Binding Affinity of Neutralizing Antibody Binds to Receptor Binding Domain of SARS-CoV-2 and British N501Y Mutant in Silico

Nianjue Zheng

Key laboratory of Regenerative Medicine of Ministry of Education, Department of Developmental & Regenerative Biology, Jinan University, Guangzhou, 510632 China.

znj1990@jnu.edu.com

Abstract

The coronavirus disease 2019 (COVID-19) outbreak has caused a serious public health incident. With the spread of acute respiratory syndrome 2 coronavirus (HcoV-19/SARS-CoV-2), multiple mutations have appeared in different countries and regions. A new British mutant strain is reported containing N501Y can increase the risk of infection. In order to research the neutralizing effect of the Fab region B38 antibody in combination with SARS-CoV-2-British mutant strain (B.1.1.7) N501Y, we performed homology modeling, molecular dynamics simulation, molecular mechanics Poisson-Boltzmann surface area to study the differences in the binding affinity of receptor binding domain of HcoV-19/SARS-CoV-2 and its Britiish mutant strain N501Y. The molecular dynamics (MD) simulation ressult shows that the N501Y variant did not cause much changes in the structure of the RBD and the antibody binding pocket. Molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) calculation indicated that N501Y variant enhances the binding energy of RBD and B38 antibody. The free energy landscape (FEL) of SARS-CoV-2-RBD and N501Y variant-RBD have been further constructed, the results show that N501Y variant can promote the stability of its complex with B38 antibody,the interaction with B38 antibody can be enhanced by conformational flexibility and hydrophilic residue. Research helps to reveal the binding mechanism of two complexes based on RBD, and provide to assistance in the development of neutralizing therapeutic antibodies.

Keywords

Receptor Binding Domain (RBD); COVID-19; British N501Y Mutant; B38 Antibody; Molecular Dynamics Simulation; MM-PBSA.

1. Introduction

The pandemic of Coronavirus disease-2019(COVID-19) caused by HCoV-19/SARS-CoV-2 has infected more than 100 million people in over 200 countries and regions, has caused more than 3.5 million people deaths as of 27 May 2021, according to data released by Johns Hopkins university in USA [1,2]. Studies have shown that 2019-nCoV, like the SARS coronavirus in 2003 and the MERS coronavirus in the Middle East Respiratory Syndrome, belongs to the β-coronavirus, posing a great threat to human life[3]. So far there is no approved specific therapy against COVID-19.

SARS-CoV-2/HcoV-19 is an RNA virus with a high mutation rate. At the end of 2020, British scientists discovered a novel mutant of SARS-CoV-2 called B.1.1.7, these mutations include the mutation of Asparagine at position 501 to Tyrosine(N501Y), which has attracted great attention around the world. Studies have shown that its transmission power has increased by more than 30% compared to the wild type. Davies et al modeling data from three regions in the United Kingdom, it is estimated that the spreading capacity of this mutant strain has a 56% higher than previous mutant strains[4]. Since the spike protein is the main target for vaccine development and therapeutic antibody development, whether these mutations will produce antibody escape and reduce the effectiveness of vaccines and antibody drugs has become the focus of researchers. The N501Y mutant is located in

the viral RBD for cell entry,experiments indicated that the N501Y could enhance the binding to the ACE2 receptors, allowing the virus to expand its host range to infect mice[5,6].

Although many vaccines are under preclinical and clinical studies, there is a long term of clinical trials, a key part for the COVID-19 treatment strategy is the development of neutralizing antibodies with expected efficacy and safety profile. B38,an neutralizing antibody from COVID-19 convalescent patient, has been confirmed by researchers that could complete competition with ACE2 for binding to SARS-CoV-2-RBD[7]. There are 36 amino acids on the RBD involved in the interaction with B38 antibody, 21 amino acids and 15 amino acids interact with heavy and light chain, respectively[8]. One of the epitope residues N501 of RBD is also on the interaction interface with B38.

To reduce the costs and accelerate the specific drugs discovery phase, in this study, we combined biological macromolecule simulation methods to assess the molecular interaction between RBD of the N501Y mutant and B38 antibody and compared our results with the wild type. These findings will not only provide efficient and reliable guidance for development of anti-COVID-19 antibody drugs, but also can provide early warning for mutations at key sites of mutants.

2. Materials and methods

2.1 Structure preparation

The complex of B38-SARS-CoV-2-RBD was downloaded from RCSB protein data bank (PDB) (https://www.rcsb.org/) ,PDB ID:7BZ5.

Coronavirus spike protein is composed of two functional subunits, S1 and S2. S1 unit binds to ACE2 through its RBD and infects the host. In this study, we explored the RBD of Spike proteins on SARS-CoV-2/ HcoV-19 and N501Y mutant. We downloaded their amino acid sequences from NCBI website with its corresponding ID numbers: 7BZ5_A(HCoV-19-RBD) and 7NEG_E (N501Y mutant-RBD) (https:// www. ncbi. nlm. nih. gov). ClustalW was performed for sequence alignment (https:// www. genome. jp/tools-bin/clustalw).

2.2 Homology Modeling

we adopted the B38-SARS-CoV2-RBD (PDB ID: 7BZ5) structure as a template, and the 3D structural model of the B38-N501Y mutant-RBD complex was built by the Homology modeling in SWISS-MODEL tool. We perform loop area optimization, hydrogenation optimization and side chain optimization on the optimal model ,respectively, evaluating the model use Ramanchandran Plot.

