Binding Profile Assessment of SARS-CoV-2 Spike Protein RBD Mutants with hACE2 Protein Using *In Silico* Methods

by

Yuzhao Zhang

Bachelor of Science, Central South University, 2019

Submitted to the Graduate Faculty of the

School of Pharmacy in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

SCHOOL OF PHARMACY

This thesis was presented

by

Yuzhao Zhang

It was defended on

Select the Date

and approved by

Junmei Wang, PhD, Associate Professor, Department of Pharmaceutical Sciences

Jaden Jungho Jun, PhD, Research Assistant Professor, Department of Pharmaceutical Sciences

Zhiwei Feng, PhD, Assistant Professor, Department of Pharmaceutical Sciences

Dissertation Director: Junmei Wang, PhD, Associate Professor, Department of Pharmaceutical Sciences

Binding Profile Assessment of SAR-CoV-2 Spike Protein RBD Mutants with hACE2 protein

Yuzhao Zhang, B.S

University of Pittsburgh, 2021

SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2), a novel coronavirus, has brought an unprecedented pandemic to the world and affected over 100 million people. The virus infects humans using its spike glycoprotein, which is mediated by a crucial area, receptor binding domain (RBD), to bind to the human ACE2 (hACE2) receptor. Mutations on RBD have been observed in different countries and classified into various types. In this research work, we studied A435S, D364Y, G476S, N354D/D364Y, N501Y, R408I, V341I, V367F, V483A, W436R, N501Y, N501Y/E484K, N501Y/K417N, and N501Y/E484K/K417N, 14 mutant types plus the prototype. Employing molecular dynamics (MD) simulation, we investigated dynamics and structures of the complexes of the prototype and mutant types of SARS-CoV-2 spike RBDs and hACE2. We then probed binding free energies of the prototype and mutant types of RBD with hACE2 protein by using an end-point molecular mechanics Poisson Boltzmann surface area (MM-PBSA) method. According to MM-PBSA binding free energy calculation results, 10 of the mutant types showed enhanced binding affinities with hACE2 compared to the prototype. Our computational protocols were validated by the successful prediction of relative binding free energies between prototype and three mutants: N354D/D364Y, V367F, and W436R. Thus, this study provides a reliable computational protocol to fast assess the existing and emerging RBD mutations. More importantly, the binding hotspots identified by using the molecular mechanics generalized Born surface area (MM-GBSA) free energy decomposition approach can guide the

rational design of small molecule drugs or vaccines free of drug resistance to interfere with or eradicate spike-hACE2 binding. We selected key residues for both RBD and hACE2 to assist binding pocket design and provide six potential binding sites. A docking study using one of the predicted binding pockets revealed COVID-19 treatment potential of molecules from the NPC library. The docking scores were also compared with binding inhibitory bioassay results from NCATS.

Table of Contents

List of Tables

List of Figures

Preface

First, I'd like to thank my two advisors, Dr. Junmei Wang, and Dr. Jaden Jun. Both of you have provided me invaluable advice and tremendous help. It's my great honor to have the chance to work with you two. Dr. Wang has taught me so many things about computational chemistry and PBPK modeling, I admire your wide-acknowledge and hard-working, and I'll take you as my example in my whole research life. I've also worked with Dr. Jun for medicinal chemistry. You always spend time with your students to talk about chemistry mechanism and synthesis skills. I loved the time when I was working with you.

I then want to express my appreciation to my committee members, Dr. Junmei Wang, Dr. Jaden Jun, and Dr. Zhiwei Feng.

I also want to thank my group members, Dr. Xibing He, and Dr. Viet Hoang Man. Both of you are so acknowledged and full of research ideas, it was my pleasure to be in one group with you. Dr. He, you always took efforts in helping me with trouble shooting and scientific writing style. I will always remember your research passion. And I'd also like to give my thanks to Beihong, who helped me get through the hard beginner time as a master's student. It's my fortune to have you as my good friend.

Last, I want to thank the ones I love. My parents: You always have my back and give me so much trust. My friends: Yujia, Lanting, and Beihong. You're the ones who kept me companied during this long working from home time. It's my fortune to have you by my side.

1.0 Introduction

1.1 Overview of COVID-19 Pandemic

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a highly pathogenic novel coronavirus, has started a worldwide pandemic since December 2019[1]. This disease is caused by the virus known as COVID-19, with a wide range of symptoms, including dry cough, fever, headache, dyspnea, and pneumonia and more[2]. As of March 14th, 2021, the virus has affected 219 countries, infected more than 100 million people, and claimed lives of 2 million, with an estimated mortality of about 2~5%[3,4]. The total number of infected people is highly underestimated since part of the mild symptom patients are self-cured without cases recorded.

