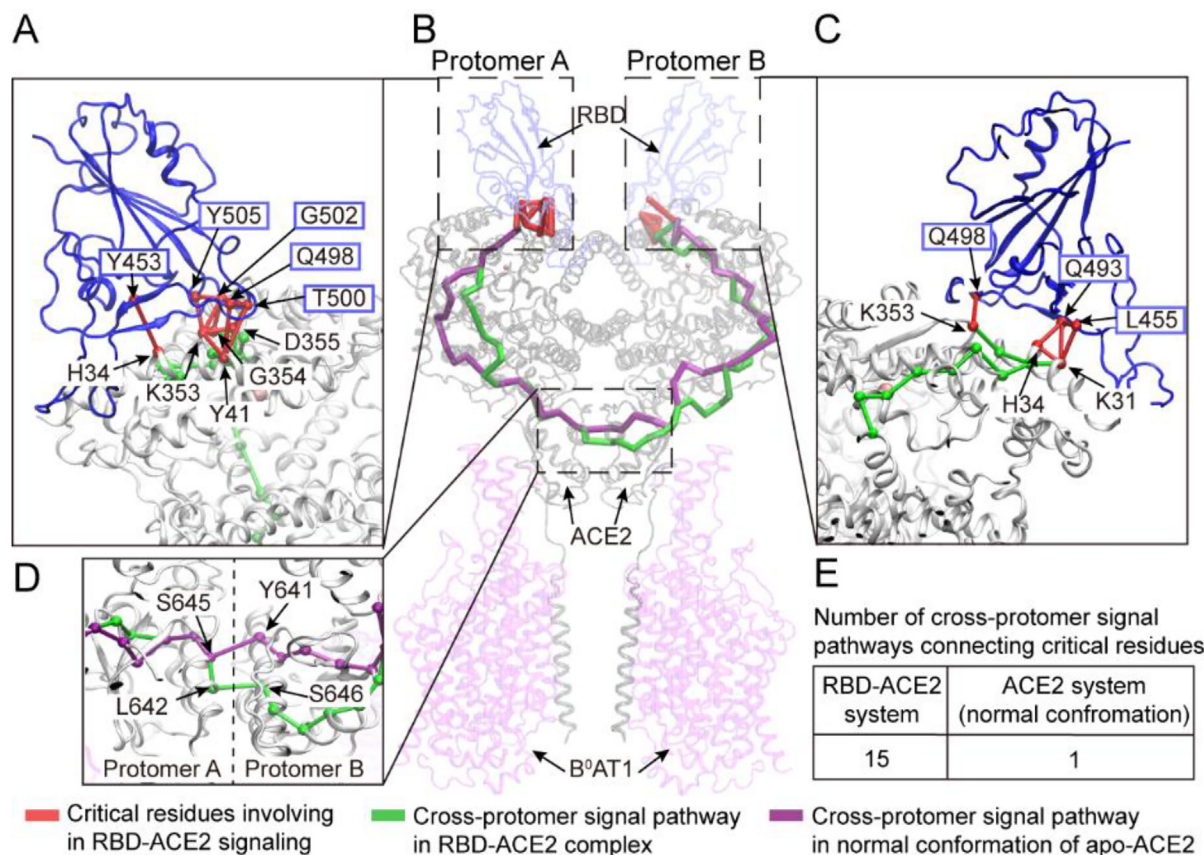

Fig. 4. (A, C) Critical residues responsible for the allosteric signal transmission between RBD and PD in protomer A (A) and B (C). (B) Signal pathways connecting critical residues belonging to different protomers. (D) Residues responsible for cross-protomer signaling in Neck. (E) Number of cross-protomer signal pathways connecting critical residues.

the aggregation of the corresponding communities is. In this way, the higher the aggregation degrees of the same color squares are, the more consistent the motion trend of communities on the same protomer is. On the X-axis of the normal conformation, the violet squares except for that representing the community 11 are clustered on both sides (Fig. 3B). This means that in the normal conformation, communities belonging to the same protomer have the tendency of the highly coordinated movement along the X-axis. At the same time, the motion correlation coefficients of interprotomer communities were negative. This suggests that the two protomers of the normal conformation of apo-ACE2 move in opposite directions along the X-axis. On the Y axis, communities of the normal conformational are less consistent in their motions. Therefore, the tendency of the normal conformational system is to move back and forth along the X-axis mainly.