2.3 Molecular dynamics (MD)simulation

The MD simulation with CHARMm 27 force filed was used the GROMACS 4.6.7 software package. The complexes of B38-SARS-CoV-2-RBD and B38-N501Y mutant-RBD were placed in the cubic lattice filled with TIP3P water molecules with an implicit model, the simulation time was set to 50ns and on a GPU computer node to calculate. Before MD simulation,we used the 500-step steepest descent (SD) methond to minimize the energy of the system[9].

The detailed parameters of the MD simulation are as follows: the integration time step is set to 2fs; the eliminiation frequency of the translation and rolation of the center of mass is set to 1 step/time; the update frequency of the non-bonded interaction is set to 10 steps/time; use the Particle-Mesh Ewald (PME) algorithm[10] to calculate the elecrostatic interactions;Fourierspacing and Coulomb radii are set to 0.135nm and 1.0nm, respectively; van der Waals (VDW) interaction truncation radius is 1.4nm; The temperature of the thermal bath for protein and non-protein components (such as solvents and ions) is set to 300 K; the system pressure is set to 1 bar by using Parrinello-Rahman barostats; the initial velocity of the atom Randomly generated according to the Maxwell distribution under 300 K conditions; using the LINCS algorithm [11]. The frequency of recording the structure frame is 10 ps/time. The generated MD trajectories were analyzed, the root mean square deviation (RMSD) and root mean square fluctuation (RMSF) are calculated by GROMACS command g_rms and g_rmsf, respectively.

2.4 MM-PBSA calculation and energy residue decomposition

The MM-PBSA algorithm [12] is the most widely used algorithm for binding free energy. The algorithm calculates and decomposes the binding free energy between the components of the complex, which can analyze the complex interaction process [13,14]. In this study, the binding free energy of complexes of the B38-wild type and B38-mutant was completed using the plug-in g_mmpbsa program of the GROMACS software package [15]. Calculate the enthalpy of system by using the molecular mechanics method(MM). Use solving the Poisson-Boltzmann (PB) equation to calculate the polar solvation energy. The contribution of non-polar solvent energy was determined by the molecular surface area (SA). The basic principle formula is as follows:

$$\Delta G_{binding} = \Delta G_{complex} - (\Delta G_{protein} + \Delta G_{ligand})$$

$$G = E_{MM} + G_{sol} - T\Delta S$$

$$E_{MM} = E_{ele} + E_{vdw}$$

$$G_{sol} = G_{polar} + G_{nonpolar}$$

Where $\Delta G_{binding}$ represent the binding free energy, $\Delta G_{complex}$, $\Delta G_{protein}$ and ΔG_{ligand} represent free energy of the protein-ligand complex, each of the protein and each of the ligand in the solvent environment,respectively. E_{MM} represents the potential energy of molecular mechanics in vacuum. G_{sol} defines the free energy of dissolution. $T\Delta S$ represents the entropy contribution of free energy. E_{MM} includes two types of bond interaction and non-bond interaction, namely bond angle, dihedral angle, electrostatic interaction (E_{ele}) and Van der Waals interaction (E_{vdw}). Free energy of dissolution includes polar part (G_{polar}) and nonpolar part ($G_{nonpolar}$)

The free energy landscape(FEL) was used to characterize the dynamics of a simulated protein state change process. Use the g_covar program of GROMACS software package to construct the Cα atomic covariance matrix of the wild-type and mutant simulation systems and diagonalize them, then further use the g_anaeig program to project the dynamic simulation trajectory onto the corresponding eigenvectors. Use the results of projection to construct FEL for B38-SARS-CoV-2-RBD and B38-N501Y mutant-RBD simulation systems respectively. The basic principle formula is as follows:

$$\Delta G(X) = -K_B T \ln P(X)$$

The reaction coordinate X represents the projection of the simulated trajectory on an eigenvector, K_B represents Boltzmann's constant, T represents the temperature of the system, and P(X) represents the conformational distribution probability on an eigenvector. In this study In, the final FEL is plotted and displayed by SigmaPlot software.

3. Results

3.1 Sequence and structure alignment analysis

We use ClustalW tool to align the sequence, the RBD of SARS-CoV-2 and N501Y variant show 84.7% sequence similarity and 83.8% sequence identity (Fig 1A). The complex structures of the Fab region of B38 antibody and SARS-CoV-2-RBD were downloaded from the protein structure database, PDB_ID: 7BZ5. The three dimensional model of B38-N501Y mutant-RBD complex was built by SWISS-MODEL tool homology modeling. Figure 1B and 1C are drawn by PyMOL mapping tool, showing the predicted complex structure of B38-N501Y mutant-RBD is very close to the experimental structure of B38-SARS-CoV-2-RBD complex. The root mean square deviation (RMSD) between the two complexes is 0.726 Å. The RBD binding sites of HcoV-19 and N501Y mutant are very conserved. There are 36 amino acids that interact with fab region B38 antibody,21 amino acids interact with the heavy chain,15 amino acids interact with the light chain [7]. Only amino acid 501 in the pair of interaction pockets of complexes is mutated (Fig 1C), Asparagine is mutated to tyrosine, both of which are hydrophilic amino acids, but the side chain groups are different. In terms of structural composition, Asn is an aliphatic amino acid and Tyr is an aromatic amino acid.

