S373P Mutation Stabilizes the Receptor-binding Domain of Spike Protein in Omicron and Promotes Binding

Bin Zheng[1], Yuelong Xiao[1], Bei Tong[2], Yutong Mao[1], Rui Ge[1], Fang Tian[1], Xianchi Dong[3,4], and Peng Zheng[1*]

1. State Key Laboratory of Coordination Chemistry, Chemistry and Biomedicine Innovation Center (ChemBIC), School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, Jiangsu, 210023, China
2. Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, Jiangsu, 210014, China
3. State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing, Jiangsu, 210023, China
4. Engineering Research Center of Protein and Peptide Medicine, Ministry of Education, Nanjing, Jiangsu, 210023, China

*Correspondence: pengz@nju.edu.cn

SUPPLEMENTARY MATERIALS

Materials and Methods

Protein expression and purification. The genes were ordered from Genscript Inc. The wtRBD and OmicronRBD constructs contain the SAS-CoV-2 spike protein (residues 319-591), followed by a GGGGS linker and an 8XHis tag in pcDNA3.4 modified vector. They were expressed in Expi293 cells with OPM-293 CD05 serum-free medium (OPM Biosciences)(*20*). In addition to these two RBDs, most RBDs (residues 333-528) were constructed as a fused polyprotein Coh-$(GB1)_2$-RBD-GB1-NGL for high-precision AFM measurement and thus expressed in *E. coli* BL21(DE3) using pQE80L vector. The *BamH*I-*Bg*lII-*Kpn*I three-restriction enzyme system was used for the stepwise construction of the genes for polyproteins. For protein purification of RBD with His-tag, culture supernatant was passed through a Ni-NTA affinity column (Qiagen). Proteins were further purified by gel filtration (Superdex^{TM} 200 Increase 10/30GL, GE Healthcare).

*Oa*AEP1(C247A) is cysteine 247 to alanine mutant of asparaginyl endoproteases 1 from *oldenlandia affinis*, abbreviated as AEP here (*1*). ELP is the elastin-like polypeptides (*2*). Their expression and purification protocols can be found in references. RBD mutants, including wtRBD(S371L), wtRBD(S373P), OmicronRBD (L371S), and OmicronRBD (P373S), were generated using the QuikChange kit. Their sequences were all verified by direct DNA sequencing.

For the RBDs-ACE2 unbinding experiment, the genes were ordered from GenScript Inc. The RBDs (WT, Omicron) construct contains the SARS-CoV-2 spike

protein (residues 319–591), followed by a GGGGS linker and an His8 tag in a pcDNA3.4 modified vector. Their sequences were all verified by direct DNA sequencing (GENERAL BIOL). A C-terminal NGL was added to the RBDs for ligation. The human ACE2 construct contains the ACE2 extracellular domain (residues 19–740) and an Fc region of IgG1 at the C-terminus followed by NGL.

AFM-SMFS unfolding experiment. The AFM cantilever/tip made of silicon nitride (MLCT-BIO-DC, Bruker Corp.) was used. The detailed protocol for AFM tip functionalization and protein immobilization on the glass coverslip can be found in references (*30, 3*). In short, the tip and glass coverslip were coated with the amino group by amino-silanization. Then, the maleimide group for cysteine coupling was added on the amino-functionalized surface using the hetero-bifunctional crosslinker sulfosuccinimidyl 4-(N-maleimidomehthyl) cyclohexane-1-carboxylate) (Sulfo-SMCC, Thermo Scientific). Next, the peptide GL-ELP_{20}-C or C-ELP_{20}-NGL was reacted to the maleimide via the cysteine, respectively. The long ELP_{20} serves as a spacer to avoid non-specific interaction between the tip and the surface as well as a signature for the single-molecule event. Finally, target protein RBDs with C-terminal NGL sequence or GB1-Doc with N-terminal GL sequence can be site-specifically linked to the coverslip or tip by ligase AEP, respectively.

Atomic force microscope (Nanowizard4, JPK) was used to acquire the force-extension curve. The D tip of the MLCT-Bio-DC cantilever was used. Its accurate spring constant was determined by a thermally-induced fluctuation method (*4*). Typically, the tip contacted the protein-immobilized surface for 400 *ms* under an indentation force of 450 pN to ensure a site-specifically interaction. Then, moving the tip up vertically at a constant velocity (1 μm/s, if not specified), the polyprotein unfolded. Then, the tip moved to another place to repeat this cycle several thousands of times. As a result, a force-extension curve was obtained, which was analyzed using JPK data process analysis software.

AFM-SMFS unbinding experiment of RBD-ACE2 on surface

FD-based AFM on model surfaces was performed in Tris-HCl buffer (pH7.4, 150 mM NaCl, 50 mM Tris-HCl) at room temperature using functionalized D tip of MLCT-Bio-DC cantilever (Bruker, nominal spring constant of 0.030 N/m and actual spring constants calculated using thermal tune). AFM (Nanowizard4, JPK) operated in the force mapping (contact) mode was used. Areas of 10 × 10 μm were scanned, ramp size set to 350 nm, and set point force of 300 pN with a contact time of 50 ms, with a resolution of 32 × 32 pixels.