The communities of the torsional conformation have the low motion consistency on both the X and Y axes (Fig. 3C). So the torsional conformation has no obvious motion trend on both axes. In this study, the open state is defined when no heavy atomic contact forms between the upper halves of the two PDs and between the upper and lower boundaries of the zinc-binding pockets (Fig. 2A and B). Because of the low consistency of the community motion, it is difficult for the torsional conformation to ensure that their communities meet the requirements of the open-state definition at the same time. Moreover, during the simulations of the apo-torsional conformation, the protomer B almost maintained its tilt angle more than the cryo-EM value (99.05%; Table 1), and two PDs have the closest contact in three ACE2 systems (73.10%; Table 1, Fig. 5C). This indicates that the torsional conformation has basically no significant motions along the X and Y axes. So, the torsional conformation has the almost zero percentage of open states (Fig. S3). It can be seen from the following text that open states may be more conducive to two-RBD binding than closed states. So, the torsional ACE2 conformation may not be a good receptor for SARS-CoV-2.

Compared with the normal conformation, the torsional ACE2 conformation is weaker in capturing zinc ions, transmitting allosteric signals, and the highly coordinated movement.

3.2. Conformational dynamics comparison between RBD-ACE2 complex and normal conformation of apo-ACE2

Like apo-ACE2 systems, the system of the RBD-ACE2 complex is divided into 13 metastates in its MSM (Fig. S4). Among them, the open states accounted for 57.1% of the conformational space, which is much higher than those of two apo-ACE2 systems. So, the RBD binding significantly increases the proportion of the open states in MSMs. As shown in Fig. 3A, the communities belonging to the same protomer have a high motion consistency on both the X and Y axes; the communities belonging to different protomers have negative correlation coefficients on both X and Y axes.

These results indicate that the two protomers of the RBD-ACE2 complex have a high tendency of the opening and closing movement on both X and Y axes. This high motion consistency increases the probability that all communities belonging to the upper halves of PDs keep longer distances than 4.5 Å. As shown in Fig. 5C and Table 1, RBD-ACE2 complex has the highest ratio of the PD-PD distance more than the cryo-EM value. As a result, the open-state proportion of RBD-ACE2 complex is highest among all three systems in this study (57.1%, Fig. S4). However, this motion is not found in two apo-ACE2 systems. This implies that the open state may be more suitable for the RBD binding than the closed one. From the structural point of view, the bending of ACE2 toward the cell membrane is also along the Y axis (Fig. 1B). Therefore, the binding of RBD may enhance the opening and closing movement of ACE2 along the Y axis, to increase the proportion of the open state and the bending amplitude of ACE2.

Like the apo-ACE2 homodimer, the RBD-ACE2 complex is also symmetrical about the z-axis (Fig. 1A). ACE2 in this complex shows the normal conformation. The dynamical network analysis shows that there are 10 critical residues responsible for transmitting allosteric effect signals at the RBD-ACE2 interface of the protomer A, of which H34, Y41, K353, G354, D355 belong to PD, and Y453, Y505, G502, Q498, T500 belong to RBD (Fig. 4A); there are 6 critical residues responsible for transmitting allosteric effect signals on the RBD-ACE2 interface of the protomer B, of which H34, K31, K353 belong to PD, and L455, Q493, Q498 belong to RBD (Fig. 4C). Therefore, although the RBD-ACE2 complex is Z-symmetric, the critical residues of the RBD-ACE2 interface are asymmetrically distributed in the two protomers.

The asymmetry also occurs in the allosteric signal pathways. Since the critical residues are responsible for the allosteric signal transmission between RBD and ACE2, these residues should be terminals of signal pathways in the ACE2 homodimer. Fig. 4B shows the allosteric signal pathways between the critical residues belonging to two protomers.

These signal pathways are almost identical except both terminals occupied by different critical residues. Compared with those in the normal conformation of apo-ACE2, these signal pathways display a right-skewed shape in the RBD-ACE2 complex. The difference shape of signal pathways may related to different motions of the two systems. In the apo-normal conformation, protomer A and B have the almost same distributions of the tilt angles more or less the cryo-EM value; while protomer A (97.74%) is more bent than protomer B in the RBD-ACE2 complex (43.40%. Table 1, Fig. 5B). These results indicate that the RBD-binding results in the asymmetry of the motions of two PDs and the allosteric signal pathways between critical residues

In the Neck interface of two protomers, residues for the cross-protomer signaling of these asymmetrical signal pathways are also different between the RBD-ACE2 complex (S645 and Y641) and the normal conformation of apo-ACE2 (L642 and S646; Fig. 6D). For

Table 1
Ratios of tilt angles and PD-PD distances.