SARS-CoV-2 belongs to the *betacoronavirus* genus, with family members including SARS-CoV and MERS-CoV. Since the 21st century, the three viruses have trespassed the species barrier and caused unprecedented epidemic situations[5-7]. Among the three coronaviruses, SARS-CoV-2 has the most severe global impact[8]. The genetic material inside SARS-CoV-2 is single-positive-strand RNA, with 1/3 of the gene codes for structural proteins (SPs) and the rest 2/3 of the gene codes for nonstructural proteins (nSPs)[9]. The major structural proteins expressed by SARS-CoV-2 can be divided into four classes: Spike protein (S); Envelope protein (E); Nucleocapsid protein (N); Membrane protein (M)[10]. Same as other *betacoronavirus* family members[11,12], SARS-CoV-2 uses spike protein as the cell entry[13,14]. The spike protein is the protrusion on the virus surface, giving the virus a crown appearance[15]. The spike protein comprises two functional subunits, S1 and S2. The S1 subunit includes the Receptor Binding Domain (RBD)

with a length of about 200 residues targeting a specific protein receptor of host cells called human angiotensin-converting enzyme (hACE2)[16,17]. Within the RBD area, a short sequence called receptor binding motif (RBM) makes direct contact with the receptor. The S2 subunit is the fusion machinery that is responsible for membrane fusion[18]. In the spike protein, RBD is thought to be essential in viral tropism and infectivity[19-21]. Just like SARS-CoV, SARS-CoV-2 infects host cells utilizing hACE2 expressed by HeLa cells as the receptor,[22] and the spike-hACE2 binding triggers a cascade of immune reactions. Normal hACE2 level in the lung is beneficial for the host to combat inflammatory disease, and also crucial for other physiological activities[23,24]. During the virus infection, hACE2 is one indispensable component and is also considered to be one promising therapeutic target.

1.2 Binding Profile Study of SARS-CoV-2 mutants

Binding mechanism study of spike protein RBD and hACE2 may provide the key to tackle this worldwide health threat. Complementary to wet lab study, computational methods can provide more details about the binding pattern and the protein-protein interaction dynamics. Also, computational tools like molecular dynamic (MD) simulation are potent in protein mutant study. The binding mechanism elucidated by virtual *in-silico* studies can provide hints on designing effective drugs and vaccines to overcome the emergent global challenge. More importantly, mutant models can be easily built through computational mutagenesis and the mutation effect can be easily predicted using the same computational protocol adopted to study the prototype. The hotspot residues, which are essential for both the prototype and mutants, can be identified by analyzing the RBD/hACE2 complexes binding profile. Drug molecules or vaccines that target hotspot

residues are likely to be free of drug resistance. It is hard to analyze which mutant is more infectious statistically as we will not know which mutant has infected the patient. Moreover, the mutant prevalence analysis among population is only based on a small fraction of the collected samples, which made it hard to capture the big picture. Also, it's very hard to collect and purify mutant protein, which is a major obstacle using wet-lab method to study spike protein mutants. Therefore, we focused on *in-silico* approaches to study the binding mechanisms of spike protein RBD/hACE2, aiming to provide insights into binding profile, and rational drug/vaccine design.

1.2.1 First-Wave SARS-CoV-2 Spike protein RBD mutants

Spike RBD is the only protein domain initializes the viral infection process of SARS-CoV-2, and it is also a highly conserved area[25]. Thus, harmful mutations on this protein domain may lead to drug resistance. Under this context, it's crucial to monitor the mutation dynamics of the virus, and the mutants should be well studied to fully capture their effect on the protein-protein interaction. From December 2019 to December 2020, strains of SARS-CoV-2 have been collected from multiple countries, and mutants have been detected and classified into nine types according to mutation positions on RBDs, including V341I, F342L, N354D/D364Y, V367F, R408I, A435S, W436R, G476S, and V483A. In this study, models of mutant RBDs were generated using the prototype SARS-CoV-2 RBD/hACE2 complex for the computational simulations. To thoroughly study the only dual amino acids mutant, N354D/D364Y, single mutation models of N354D, and D364Y were also built for the subsequent studies, including molecular dynamics (MD) simulations of the protein complexes, end-point MM-PBSA-WSAS binding free energy calculations, and end-point MM-GBSA binding free energy decomposition analysis (see the Methods section). The binding affinity calculations can reveal which mutations are able to strengthen the protein-protein

binding. The conformation analysis shed light to vaccine resistance. Binding free energy decomposition study further reveals how binding profile is altered by point mutations on RBD. The hotspot residues from both spike RBD and hACE2 were identified by analyzing the binding free energy decomposition data. The common hotspots occurring to both the wild type and harmful mutants should be the major targets in binding pocket definition for structure-based drug design and vaccine development.