Bell-Evans model to extract kinetics. The RBD-ACE2 complex dissociation in the AFM experiment is a non-equilibrium process that can be modeled as an all-or-none two-state process with force-dependent rate constant k(F). The rate constant can be described by Bell-Evans' model (*34*):

$$k(F) = k_{\text{off}} \exp\left(\frac{F\Delta x}{k_b T}\right) \quad (1)$$

$k(F)$ is the protein unfolding rate under a particular force F, k_{off} is the unfolding rate constant under zero force, Δx is the distance between the unfolded state and the transition state. For the dynamic force spectroscopy measurements, the slope a of the force−extension curves immediately before the unfolding event (~2 nm) was first determined to obtain the average loading rate ($r = av$, where v is the velocity). The Bell-Evans model was used to fit all the data (1), yielding the spontaneous unfolding rate, and the distance from the folded state to the transition state with the following equation:

$$F = \frac{k_b \mathrm{T}}{\Delta x} \ln\left(\frac{\Delta x}{k_{\text{off}} k_b \mathrm{T}}\right) + \frac{k_b \mathrm{T}}{\Delta x} \ln(\mathrm{r}) \qquad (2)$$

Four different pulling velocities, 0.2 μm/s, 0.4 μm/s, 1 μm/s, and 4 μm/s were used. The relationship between the most probable unfolding force and loading rate was obtained on a log scale, which is fitted by a linear line as equation (2).

MD simulation for RBD unfolding. To reveal the unfolding pathway of RBD under mechanical load, we performed steered molecular dynamics (SMD) simulations using NAMD(*5, 6*). Simulation systems were prepared using CHARMM-GUI (*7*). The structures of the RBDs were prepared following established protocols. For the WT, the structure had been solved by Electron Microscopy at 2.60 Å resolution and is available at the protein data bank (PDB ID: 6zge). The Omicron had also been solved by Electron Microscopy, at 3.5 Å resolution, and is available at the protein data bank (PDB ID: 7tl1). After extracting residues 333 to 528 in part A both of them, the RBDs were solvated and the net charge of the proteins were neutralized using a 150 mM salt concentration of sodium chloride. Disulfide bonds and the glycans at N343 were included following the literature information (*8*). SMD simulations were performed employing the NAMD molecular dynamics package. The CHARMM36 force field along with the TIP3P water model was used to describe all systems. The simulations were performed assuming periodic boundary conditions in the NpT ensemble with temperature maintained at 298 K using Langevin dynamics for pressure, kept at 1 bar, and temperature coupling (*9*). Before the MD simulations, all the systems were submitted to an energy minimization protocol for 5,000 steps. MD simulations with position restraints in the protein backbone atoms were performed for 1.0 *ns* and served to pre-equilibrate systems before the 10 *ns* equilibrium MD runs. To characterize the unfolding pathway of RBDs, ten times of SMD simulations with constant stretching velocity employed a pulling speed of 5.0 Å/ns and a harmonic constraint force of 7.0 kcal/mol/Å was applied for 30.0 ns. In this step, SMD were employed by harmonically restraining the position of the C-terminus and pulling on the N-terminus of the RBDs (WT or Omicron). Each system was run 10 times. Simulation force-extension traces were analyzed analogously to experimental data. Data were analyzed by python-based Jupyter notebooks. Jarzynski's equality is applied to potential mean force (PMF) from SMD simulations (*10*). We also performed 60 times of SMD simulations with a pulling speed of 20 Å/ns to calculate the PMF.

MD simulations were performed utilizing the GROMACS 2021 package. And all simulation systems, including the mutation (G339D, S371L, S373P, or S375F), is established with similar procedures mentioned above including energy minimization and pre-equilibrium. All the input files were generated from CHARMM-GUI. The system's environment changed at 298 K (NVT ensemble) and subsequently at 298 K and 1 bar (NPT ensemble). We performed 300 ns MD simulations 10 times. And the number of hydrogen bonds formed between β2/β10 and β1/β5 (residue 353 to 363, 394 to 400, 523 to 526) was analyzed by the plug-in program *hbond* in GROMACS(*11*). The cutoff parameters of the hydrogen bond analysis program were defaulted values (3.5 Å and 30°). It means that when the distance between donor D and acceptor A was shorter than 3.5 Å, as well as the bond angle H–D is smaller than 30.0°, it is regarded as a hydrogen bond. Hydrogen bond formation for every residue between β2/β10 and β1/β5 is also generated by the same plug-in program, executed by a Python script. The distance between the Cα atoms of the 373 and 343 residues during the MD simulations of WT and Omicron was calculated by another plug-in program *distance* and organized by a python script to form a scatter plot with the number of hydrogen bonds. Then we use the pearsonr function in the scipy module to evaluate the correlation between the number of hydrogen bonds and the distance between residue 371 and 343 (*12*). RMSF and energy analysis were performed using gromacs built-in programs *rmsf* and *energy*