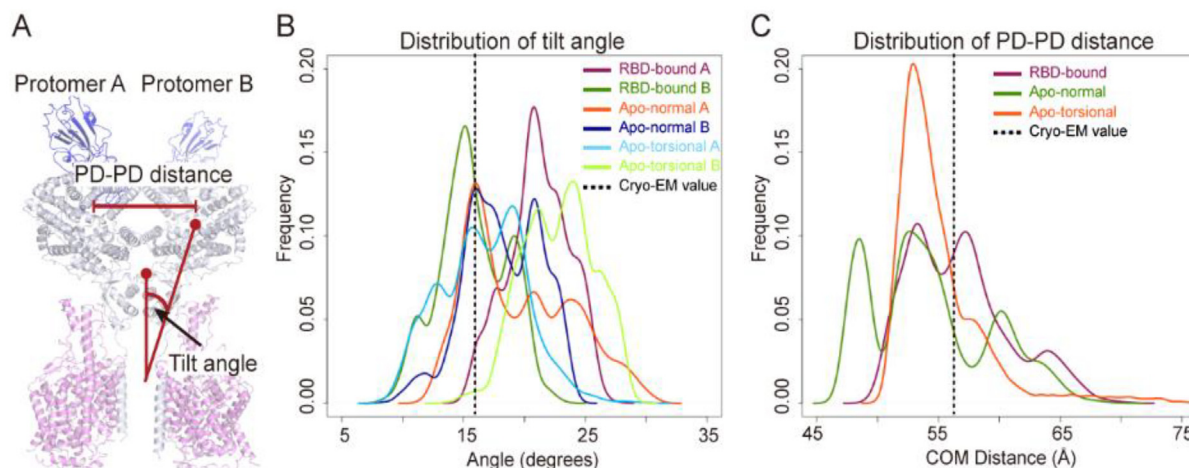| System names	Tilt angles				PD-PD distances	
	Ratios of angles less than cryo-EM values (%)		Ratios of angles more than cryo-EM values (%)		Ratios of distances less than cryo-EM values (%)	Ratios of distances more than cryo-EM values (%)
	Protomer A	Protomer B	Protomer A	Protomer B		
RBD-ACE2 complex	2.26	56.70	97.74	43.30	47.83	52.17
Apo-normal ACE2	25.42	25.32	74.58	74.68	68.88	31.12
Apo-torsional ACE2	42.50	0.95	57.50	99.05	73.40	26.60

Fig. 5. Distributions of tilt angles and PD-PD distances of RBD-ACE2 complex, normal and torsional ACE2 conformations. (A) Definitions of the tilt angle and PD-PD distance. (B) Distributions of tilt angles for both protomers of three ACE2 systems. (C) Distributions of PD-PD distances of three ACE2 systems. The Cryo-EM values are marked by dash black lines. PD-PD distance is the distance between the COMs of two PDs. The three endpoints of the tilt angle are the COMs of a single PD, two TMs and Neck.

any critical residue pairs belonging to different protomers, there is always a signal pathway connecting them in the RBD-ACE2 complex. However, in the normal conformation of apo-ACE2, there is only an allosteric signal pathway from G354 to K353 (Fig. 4E). These results indicate that the binding of RBD on the two PDs enhances the signal transmission between them.

Since the ACE2 bending is accompanied by the TM tilting, every critical residue should transmit allosteric signals to both TMDs. As show in Fig. 6B, 6C and S1, the parts of such signal pathways in head domains are the same as the signal pathways connecting critical residues described above. As a result, the signal pathways connecting critical residues and TMDs also presents a right-skewed shape. Among these paths, the residues (L642 and S646) for signal transmission at the Neck interface (Fig. 6D and E) are also the same as the signal pathways connecting critical residues (Fig. 4D). How-

ever, the signal pathways in the normal conformation of apo-ACE2 present a different situation. For the signal pathways connecting the critical residues of protomer B and TMD of protomer A, the cross-protomer signal is transmitting through Y649 and Y641 (Fig. 6D). For the signal pathways connecting the critical residues of protomer A and TMD of protomer B, the cross-protomer signal is transmitting through S645 and Y641 (Fig. 6E). This indicates that the binding of RBD makes the allosteric signal adopt a uniform pathway across the Neck interface.