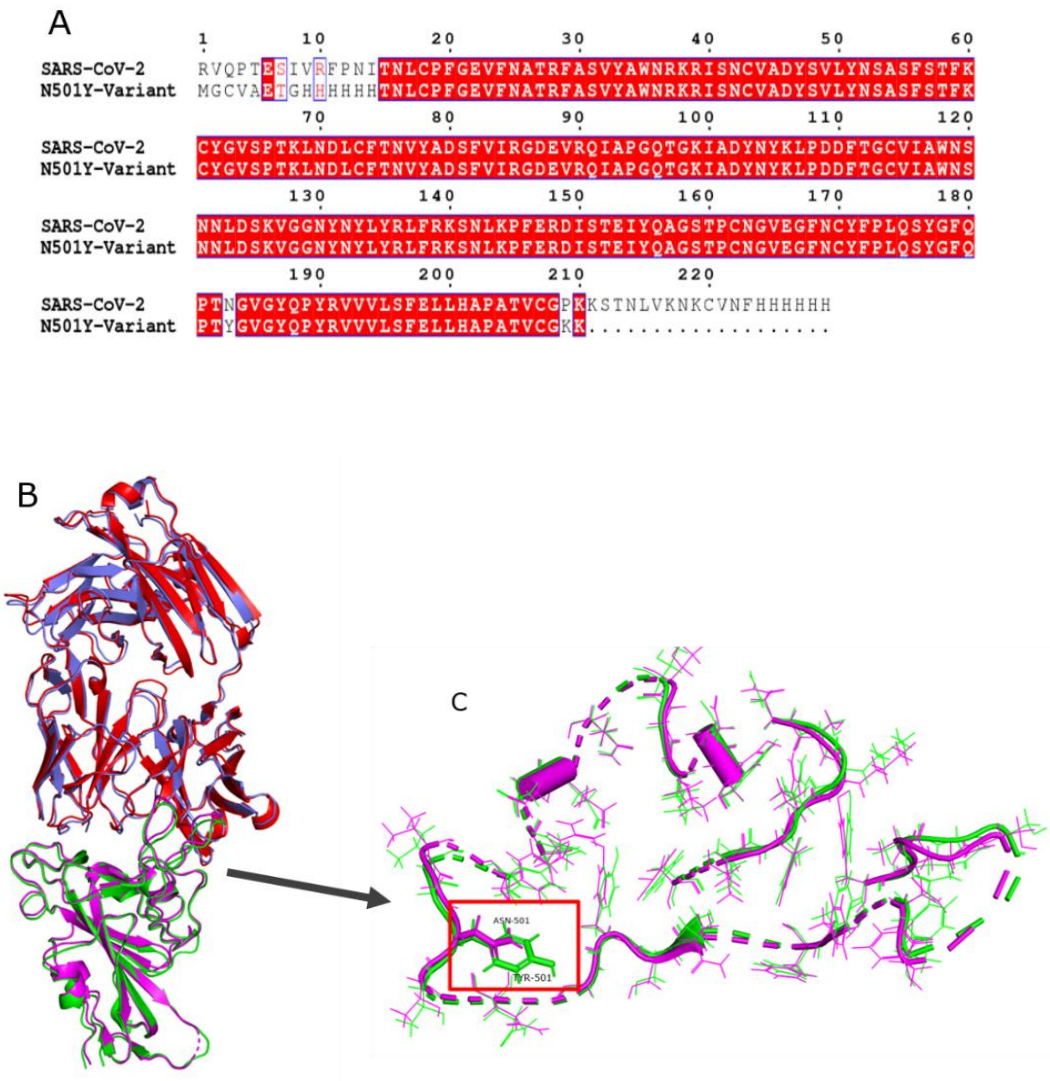

Fig. 1 Sequence and crystal structure alignment. (A). sequence alignment of SARS-CoV-2-RBD and N501Y variant-RBD. (B).Structural comparison between predicted B38-N501Y mutant-RBD complex and experimental B38-SARS-CoV-2-RBD complex. N501Y mutant-RBD is colored green and its interacting B38 antibody is colored red, SARS-CoV-2-RBD is colored magenta and its B38 antibody is colored slate., respectively. (C). The structural alignment of the interaction sites on the RBD of SARS-CoV-2 and N501Y variant that bind to the B38 antibody. Binding sites of SARS-CoV-2-RBD and N501Y variant are colored magenta, green, respectively. The red box indicates the mutation site.