SMD simulations of RBD-ACE2 complex

The SMD simulation systems for RBD and ACE2 were established following the aforementioned method. Both the OmicronRBD-ACE (pdb code: 7wbl) and wtRBD-ACE2 (pdb code: 6m0j) complexes underwent energy minimization, 1 ns of pre-equilibration, and 10 ns of production MD simulation using NAMD2. The C-terminus of RBD and ACE2 were designated as the pulling and fixing groups, respectively. The pulling was conducted at a rate of 1 nm/ns, with each SMD simulation lasting for 8 ns. A total of 10 simulations were conducted, ensuring complete dissociation of the complex in each simulation.

Supplementary Text

Protein sequences:

His6-Coh-GB1-GB1-RBD**(XX)**-GB1-NGL

MRGSHHHHHHGSMGTALTDRGMTYDLDPKDGSSAATKPVLEVTKKVFDTA
ADAAGQTVTVEFKVSGAEGKYATTGYHIYWDERLEVVATKTGAYAKKGAA
LEDSSLAKAENNGNGVFVASGADDDFGADGVMWTVELKVPADAKAGDVYP
IDVAYQWDPSKGDLFTDNKDSAQGKLMQAYFFTQGIKSSSNPSTDEYLVKAN
ATYADGYIAIKAGEPRSMDTYKLILNGKTLKGETTTEAVDAATAEKVFKQYA
NDNGVDGEWTYDDATKTFTGTERSMDTYKLILNGKTLKGETTTEAVDAATA
EKVFKQYANDNGVDGEWTYDDATKTFTGTERSXXRSMDTYKLILNGKTLK
GETTTEAVDAATAEKVFKQYANDNGVDGEWTYDDATKTFTVTERSNGL

XX: **WT**
TNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSP
TKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAW
NSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCY
FPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPK

XX: **WT(C4^{-})**
TNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKAYGVSP
TKLNDLAFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGAVIAW
NSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCY
FPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVAGPK

XX: **WT(S371L)**
TNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNLASFSTFKCYGVSP
TKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAW
NSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCY
FPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPK

XX: **WT(S373P)**
TNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSAPFSTFKCYGVSP
TKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAW
NSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCY
FPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPK

XX: **Omicron**
TNLCPFDEVFNATRFASVYAWNRKRISNCVADYSVLYNLAPFFTFKCYGVSP
TKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGNIADYNYKLPDDFTGCVIAW
NSNKLDSKVSGNYNYLYRLFRKSNLKPFERDISTEIYQAGNKPCNGVAGFNC
YFPLRSYSFRPTYGVGHQPYRVVVLSFELLHAPATVCGPK

XX: **Omicron(4C-)**
TNLCPFDEVFNATRFASVYAWNRKRISNCVADYSVLYNLAPFFTFKAYGVSP
TKLNDLAFTNVYADSFVIRGDEVRQIAPGQTGNIADYNYKLPDDFTGAVIAW
NSNKLDSKVSGNYNYLYRLFRKSNLKPFERDISTEIYQAGNKPCNGVAGFNC
YFPLRSYSFRPTYGVGHQPYRVVVLSFELLHAPATVAGPK

XX: **Omicron(L371S)**
TNLCPFDEVFNATRFASVYAWNRKRISNCVADYSVLYNSAPFFTFKCYGVSP
TKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGNIADYNYKLPDDFTGCVIAW
NSNKLDSKVSGNYNYLYRLFRKSNLKPFERDISTEIYQAGNKPCNGVAGFNC
YFPLRSYSFRPTYGVGHQPYRVVVLSFELLHAPATVCGPK

XX: **Omicron(P373S)**
TNLCPFDEVFNATRFASVYAWNRKRISNCVADYSVLYNLASFFTFKCYGVSP

TKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGNIADYNYKLPDDFTGCVIAW
NSNKLDSKVSGNYNYLYRLFRKSNLKPFERDISTEIYQAGNKPCNGVAGFNC
YFPLRSYSFRPTYGVGHQPYRVVVLSFELLHAPATVCGPK