1.2.2 Second Wave SAR-CoV-2 Spike Protein RBD mutants

Since December 2020, N501Y has been detected in numerous countries including the UK, the US, and Canada, and is thought to be 75% more infectious than the wild-type[26]. The new mutant was officially reported by the UK where a rapid increase of COVID-19 cases occurred in December 2020[27]. According to the report, the new mutant accounts for about 60% of new infections in London, suggesting that the new variant is highly transmissible[28]. After the report of N501Y, UK imposed a harsh lockdown policy to prevent the new variant from spreading. Now, the world is closely monitoring this new mutant. As of Jan 8th, 2021; 63 N501Y infections were reported in the US.[29] Though the variant has not been thought to cause more severe symptoms yet, it must be taken under control due to its enhanced infectivity.

After the report of N501Y, another 3 mutants related to N501Y mutant were also detected: N501Y/E484K, N501Y/K417N, and N501Y/E484K/K417N[26]. All these three N501Y related mutants were also believed to have enhanced binding affinity with hACE2. In this study, these four second-wave mutants including N501Y were added on the first-wave mutants to provide a greater view of spike protein mutants.

The 501V2 variant was also known as South Africa variant. It was first detected in the Nelson Mandela Bay and reported on December 2020. The 501V2 including three mutations, N501Y, E484K and K417N. Two of the three mutations locate at the RBM of the RBD, the N501Y and the E484K. The K417N mutation also has called great attention as it locates at the area where the neutralizing antibody targets at. The variant has spread to the world in a fast speed. As of 20th February 2021, the variant has accumulated more than 2000 cases[30]. It's no doubt that the study of this mutant is under emergency.

At the structural level, the second-wave mutants probably bind more tightly to the human angiotensin-converting enzyme 2 (hACE2) so that it may have better chance to infect people than other types do. Although it is not conclusive, the N501Y variants can reduce neutralization sensitivity to convalescent sera and monoclonal antibodies[31,32]. As the crystalized structure of the new mutant hasn't been discovered yet, it's our virtue to apply all the resources to uncover the mystery of the new mutants. In this work, a model was built for the 501V2 mutants RBD/hACE2 based on the prototype RBD/hACE2 complex crystal structure. Extensive molecular dynamic (MD) simulations were performed to study the binding characteristics of the mutant to hACE2. The binding affinity of the N501Y and its related mutants were then compared to the prototype RBD with hACE2 applying the same computational protocol built for the first-wave RBD mutants. The aim is to demonstrate the distinct structural features of the mutant and its potential effect on the vaccines, to elucidate the binding free energy change at the residue level of the RBD, and to provide reasonable structures of N501Y mutant to community for the sake of structure-based drug design.

1.3 Computational Tools in Drug Design

Computational techniques such as MD simulation and molecular docking have been extensively applied in the early-stage of drug discovery. Different from wet lab method, computational method is famous for its high output and efficiency. More importantly, computational tools can provide a unique molecular perspective to understand the binding dynamics and mechanism.

1.3.1 Molecule dynamics simulation

MD simulation is a technique that studies the movement of molecules in a complex over time to help understand the dynamic activity of biomolecules in full atomic detail and at very fine temporal resolution[33]. MD simulation was universally used in an extensive content, including conformational change, protein-ligand binding, protein folding, and atomic position. The most important part of MD simulation is to understand the response of biomolecules at an atomic level, which is a great solution to study protein mutants[34]. Another advantage of MD simulation is the carefully restrained conditions, such as the receptor's conformation, the binding site of the ligand, and the mutations on the protein. All of these virtues made MD simulation an ideal tool to study the molecular dynamics mechanism. In this study, we used MD simulation as the main technique to study the mutants of SARS-CoV-2 spike protein. The goal of applying MD simulation is to capture each mutants' movement when binding to the human receptor, hACE2, in order to study the differences and commonality among the mutants.

1.3.2 Molecular docking study

Molecular docking is a computational drug design method that predicts and studies the binding between two components[35]. The content of the components varies from case to case: it can be fragment[36], small molecule[37], peptide[38], or even macromolecule[39]. Docking method enables researchers to identify novel targets of therapeutic interest, provides insight into the binding process, or even helps structure-activity relationship study. In this study, we applied molecular docking to find potential hits for SARS-CoV-2 using small molecules and spike protein RBD as ligands and receptors.