3.2 Structural stability analysis

3.2.1 RMSD analysis

The root mean squared deviation (RMSD) represents the distance of the same atom in different structures. The RMSD of the protein can reveal the position change between the conformation and the initial conformation of the protein during the simulation; the change trend of the RMSD of the protein and the ligand is also an important indicator for judging whether the simulation is stable. In this study, 50ns steps in the MD simulation for two antigen-antibody structures. The RMSD value of the backbone atom of the entire complex structure was calculated (Fig.2A), which can ignore the influence of amino acid, B38-SARS-CoV-2-RBD tends to be in equilibrium after 40ns, compared to B38-N501Y mutant-RBD after 45ns steps. After 35ns step, the conformational fluctuation of RMSD value of the B38-N501Y mutant-RBD is more obvious than wild type, and the RMSD values of B38-N501Y mutant-RBD increased significantly as high as 0.6Å. The RMSD value of the antigen-

antibody interface region(inter-group) was also calculated as show in Figure 2B, the RBD of N501Y mutant has lager fluctuations than SARS-CoV-2-RBD after 10ns steps. It indicates that the conformational stability of B38-N501Y variant -RBD is slightly lower than that of B38-SARS-CoV-2-RBD. Whether it is the RMSD curve of the backbone atoms or the RMSD curve of the interface region is stable and does not exceed 1.25 Å. These RMSD results suggest that 50ns MD simulation is enough to equilibrate the complexes of B38-SARS-CoV-2-RBD and B38-N501Y mutant-RBD. The trajectory from MD simulation is reliable for further analysis.

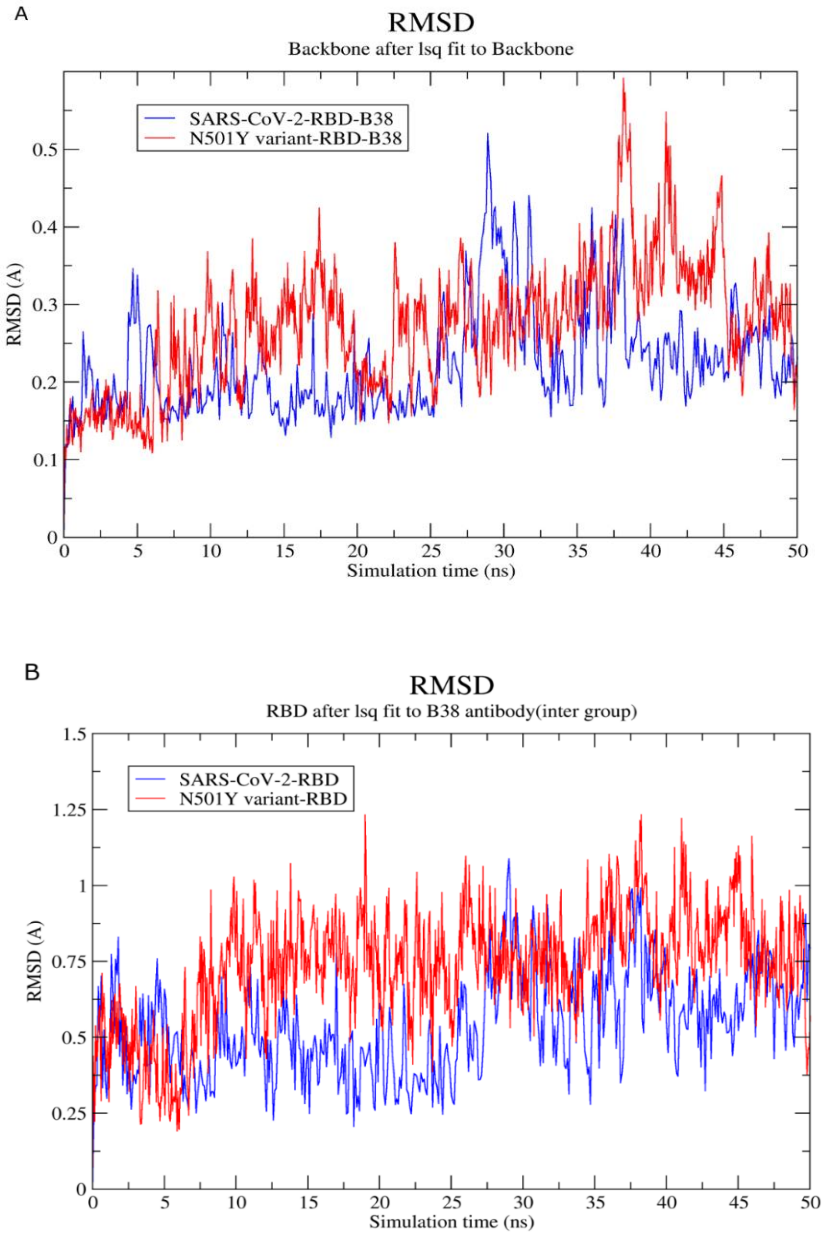

Fig. 2 RMSD curve (A) The RMSD curve of backbone atoms of wild type and mutant. (B) The RMSD curve of the antigen-antibody interface region (inter group)

We performed cluster analysis on RMSD,to build an RMSD matrix. We calculated the RMSD for each combination of structures in a trajectory file, the similar structural features are formed into a cluster. The gromacs tool g_rms is called with two trajectories, then use the GROMACS command xpm2ps to color a matrix for the RMSD cluster (Fig 3). The color blue means that the lower the RMSD, the more stable the structure. The color red means that the greater the RMSD distance, the

more unstable the structure.The RMSD matrix of mutation is more blue areas than the wild type, indicating that the protein structure of the mutant is more stable.