Supplementary Figures

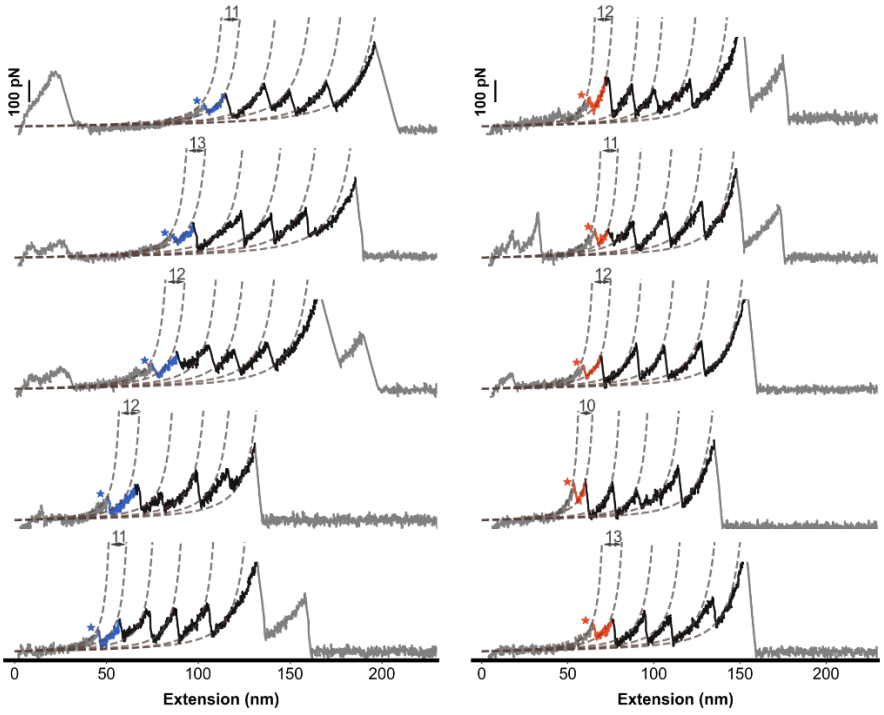


Figure S1. More representative curves of RBDs. Representative curves showing a series of sawtooth-like peaks from the unfolding of the RBD with a ΔLc of ~11 nm and marker protein GB1 with a ΔLc of ~18 nm. The Omicron and WT variants are colored orange and blue, respectively.

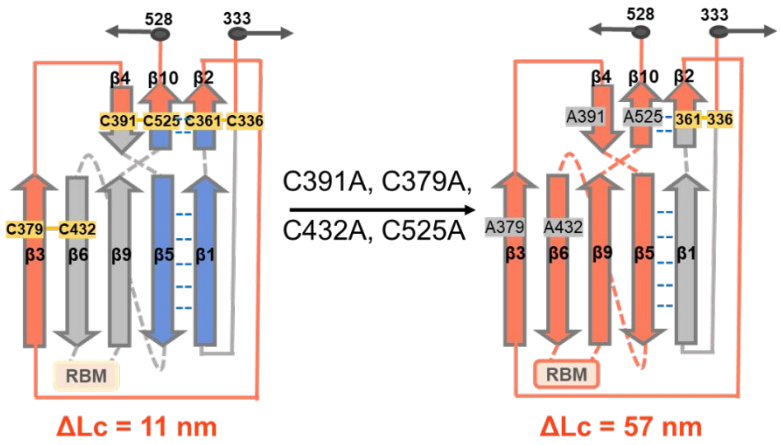


Figure S2. The unfolded fragments of RBD and RBD (4C^{-}) upon unfolding. **.** The comparison between RBD and Cys-delete RBD mutant cartoon structure indicates that a larger ΔLc (57 nm) upon unfolding is expected due to the release of a more extensible protein structure (colored in orange) upon the deletion of two disulfide bonds.

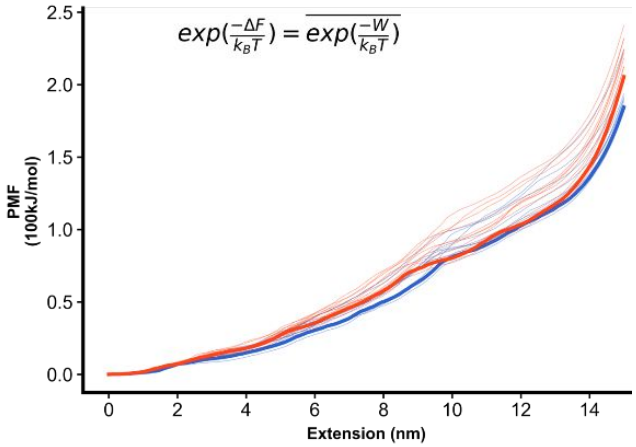


Figure S3. PMF calculation shows that Omicron is more stable than WT. PMF was calculated from irreversible pulling (v =5 Å/ns) of 10 SMD trajectories. Light color lines represent the work done in the simulations.