In the RBD-ACE2 complex, the signal pathway connecting the critical residues to the TMD of the protomer B passes through B⁰AT1 (Fig. 6C, G, and S1). This kind of signal pathways enters B⁰AT1 from residue R678 at the junction of Neck and TMD, and then returns to ACE2 through residue F746 located in TMD (Fig. 6G and S1). This phenomenon does not occur in apo-ACE2.
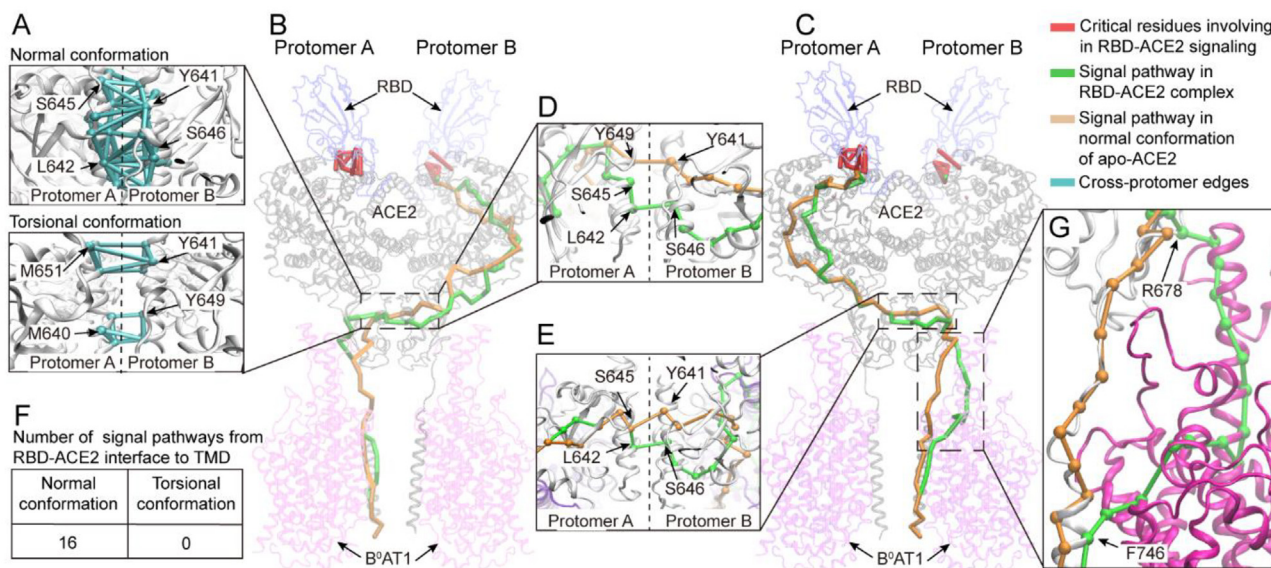

Fig. 6. (A) Cross-protomer edges in the Neck domain of normal and torsional conformations. (B, C) Signal pathways from the critical residues to the TMD of the other protomer in RBD-ACE2 complex and the normal conformation of apo-ACE2 homodimer. (D, E) Parts of signal pathways displayed in the panel B, C and residue for the cross-protomer signal transmission. (F) Number of signal pathways from the RBD-PD interface to TMD in normal and torsional conformations. (G) Signal pathway towards TMD goes through B⁰AT1. Red: critical residues and edges involving RBD-PD signaling. Green: signal pathways in RBD-ACE2 complex. Yellow: signal pathways in normal conformation of apo-ACE2 homodimer. Cyan: cross-protomer edges. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

As mentioned in Section 1 of Methods, B⁰AT1 mimics the role of the cell membrane in this study. This suggests that the cell membrane may also influence the conformational change of ACE2 by participating in allosteric signal transmission. In the study of Emilia Barros et al., a small loop of the head domain during the ACE2 bending extends into the cell membrane [22]. Although the loop does not contain R678, this phenomenon suggests that the head domain may be in direct contact with the cell membrane. This contact may be the structural basis for the transmission of allosteric signals to TMD via the cell membrane. So, it seems that both B⁰AT1 and cell membrane may participating in allosteric signal transmission from RBD-binding sites.

The ACE2 homodimer formed with two B⁰AT1 molecules display the different motion from that inserting into the cell membrane. The tilt angles vary from 0 to 60 degrees in the cell membrane environment, while the tilt angles are between 5 and 35 degrees in this study (Fig. 5B) [22]. This may be due to the fact that the B⁰AT1 is taller than the cell membrane and extends to the vicinity of the Neck domain (Fig. 2A), thus limiting the degree of bending of the PDs. As described above, in Protomer B, B⁰AT1 forms a stable atomic contact with the Neck domain to transmit allosteric signals. As a result, the motion of the protomer B is constrained so that its tilt angle is less than cryo-EM value during 56.7% of the simulation time (Table 1). On the other hand, the protomer A almost maintain the tilt angle greater than the cryo-EM value during the MD simulation (Table1). That is to say, Protomer A always bent away from Protomer B, and the degree of bending was much greater than that of Protomer B. This was not the case in the simulation of the membrane environment [22]. In addition, under the action of B⁰AT1, both the RBD-ACE2 complex and the apo-normal conformation, the span of PD-PD distances are larger than that of the membrane environment, and the ratios of PD-PD distances greater than the cryo-EM value is higher (Fig. 5C; Table 1). These

results suggest that B⁰AT1 and cell membrane may play different roles in RBD binding ACE2, and further research is needed.

In summary, although RBD-ACE2 complex has a Z-axis symmetric structure, the RBD binding resulted in the significant opening and closing movement along the X and Y axes, and significantly asymmetric allosteric signal pathways.