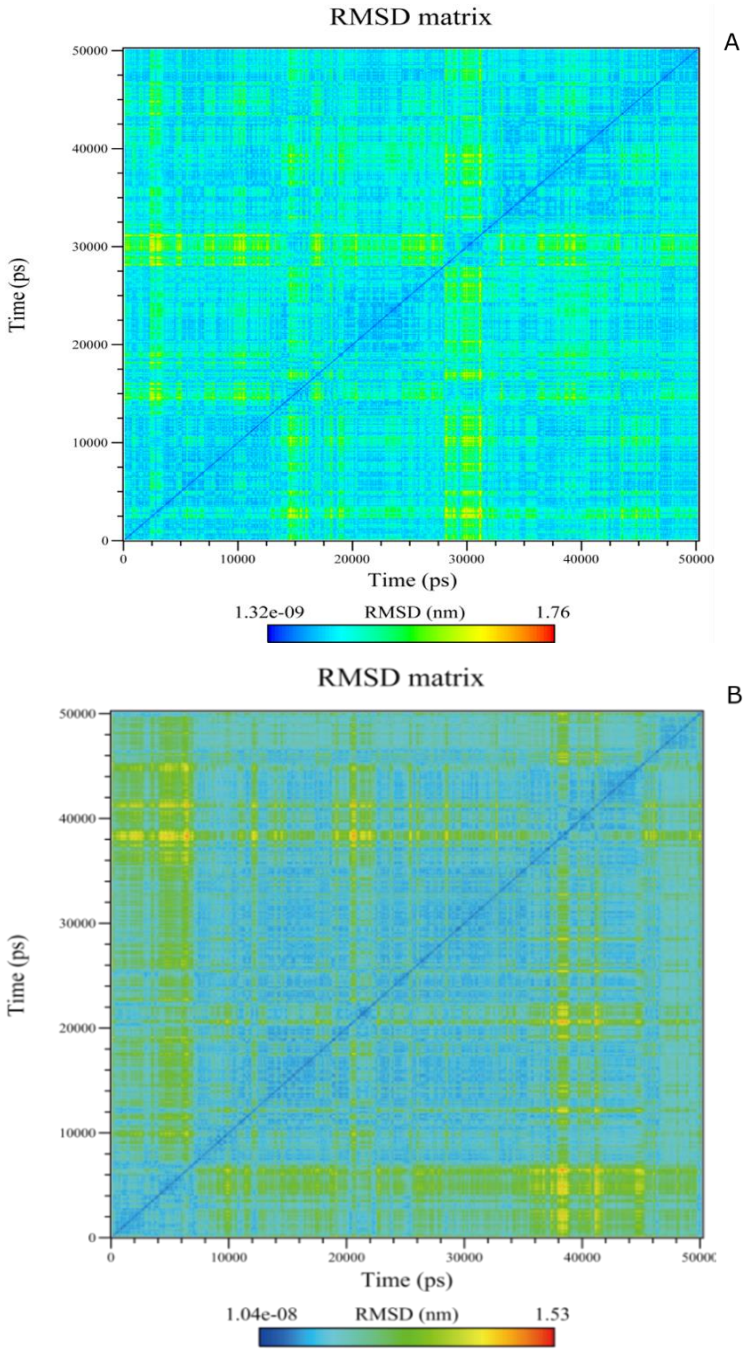

Fig. 3 RMSD matrix. (A) B38-SARS-CoV-2-RBD complex. (B) B38-N501Y mutant-RBD complex

3.2.2 RMSF analysis

The root mean squared fluctuation (RMSF) is the average of atomic position changes over time, which characterize the flexibility and intensity of movement of protein and amino acid during the entire simulation process. The per-residues RMSF of two complexes were calculated as show Fig.4. The lower flexible region is mainly concentrated in the amino acid positions ASP364 to ASN394, most of these are exposed on the surface of complexes belong to loop region. however, the residues that interact with the B38 antibdoy show low RMSF values, it is obvious that the residues with low RMSF values could tightly interact with the B38 antibody to form stable complexes.