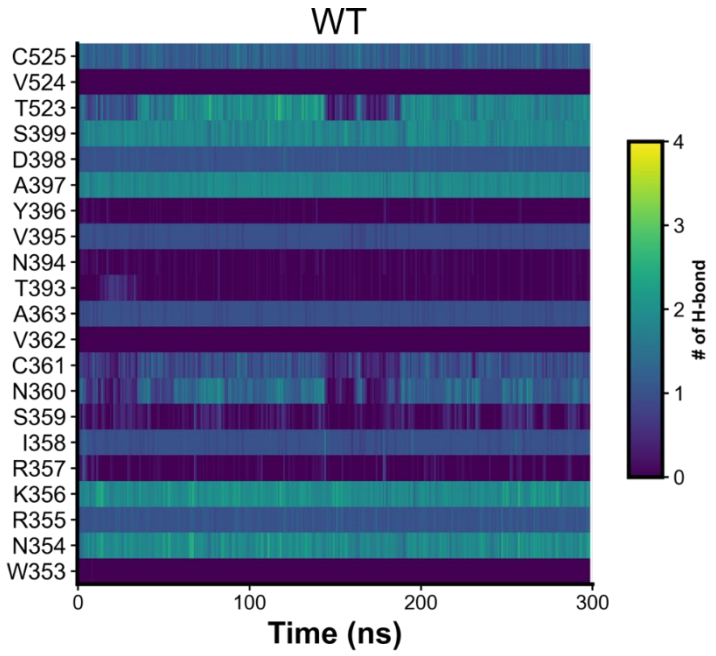


Figure S4. The evolution of H-bond number for each residue involved in β2/β10 and β1/β5 during the MD simulations for wtRBD.

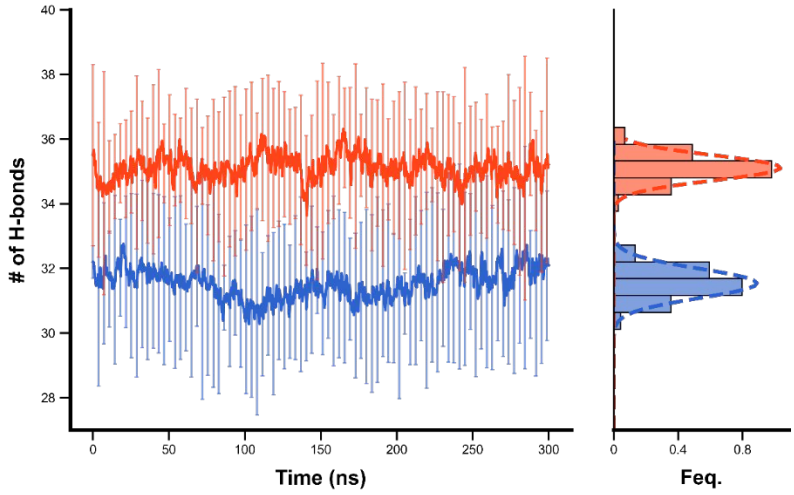


Figure S5. Overall MD simulations during 300 ns. It shows the average number of hydrogen bonds formed in the β-core region. The average number of hydrogen bonds formed by OmicronRBD in the β-core region is still larger than that of wtRBD.

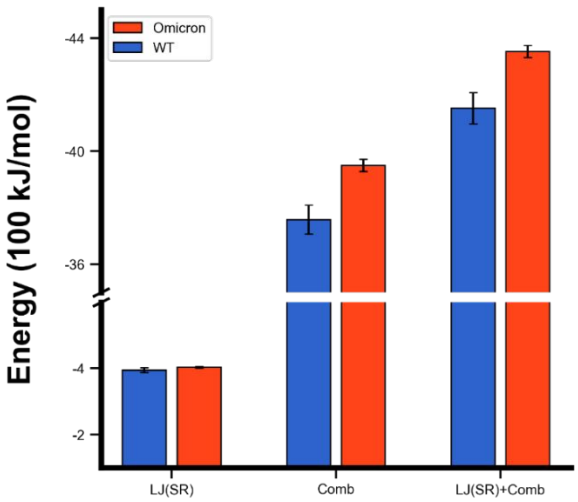


Figure S6. Energy analysis of β-strands during 10 times MD simulations. The coulomb interaction of Omicron RBD is stronger than wtRBD. Thus, the Coulomb interaction might play an important role in the stability of the β-strands.

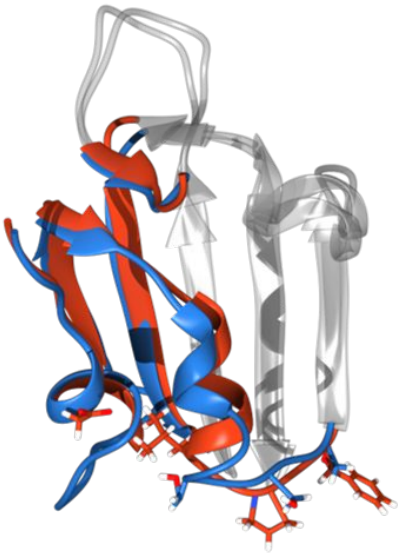


Figure S7. Overlap of the Cryo-EM structure of Omicron (red) and WT (blue). A slight conformational change closing to the β-core region of RBD is observed between Omicron and WT. The four mutations are depicted.