This study applied a series of computational methods to investigate the binding profile of 15 SARS-CoV-2 RBD mutants and the wild-type SARS-CoV-2 RBD. The study subjects include mutant binding affinity towards the human target, the complex stability of RBD/hACE2, the hotspot residues, and the preliminary drug screening. We provided insights into residue-residue interactions and protein-protein interaction inside the RBD/hACE2 complex to better understand why mutants vary with binding affinity. Also, through a series of post-simulation analysis methods, we provided guidance for drug design and vaccine design. A molecular docking study was also accomplished to screen potential therapeutic molecules for interrupting RBD/hACE2 binding. We utilized experimental data from previous reports to validate and strengthen the reliability of our computational methods and models.

2.0 Methods

2.1 Binding Profile Study of SARS-CoV-2 Mutants

We studied SARS-CoV-2 mutants binding profile using a series of computational tools. We built the prototype RBD/hACE2 model using the crystalized structure reported and then modified it to construct mutant type models. We performed multiple independent MD runs for each system to understand the binding activity dynamics better. Various post-MD analyses such as MM/PBSA-WASA and MM/GBSA methods were applied to analyze the MD simulations and give clues in binding affinity, structure stability, and drug design.

2.1.1 Molecular simulation system setup

The complex of prototype SARS-CoV-2 RBD/hACE2 complex structure (PDBID 6M17)[40] was obtained from Protein Data Bank[41]. The models of mutant type SARS-CoV-2 spike glycoprotein were built based on the prototype RBD/hACE2 complex by mapping the common atoms of the original and mutation residues and manually rotating side chains to maximize the favorable interactions with the surrounding residues. Three types of glycosylated residues, called AS1, AS2, and AS3, are modified ASN residue with one N-Acetylglucosamine (NAG), two NAG, and three NAG residues, respectively, were introduced. A modified GLU and HIS residues which are covalently bonded to Zn^{2+}, were also prepared using programs in the Antechamber package[42]. The atomic partial charges of those nonstandard residues were derived by the RESP[43] program to fit the HF/6-31G* electrostatic potentials generated using the Gaussian 16 software package[44].

FF14SB[45] was used for modeling proteins except for NAG, which was described by GAFF[46]. In total, we studied 15 RBD/hACE2 systems. Each system contained a copy of RBD/hACE2 protein, 120 Cl^- and a certain number of Na^+, which neutralized the whole MD system and about 40,000 TIP3P[47] water molecules. The simulation systems were rectangles with sizes of roughly 110 × 110 × 110 Å after equilibrium.

2.1.2 Molecular dynamic simulation

We performed Molecular mechanics (MM) minimization and the sequential molecular dynamics (MD) simulations using the AMBER18 package[48]. First, five 10000-step restrained minimizations were performed with the restraining forces on the main chain atoms gradually decreased from 20 to 10, 5, 1, and 0 kcal/mol. The followed MD simulations have four phases, including the relaxation phase, the heating-up phase, the equilibrium phase, and the sampling phase. In the relaxation phase, five 200-picosecond MD simulations using the same restraining forces to the main chain atoms as in minimization stage. Then the MD system was heated up progressively from 50K to 250K at steps of 50K in a series of 1-nanosecond MD simulations. In the next equilibrium phase, the system was equilibrated at 298K, 1 bar for 10 ns. Last, a 100-nanosecond MD simulation was performed at 298K, 1 bar to produce an isothermal–isobaric ensemble. The last phase of the simulation was repeated for all the mutant types. For the important mutant type: prototype, N501Y, N501Y/E484K, N501Y/N417N, and N501Y/E484K/N417N, the last phase was repeated another four times; the other mutant types were repeated one time. The repeated MD runs produced independent MD trajectories by using different random number seeds for temperature regulation using Langevin dynamics[49] with a collision frequency of 5 ps^{-1}. The integration of the equations of motion was conducted at a time step of 1 fs for the relaxation phase

and 2 fs for the other three phases. The Particle Mesh Ewald (PME) method[50] was used to calculate the full electrostatic energy of a unit cell in a macroscopic lattice of repeating images. In total, we collected 10,000 snapshots (for mutants have one repeated MD run) or 50000 snapshots (for mutants have four repeated MD runs) from the sampling phase for post-analysis using the Cpptraj module implement in the AMBER software package[51].