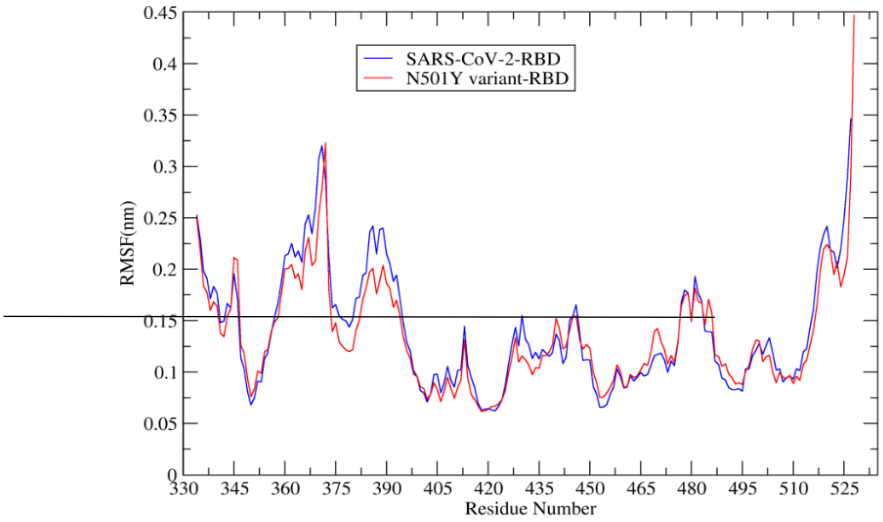

Fig. 4 The RMSF value of per-residue of SARS-CoV-2-RBD and N501Y mutant-RBD

3.2.3 Principal Components

Protein structural dynamics and conformational space are multidimensional and heterogeneous, and their dimensionalities can be reduced by constructing a set of suitable eigenvectors (PC1 and PC2) that gives a vector description of each component of the motion according to principal component analysis (PCA) [16]. We use the MD trajectory on the first eigenvector(PC1) and the second eigenvector (PC2) to construct the Free Energy Landscapes (FEL) to characterize their maximum possible movement of proteins during the simulation process (Fig.5). From figure 5A, the area of blue basin is sacttered and very small, however, B38-N501Y mutant-RBD shows a large basin area, the range is PCA1(-5,7), PCA2(-1,2), and has less free energy minimization zone area (Fig. 5B). This shows that the N501Y mutant has collected a richer conformational space compared wild type. The mutation of N501Y may increase the flexible area of the structure to a certain extent, making it more tightly bound to the B38 antibody.

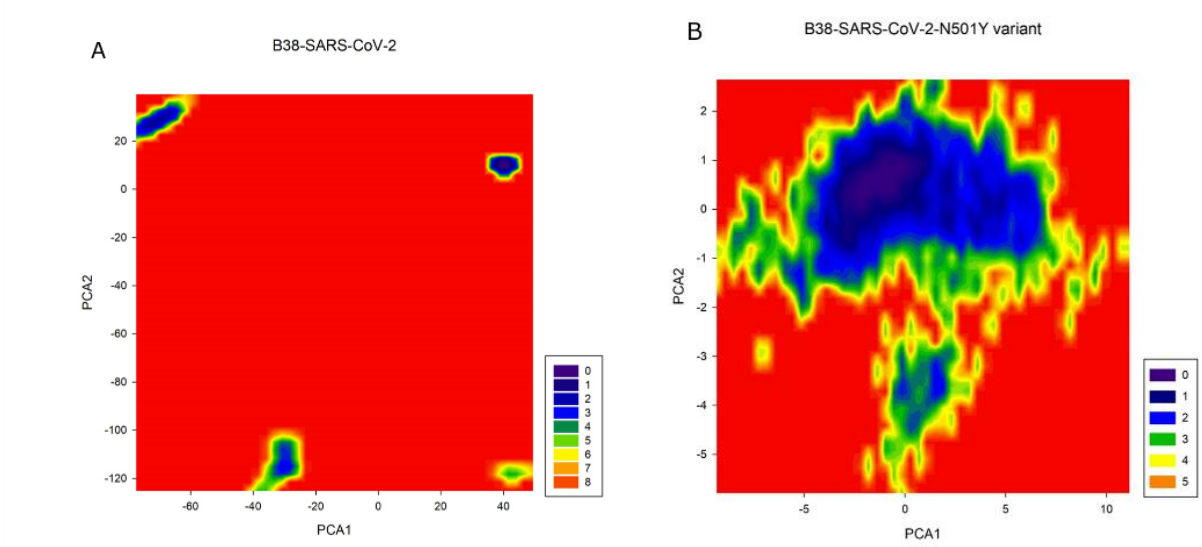

Fig. 5 The Principal Component Analysis (PCA) of RBD structural dynamics properties of B38-SARS-CoV-2 and B38-N501Y mutant-RBD complexes (A and B)

3.2.4 Differences before and after simulated alignments protein

In this study, the protein conformations of the MD simulation from the first frame of trajectory and the last frame of trajectory were superimposed and make a procupine map. From figure 6A, the 3D structure of B38-SARS-CoV-2-RBD exhibits a larger b factor, especially in the yellow-green area of the B38 antibody. The white arrow indicates the trend of the simulation process, the simulation process of B38-N501Y variant -RBD is relatively stable and the white arrow of RBD is to promote the direction of antigen-antibody binding. Conversely,the white arrow of RBD of the SARS-CoV-2 moves in a way that deviates from the antigen-antibody binding.