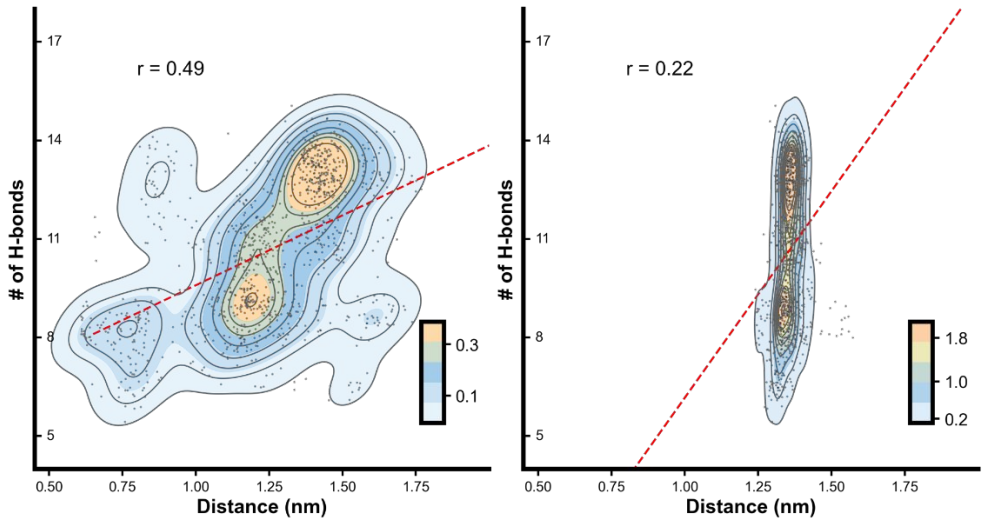


Figure S8. The relationship between the distances and the number of hydrogen bonds for different residues. The relationship between the distances of residues 371(left) / 375(right) to 432 and the number of hydrogen bonds is week.

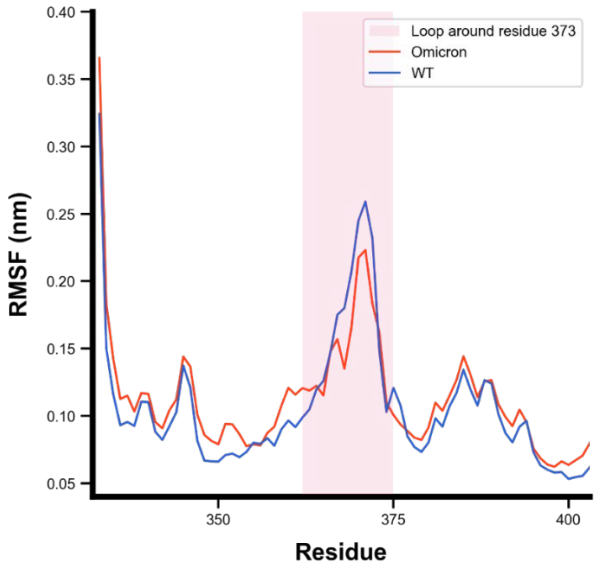


Figure S9. The average RMSF value of 10 times MD simulations. The RMSF of helix around residue 373 in OmicronRBD was lower than wtRBD, suggesting that the mutation of OmicronRBD could enhance the stability of this region and change the movement of specific residues.

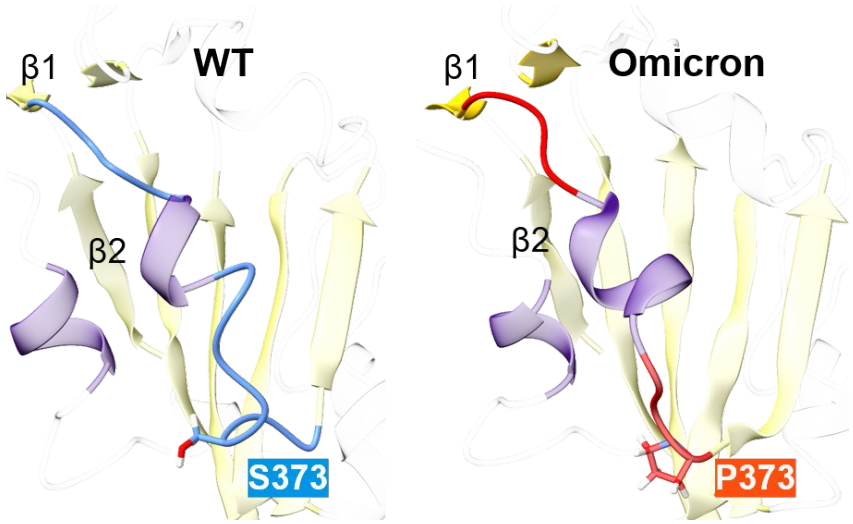


Figure S10. A focus view of peptide chain around residue 373. The rigid proline 373 in OmicronRBD might open the loop between the nearby helix and β-strand, so that it can maintain a relatively large distance from the helix.