3.3. Analysis of hydrogen bonds and binding free energy of RBD-ACE2 interface

According to the concepts of dynamical network analysis, the critical residue pairs located at the RBD-ACE2 interface have to maintain the contact of heavy atoms during 75% of the simulation time at least. An important interaction that holds the residue pairs to enhance the molecular binding affinity is the hydrogen bond (H-bond). As shown in Fig. 7C and D, stable H-bonds do exist between some of the critical residues. For example, the H-bond between D355 and T500 in the protomer A lasted nearly half the simulation time (49.96%). At the same time, their adjacent critical residue pair, K353-G502 also forms a stable H-bond (45.08% of the simulation time; Fig. 7B, C, S10 and S11). With the help of their own H-bonds, these critical residue pairs form the binding motif to maintain their edges in the dynamical network. However, there are also some critical residues between which H-bonds last for a short time. For example, the H-bond between K353 and Y505 lasted only 1.75% of the simulation time in the protomer A (Fig. 7C). Meanwhile, K353 also forms H-bonds with G502 and G496 on RBD. These two H-bonds last 45.08% and 29.84% of the simulation time, respectively (Fig. 7C). This implies that heavy atomic contact between K353 and Y505 should be maintained by these adjacent H-bonds. This can lead to the formation of the edge connecting K353 and Y505 in the dynamical network analysis. The critical residue pair H34-Y453 has a similar condition: stable
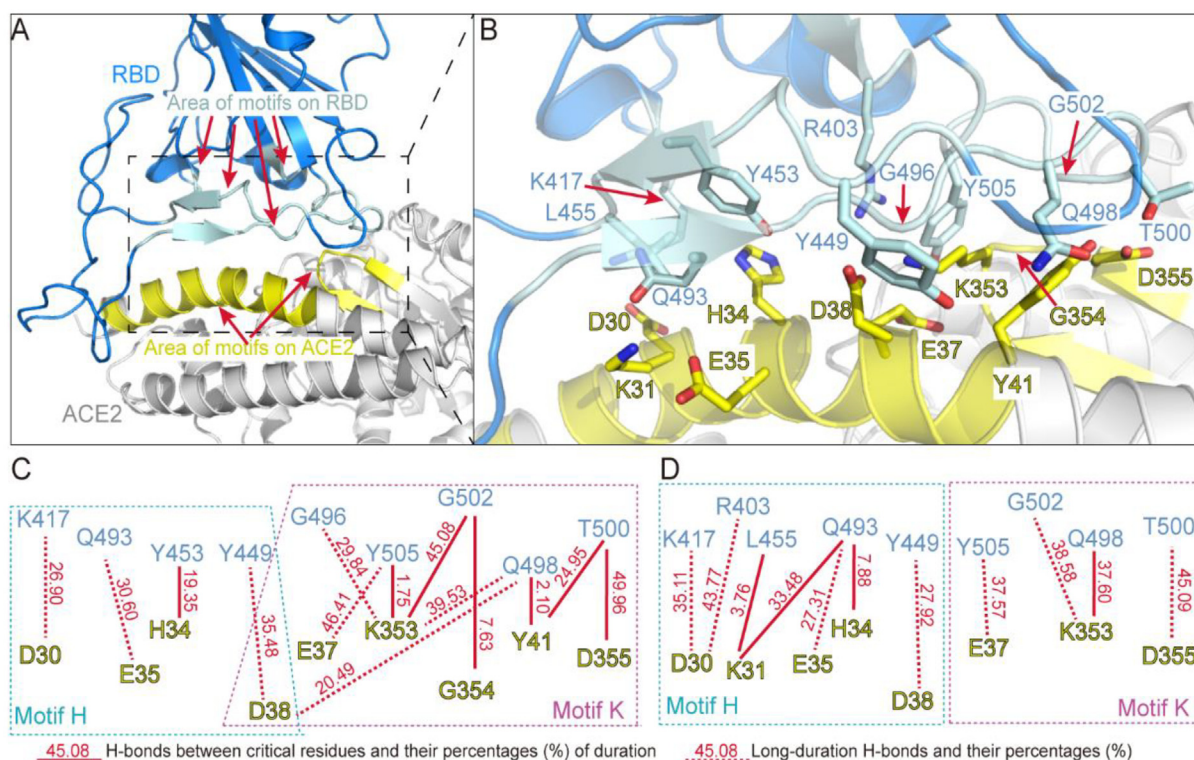

Fig. 7. Two types of binding motifs containing critical residues. (A) Areas of binding motifs (aquamarine and yellow areas) on the RBD-ACE2 interface. (B) Part of residues that form binding motifs (stick model). (C, D) H-bonds networks of binding motifs in the protomer A and B, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

H-bonds of residue pairs, D38-Y449 and E35-Q493 formed near it (Fig. 7C), although its own H-bond lasted only 19.35 % of the simulation time. In this way, these adjacent residues and critical residues also form the binding motifs. Therefore, the binding motifs containing critical residues contribute to the stability of the edges connecting RBD and PD.