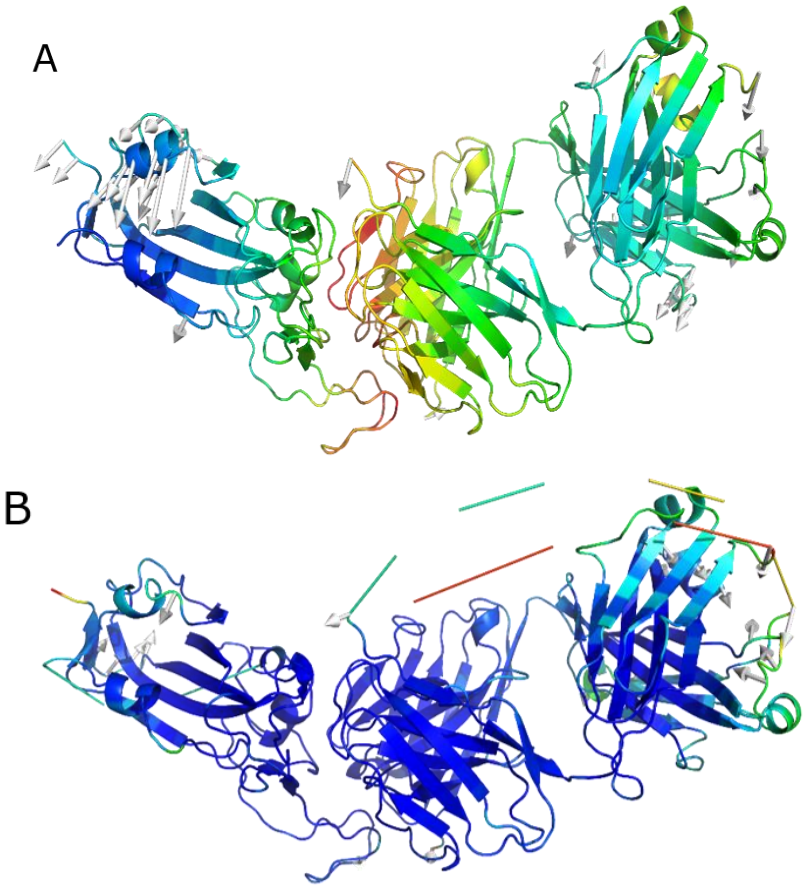

Fig. 6 Structure comparison before and after simulation. (A) B38-SARS-CoV-2-RBD. (B) B38-N501Y mutant-RBD

3.3 MM-PBSA and Energy Decomposition

3.3.1 Binding Free Energy Calculation

In order to quantitatively describe the effect of N501Y mutation on the interaction between RBD and B38 antibody, we adopted the MM-PBSA algorithm to calculate and analyzed in detail the difference among each of the energy components. As shown in Table 1, the binding free energy of cmplexes of wild type and mutant are −87.200 kJ/mol and −124.736 kJ/mol, respectively, indicating that B38 antibody binds to the N501Y mutant more tightly, so the N501Y mutation has a Facilitate the binding between RBD and B38 antibody. Further analysis shows that the main force driving the binding of RBD and B38 antibody is electrostatic interaction (E_{ele}) and Van der Waals interaction (E_{vdw}). For the B38-SARS-CoV-2-RBD complex, E_{vdw} and E_{ele} are -301.133kJ/mol and -223.539 kJ/mol, respectively. For the B38-N501Y mutant-RBD complex, E_{vdw} and E_{ele} are -470.165kJ/mol and -237.043kJ/mol, respectively. Therefore, it can be concluded that B38 has a higher binding energy affinity for N501Y mutant-RBD.

Tab. 1 The binding free energy of B38-SARS-CoV-2-RBD and B38-N501Y mutant-RBD complexes

	van der Waal energy	Electrostattic energy	Polar solvation energy	SASA energy	Binding energy
B38-HCoV-19-RBD	-301.133 +/- 120.078	-223.539 +/- 96.519	476.910 +/- 242.734	-39.438 +/- 14.095	-87.200 +/-59.590
B38-N501Y Mutant-RBD	-470.165 +/- 33.058	-237.043 +/- 39.547	643.755 +/- 81.637	-61.284 +/-2.072	-124.736 +/- 85.028

3.3.2 Residues Energy Decomposition

In order to evaluate the binding power of the antibody and the antigen, we used python's MMPBSA. py script to calculate the binding free energy of per-residure on the RBD against antibody, and the specific energy contribution of per-residue can be seen in Figure 7. the greater the negative value of residues(red column), the greater the contribution to antigen-antibody binding.The residue binding energy less than cutoff value (usually -20 KJ/mol) can be considered that play a positive role in the binding affinity of the complex. The energy contribution value (red column) of N501Y variant-RBD is lower than wild type, it indicated that the conformational flexibility of N501Y variant-RBD is higher,which is more conducive to antigen-antibody binding.