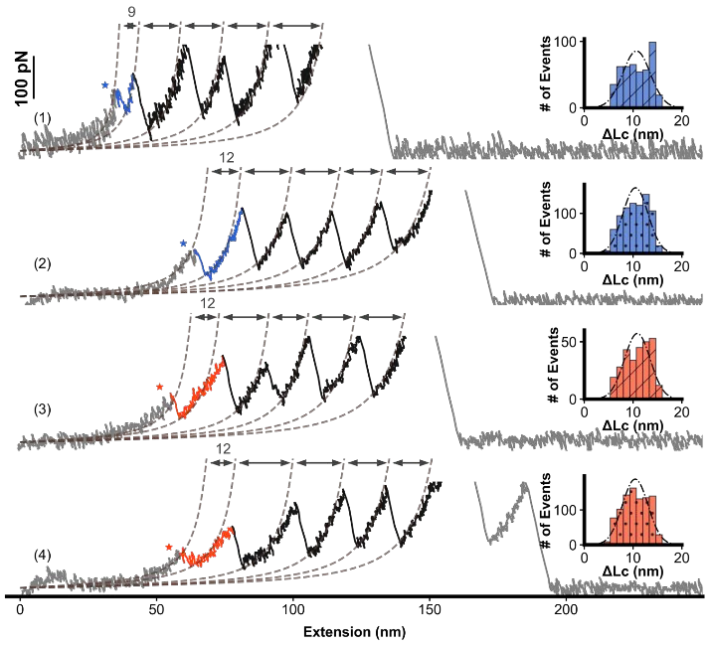
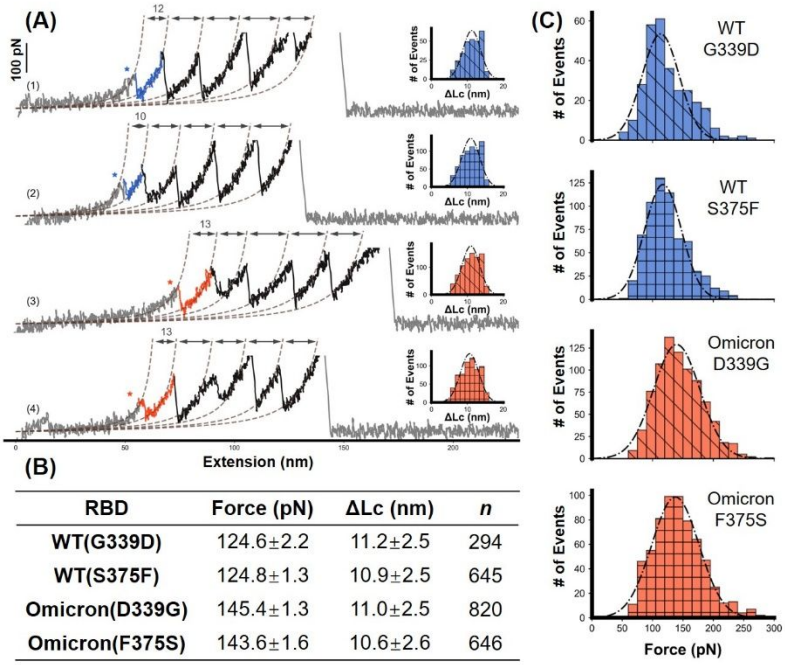


Figure S11. Representative curves of single mutation of residue 371 or 373 on RBD. Force-extension curves of RBDs of WT(S371L), WT(S373P), Omicron(L371S), and Omicron(P373S) showed the stepwise unfolding of the polyprotein, including the RBD, respectively. The histogram of ΔLc of RBD is shown in the inset.



(B)

RBD	Force (pN)	ΔLc (nm)	*n*
WT(G339D)	124.6±2.2	11.2±2.5	294
WT(S375F)	124.8±1.3	10.9±2.5	645
Omicron(D339G)	145.4±1.3	11.0±2.5	820
Omicron(F375S)	143.6±1.6	10.6±2.6	646

Figure S12. AFM unfolding results of single mutation of residue 339 or 375 on RBD. A) Force-extension curves of RBDs of WT(G339D), WT(S375F), Omicron(D339G), and Omicron(F375S) showed the stepwise unfolding of the polyprotein, including the RBD, respectively. The histogram of ΔLc of RBD is shown in the inset. **B-C)** AFM unfolding results of the four mutants are shown in details.