According to the spatial distribution of critical residues (Fig. 4A and C), binding motifs can be classified into two types: those containing K353 (named Motif K) and those containing H34 (named Motif H). The area of binding motifs on ACE2 consists of an alpha-helix (residue 24 ∼ 41; Fig. 7A) and a beta-sheet (residue 350 ∼ 356; Fig. 7A). The former is a part of the mini-proteins that are potential inhibitors for RBD [23]. This helix displayed little conformational changes, whether or not RBD were bound (backbone RMSD < 2.0 Å; Figs. S6 and S7). The conformational changes of the binding motifs on ACE2 mainly comes from the relative displacement of the two parts (Figs. S6 and S7). The area of binding motifs on RBD consists of four sequences: residues 448 ∼ 455, 491 ∼ 506, 402 ∼ 404 and 416 ∼ 418. The conformation of this area is also relatively stable in most of the simulation time. The large conformation change comes from the movement of RBD toward the lateral side of ACE2 (Fig. S6). Compared with protomer A, RBD has a higher probability of such movement at the interface of protomer B (Figs. S6 and S7). So, the interfaces of two protomers presents the asymmetrical H-bond networks (Fig. 7C and D).

In the protomer A, the binding motif containing K353 did not form the hydrogen bond with Q498 in the Cryo-EM structure (Fig. S10A). However, in the simulated frame at 952 ns, the hydrogen bond of the residue pair K353-G502 breaks; Q498 and N501 in RBD move towards K353 to form new hydrogen bonds (Fig. S10B). For the binding motif containing H34, the hydrogen bonds of residue pairs D38-Q498 and H34-Q493 formed at 1550 ns (Fig. S10D), which were not present in the Cryo-EM structure (Fig. S10C). Moreover, the longest lasting hydrogen bond, D355-T500 is not actually present in the Cryo-EM structure (Fig. 7C and S11). The above situation also exists in the binding motifs of the protomer B (Fig. S12). Besides, at the RBD-ACE2 interface, none of the H-bonds lasts more than 75% of the simulation time (Fig. 7C and

D). Thus, the binding motifs containing critical residues change over time.

In order to show the residues with the high binding free energy on the RBD-ACE2 interface, MM/GBSA was used to calculate the binding free energy per residue of each metastate. As shown in Fig. 8 and S13, residues which contribute to the RBD affinity and have negative binding free energy in all 13 metastates are Y505, G496, Y489, F486 and F456; the residues which contribute to ACE2 affinity and have negative binding free energy in all 13 metastates were D355, Y83, Y41, D38, E37, F28, T27 and Q24. Among them, D355, Y83, E37 and F28 are highly conserved in bats, pangolins, minks, cat, dog and primates which can be naturally infected by SARS-CoV-2 [24,25]. The changes in the binding energy of Y41H, D38E and Q24L mutations were all within 2 kcal/mol [24,25]. So, with the exception of T27, such residues are stable sources of binding free energy of RBD and ACE2. Most of them involve in the formation of binding motifs. So, binding motifs are important for both allosteric signaling and RBD binding. In combination with the H-bond analysis above, some critical residues contribute less to the binding affinity, so they do not appear in Fig. 8. They are K353, G354, K31 in ACE2, and Y453, L455 in RBD. So, mutations in these residues should not affect RBD binding, but may interfere with motion signaling. Therefore, through comprehensive consideration of dynamic binding motifs and reasonable modification of these critical residues, it may be possible to design high affinity inhibitors that can interfere with Spike protein binding.

In summary, the dynamic binding motifs ensure the stable heavy atomic contacts of critical residue pairs and can be used as a templates or targets for the inhibitor design targeting RBD or ACE2.