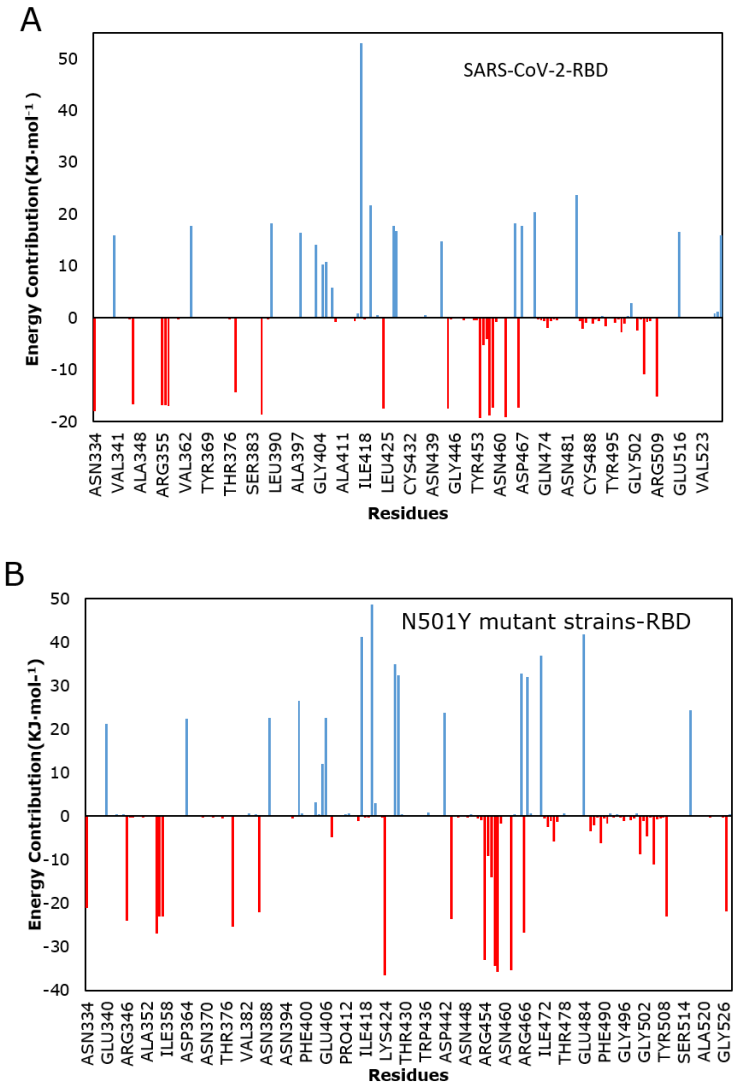

Fig. 7 The per-residue binding free energy contribution of (A) SARS-CoV-2-RBD. (B) N501Y mutant strains-RBD. the red column contribution value is negative, and the blue column contribution value is positive.

Further analyze the energy of these key residues, we found that there are 15 amino acid with the binding free energy of residues less than cutoff value in the N501Y mutant-RBD, they are ASN334, ARG346, ARG355, LYS356, ARG357, LYS378, LYS386, LYS424, LYS444, ARG454, ARG457, LYS458, LYS462, ARG509, LYS527. Whereas in the SARS-CoV-2-RBD, no residues contribution value less than cutoff value, the lowest binding free energy is ARG454 (-19.318 KJ/mol). From the key amino acids mentioned above, hydrophilic and uncharged polar amoni acids could be considered as primary driving forces for RBD and B38 antibody.

4. Conclusion

Mutations in the amino acid sequence of a protein often have an important impact on its structure, dynamics and function. Although the mutation of N501Y only mutated one amino acid on the RBD, it enhanced the flexible configuration of RBD-B38 complex, further make the hydrophilic amino acids show stronger binding free energy. In order to explore the reason for the increased conformational flexibility of RBD after mutation at position 501, the intramolecular interactions of the two complexes were calculated separately. The results showed that the van der Waals force and electrostatic potential energy of the mutant were significantly increased, which led to the increase in conformational flexibility. The higher conformational flexibility can easily realize the relaxation of the substrate binding pocket, so that the B38 antibody can enter and exit the binding pocket more conveniently, thereby improving the efficiency of binding to the substrate.

Acknowledgements

No funding

References

[1] Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579(7798):270–3. 2.

[2] Chan JF-W, Yuan S, Kok K-H, To KK-W, Chu H, Yang J, et al. A familial cluster of pneumonia associated withthe 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 2020;395(10223):514–23

[3] Special Expert Group for Control of the Epidemic of Novel Coronavirus Pneumonia of the Chinese Preventive Medicine Association. An update on the epidemiological characteristics of novel coronavirus pneumonia (COVID-19) [J]. Chinese Journal of Epidemiology, 2020, 41(02): 139-144.

[4] Davies N G, Barnard R C, Jarvis C I, et al. Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/ 01 in England. 2020.

[5] Xie X. et al. An infectious cDNA clone of SARS-CoV-2. Cell Host Microbe 27, 841–848 (2020).

[6] Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367, 1260–1263 (2020).

[7] Wu Y, Wang F, Shen C, et al. A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2[J]. Science, 2020, 368(6496):eabc2241.

[8] Wu Y, Wang F, Shen C, et al. A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2[J]. Science, 2020, 368(6496):eabc2241

[9] Watowich S J, Meyer E S, Hagstrom R, et al. A stable, rapidly converging conjugate gradient method for energy minimization[J]. Journal of Computational Chemistry, 1988, 9(6):650-661.

[10] Yuan X H, Wang Y C, Qu Z Y, et al. Phylogenetic and structural analysis of major surface proteins hemagglutinin and neuraminidase of novel avian influenza virus A H7N9 from chinese patient[J]. Chemical Research in Chinese Universities, 2013, 29(5).

[11] Yuan X H, Wang Y C, Jin W J, et al. Structure-Based High-Throughput Epitope Analysis of Hexon Proteins in B and C Species Human Adenoviruses (HAdVs) [J]. PLoS ONE, 2012, 7(3): e32938.

[12] DA Case, H Gohlke, C Kiel. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes.

[13] Kollman P A, Massova I, Reyes C, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models[J]. Accounts of Chemical Research, 2000, 33(12):889-897. DOI: 10.1021/ar000033j.

[14] Xu Y, Wang R. A computational analysis of the binding affinities of FKBP12 inhibitors using the MM-PBSA method[J]. Proteins: Structure, Function, and Bioinformatics, 2006, 64(4): 1058-1068. DOI:10.1002/prot.21044

[15] Kumari R, Kumar R, Lyn A. g_mmpbsa—AGROMACS Tool for High-Throughput MM-PBSA Calculations[J]. Journal of Chemical Information and Modeling, 2014, 54(7): 1951-1962.

[16] Yu W, Wu X, J Ren, et al. Mechanistic Insights to the Binding of Antibody CR3022 Against RBD from SARS-CoV and HCoV-19/SARS-CoV-2: A Computational Study[J]. Combinatorial Chemistry & High Throughput Screening, 2020, 23.