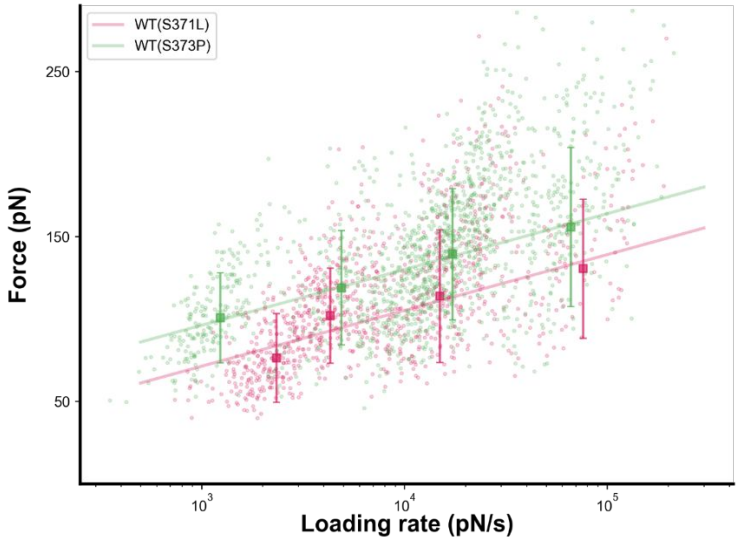


Figure S13. Dynamic force spectrum of the unfolding of RBDs. The unfolding forces of the WT(S371L) and WT(S373P) show a linear relationship with the logarithm of the loading rate. The k_{off} was 0.53±0.79s^{-1} for the WT(S371L) and 0.10±0.04 s^{-1} for the WT(S373P). The WT(S371L) and WT(S373P) are colored pink and green, respectively.

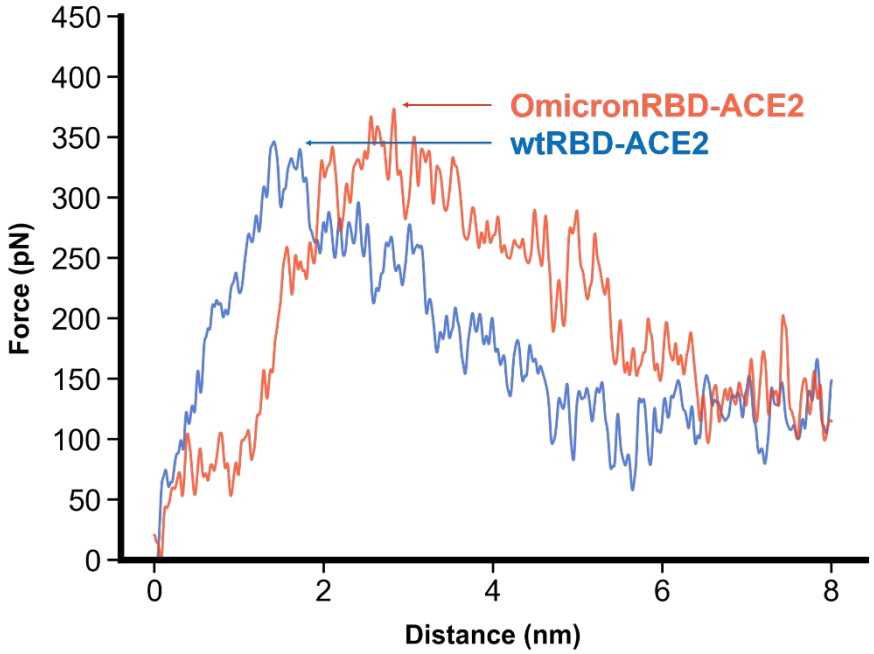


Figure S14. SMD simulations of RBD-ACE2 complex dissociation. Average dissociation force profiles (10 times) of wtRBD-ACE2 (blue) and OmicronRBD-ACE2 (orange) as a function of the pulling distance. The dissociation force of OmicronRBD-ACE2 is larger than wtRBD-ACE2, consistent with a previous study[13].

Supplementary References:

1. R. Yang *et al.*, *J. Am. Chem. Soc.* **139**, 5351-5358 (2017).
2. W. Ott *et al.*, *Acs Nano* **11**, 6346-6354 (2017).
3. S. Shi *et al.*, *CCS Chem.* **4**, 598-604 (2022).
4. J. L. Hutter, J. Bechhoefer,. *Revi. Sci. Instrum.* **64**, 1868-1873 (1993).
5. W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **14**, 33-38, 27-38 (1996).
6. J. C. Phillips *et al.*, *J. Chem. Phys.* **153**, 044130 (2020).
7. R. B. Best *et al.*, *J. Chem. Theory Comput.* **8**, 3257-3273 (2012).
8. J. Lan *et al*. *Nature* 581.7807: 215-220 (2020).
9. M. Kim, E. Kim, S. Lee, J. S. Kim, S. Lee, *J. Phys. Chem. A* **123**, 1689-1699 (2019).
10. S. Park, F. Khalili-Araghi, E. Tajkhorshid, K. Schulten, *J. Chem. Phys.* **119**, 3559-3566 (2003).
11. D. Van Der Spoel *et al.*, *J. Comput. Chem.* **26**, 1701-1718 (2005).
12. P. Virtanen *et al.*, *Nat. Methods* **17**, 261-272 (2020).
13. S. Kim et al. *J. Comput. Chem.* **44**, 594-601 (2022).

Movies S1 to S4

S1: Movie of Steering molecular dynamics (SMD) simulations of wtRBD
S2: Movie of Steering molecular dynamics (SMD) simulations of OmicronRBD
S3: Movie of MD simulations to determine the number of H-bonds of the wtRBD
S4: Movie of MD simulations to determine the number of H-bonds of the OmicronRBD