4. Discussion

In this study, microsecond-scale conventional and accelerated MD simulations were performed for the full-length all-atomic systems constructed from the RBD-ACE2 complex, the normal and torsional conformations of apo-ACE2. Based on the long-time
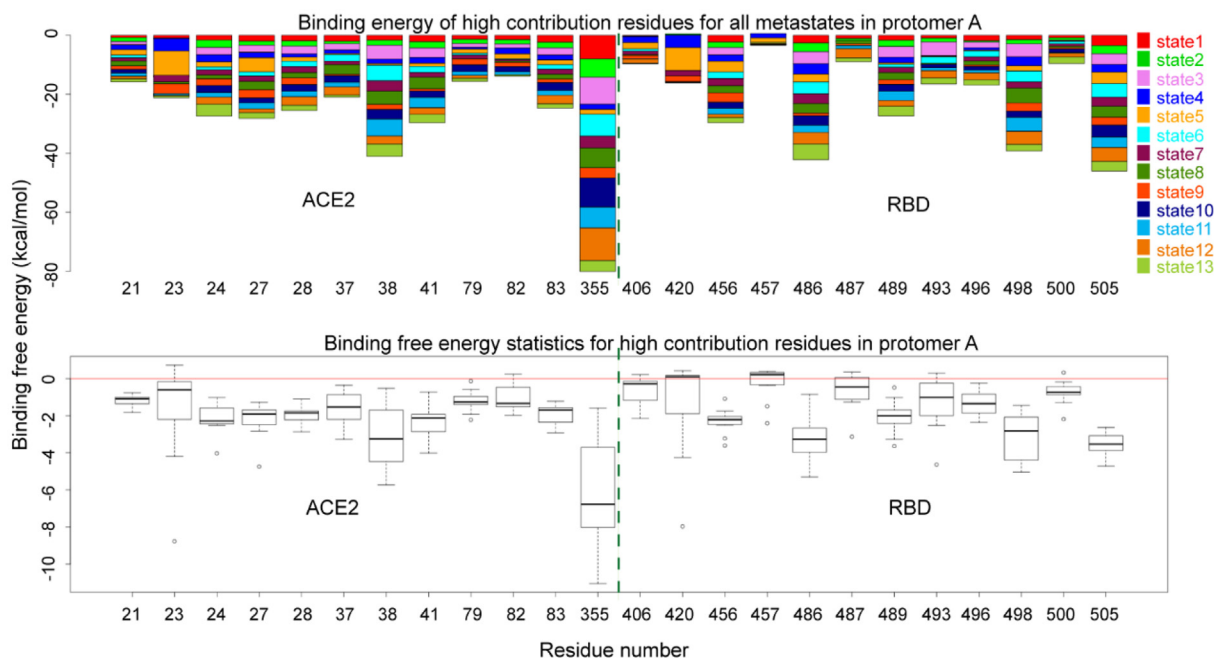

Fig. 8. Binding free energy and its statistics of high contribution residues for all metastates in protomer A. The values of binding free energy are 0 at the red horizontal line. Residues on the left of the green vertical dotted line belong to ACE2, and residues on the right belong to RBD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

trajectories, dynamical network analysis, MSM and MM/GBSA were used to analyze the allosteric signal transmission, ACE2 motion trend and binding free energy per residue at the RBD-ACE2 interface, respectively.

The conformational dynamics comparison between the normal and torsional conformations of apo-ACE2 shows that (1) the ability to trap zinc ions is weaker in the torsional conformation compared with the normal one; (2) the number of communities is doubled, which leads to difficulties in transmitting allosteric signals in the torsional conformation; (3) there is no significant motion trend on both X and Y axis for the torsional conformation; (4) compared with the normal one, The proportion of open states and the conformational diversity are lower in the torsional conformation. Combined with the results of RBD-ACE2 complex, these results suggest that the torsional conformation may not be suitable for the two-RBD binding.

The conformational dynamics comparison between the RBD-ACE2 complex and the normal apo-ACE2 homodimer showed that (1) RBD-ACE2 complex display a significant opening and closing movement along the X and Y axes, leading to an increase in the proportion of open states; (2) the RBD binding resulted in significant asymmetry of the critical residue distribution and allosteric signal pathways; (3) the RBD binding leads to the enhancement of the cross-protomer signal transmission and a uniform signal pathway in the Neck domain; (4) B^0AT1 may participating in allosteric signal transmission from RBD-binding sites; (5) B^0AT1 and cell membrane may play different roles in RBD binding ACE2. Despite these results, further research is needed to determine whether the bending of ACE2 promotes viral invasion of SARS-CoV-2.

At present, a feasible approach of designing drugs targeting ACE2 or RBD is to design small protein molecules or peptides with similar or stronger affinity than RBD or ACE2 based on the structure of the RBD-ACE2 interface, so as to achieve the purpose of competitive binding with RBD or ACE2 [23]. Compared with protein inhibitors, small peptides have little immunogenicity and are relatively easy to synthesize. The analysis of hydrogen bond and binding free energy per residue at the RBD-ACE2 interface showed that the binding motifs containing critical residues were important in both the signal transmission and the RBD binding. Therefore, these binding motifs can be used as targets for peptide inhibitors or fragments in the peptide library.

The simulated trajectories and MSMs of this study provide multiple conformations of these binding motifs in the aqueous environment. These conformations provide much richer atomic details of binding motifs than the static structures. At the same time, MM/GBSA calculations based on metastates of MSMs found more high affinity residues than the FEP analysis [26], and also gave the binding energy contribution of each residue at the RBD-ACE2 interface. These values vary with the different metastates. This information can be act as the basis for designing peptide targeting the specific conformation of binding motifs. Therefore, the peptide inhibitor design based on multiple conformations targeting ACE2 and RBD should have a higher success rate and reliability than the drug design based on a single structure.

5. Conclusion

In this study, molecular dynamics simulation and other computational biology methods were used to compare and analyze the conformational dynamics of RBD-ACE2 complex, the torsional and normal conformations of apo-ACE2 homodimer. In this way, differences in their allosteric signal pathways and motion trends were described. These works have improved our understanding of the interaction between RBD and ACE2, and may also contribute to the further study of the cell entry of SARS-CoV-2 in the future. Moreover, through the analysis of RBD-ACE2 binding interface, this study also gave the binding motifs involving in allosteric signal transmission and RBD binding. These multi-conformational binding motifs can be used as targets or templates for the design of peptide inhibitors targeting ACE2 and RBD.

CRediT authorship contribution statement

Qing Liu: Investigation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing, Conceptualization. **Yuwei Wang:** Investigation, Writing - review & editing. **Elaine Lai-Han Leung:** Supervision, Conceptualization, Writing - review & editing, Funding acquisition. **Xiaojun Yao:** Supervision, Conceptualization, Writing - review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.09.032.

References

[1] Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 2020;181 (2). 281-292 e6.

[2] Liu S, Xiao G, Chen Y, He Y, Niu J, Escalante CR, et al. Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors. Lancet 2004;363(9413):938–47.

[3] Huang Y, Yang C, Xu XF, Xu W, Liu SW. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacol Sin 2020;41(9):1141–9.

[4] Kawase M, Kataoka M, Shirato K, Matsuyama S, Gallagher T. Biochemical analysis of coronavirus spike glycoprotein conformational intermediates during membrane fusion. J Virol 2019;93(19). https://doi.org/10.1128/JVI.00785-19.

[5] Yan R, Zhang Y, Li Y, Xia Lu, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 2020;367(6485):1444–8.

[6] Hamelberg D, Mongan J, McCammon JA. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. J Chem Phys 2004;120(24):11919–29.

[7] Chodera JD, Singhal N, Pande VS, Dill KA, Swope WC. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. J Chem Phys 2007;126(15):155101. https://doi.org/10.1063/1.2714538.

[8] Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 2000;33(12):889–97.

[9] Tsui V, Case DA. Theory and applications of the generalized born solvation model in macromolecular simulations. Biopolymers 2000;56(4):275–91.

[10] Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 2020;581 (7807):215–20.

[11] Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234(3):779–815.

[12] Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des 2013;27(3):221–34.

[13] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79 (2):926–35.

[14] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 2015;11(8):3696–713.

[15] Case DA, Brozell SR, Cerutti DS, Cheatham III TE, Cruzeiro VWD, Darden TA, et al. AMBER 2018. San Francisco, San Francisco: University of California; 2018.

[16] Miao Y, Nichols SE, McCammon JA. Free energy landscape of G-protein coupled receptors, explored by accelerated molecular dynamics. Phys Chem Chem Phys 2014;16(14):6398–406.

[17] Rivalta I, Sultan MM, Lee N-S, Manley GA, Loria JP, Batista VS. Allosteric pathways in imidazole glycerol phosphate synthase. PNAS 2012;109(22): E1428–36.

[18] Girvan M, Newman MEJ. Community structure in social and biological networks. PNAS 2002;99(12):7821–6.

[19] R Core Team. R: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria; 2020.

[20] Scherer MK, Trendelkamp-Schroer B, Paul F, Perez-Hernandez G, Hoffmann M, Plattner N, et al. PyEMMA 2: a software package for estimation, validation, and analysis of Markov models. J Chem Theory Comput 2015;11(11):5525–42.

[21] Yuan C, Chen H, Kihara D. Effective inter-residue contact definitions for accurate protein fold recognition. BMC Bioinf 2012;13:292.

[22] Barros EP, Casalino L, Gaieb Z, Dommer AC, Wang Y, Fallon L, et al. The flexibility of ACE2 in the context of SARS-CoV-2 infection. Biophys J 2021;120 (6):1072–84.

[23] Cao L, Goreshnik I, Coventry B, Case JB, Miller L, Kozodoy L, et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. Science 2020;370 (6515):426–31.

[24] Melin AD, Janiak MC, Marrone 3rd F, Arora PS, Higham JP. Comparative ACE2 variation and primate COVID-19 risk. Commun Biol 2020;3(1):641.

[25] Mahdy MAA, Younis W, Ewaida Z. An overview of SARS-CoV-2 and animal infection. Front Vet Sci 2020;7:596391.

[26] Zou J, Yin J, Fang L, Yang M, Wang T, Wu W, et al. Computational prediction of mutational effects on SARS-CoV-2 binding by relative free energy calculations. J Chem Inf Model 2020;60(12):5794–802.