2.1.3 MM-PBSA-WSAS free energy calculation and MM-GBSA energy decomposition

Molecular Mechanics/Poisson Boltzmann Surface Area (MM-PBSA) is an end-point method[52-58] for free energy calculations, with the solvation free energy being calculated using the PBSA method and the conformational entropy being estimated using the WSAS method[59]. For a molecule in a solvent, we calculated the free energy using the following equations.

$$\Delta G_{MM-PBSA-WSAS} = \Delta H - T\Delta S$$

$$= \Delta E_{int} + \Delta E_{vdw} + \Delta E_{eel} + \Delta G_p^{sol} + \Delta G_{np}^{sol} - T\Delta S$$

ΔE_{int} stands for the internal energy contribution, which is canceled out when applying the "Single Trajectory" sampling protocol as we did in this study[60]; ΔE_{vdw} and ΔE_{eel} are the van der Waals and gas phase electrostatic energies, respectively; $\Delta G_p^{sol}$and ΔG_{np}^{sol} stand for the polar and nonpolar components of the solvation free energy, respectively. T is the absolute temperature; ΔS is the change of the conformational entropy. ΔG_p^{sol} is calculated by solving the Poisson-Boltzmann equations using the Delphi program[55]. ΔG_{np}^{sol} is estimated using solvent accessible surface area with the surface tension coefficient of 0.00542 kcal/(mol·Å^2) and a constant of 0.92 kcal/mol[61]. For each MD trajectory, MM-PBSA-WSAS calculations were performed for 200 evenly selected

snapshots. For MM-GBSA energy decomposition[62], on the other hand, we performed for all the 10,000/5,0000 snapshots (as mentioned above). For MM-GBSA analysis, the polar component of solvation free energy is calculated using the Generalized Born model developed by Hawkins et al.[63]. The internal and external dielectric constants were set to 1 and 80, respectively, for both PBSA and GBSA calculations. The free energy decomposition analysis was performed using an internal program. An RBD residue becomes a hotspot when its interaction energy with hACE2 is smaller than a cutoff, -0.1 kcal/mol. The same rules apply; an hACE2 residue is a hotspot when its interaction energy with RBD is more minus than the cutoff. The hotspot residues with extremely great contributions were selected as crucial residues to provide inspiration for binding pocket identification.

2.2 Identification of RBD Hits

We identified hits for RBD on SARS-CoV-2 spike protein using one of the binding pockets suggested from MD simulation analysis results. The molecules used in this section were FDA-approved drugs. To dig deeper in this subject, the docking study result was associated with experimental data to find the underlying ligand-receptor mechanism.

2.2.1 Molecular docking

We used the NCGC Pharmaceutical Collection (NPC) as the drug screening library. The NPC molecules underwent bioassay screening performed by National Center for Advancing Translational Sciences. Molecules collected from NPC underwent preparation to output as many

conformations as possible. All the conformations were later used in the docking study using Glide Docking. The binding pocket was selected using the one pre-docked ligand as reference located at the position shown in Figure 10A. The Glide docking used the standard precision version of docking scoring function; the scaling factor was set in 0.8; the partial charge cut-off was 0.15; the sample nitrogen inversions and sample ring conformations options were turned on; intramolecular hydrogen bonds were rewarded, and Epik state penalties was added to the docking score; for each ligand input, at most 2 poses were written out, and the best score was kept for further analysis. To validate our docking results and find the underlying mechanism, the docking scores were compared to experimental data collected from NCATS.

3.0 Results and Discussion

We first performed multiple MD simulations to assess the structural stabilities of the wild type and mutant types protein complexes. We then applied the collected MD snapshots to calculate the binding free energy of spike RBD/hACE2 interactions, and to conduct binding free energy decomposition. The residues of great energy contribution were selected to provide guidance for binding pocket design. Then we performed molecular docking using drugs from NIH NCATS database to one selected binding pocket on the RBD to find promising therapeutic molecule.

3.1 Binding Profile Characterization of RBDs/hACE2

3.1.1 Structural Stability of Prototype and Mutant types of RBD/hACE2 complex during MD Simulation

Five independent MD runs were performed for the prototype spike RBD/hACE2 complex. For each of the mutant types including V341I, F342L, N354D, N354D/D364Y, D364Y, V367F, R408I, A435S, W436R, G476S, and V483A, we carried out two independent MD runs. For the N501Y related mutants including N501Y, N501Y/E484K, N501Y/K417N, and N501Y/E484K/K417N, we carried out five independent MD runs. Each independent MD run lasts 100 nanoseconds.

The root-mean-square deviations (RMSDs) of the RBD and hACE2 in each system along with the simulation time were calculated from the MD trajectories and shown in **Figures S1** and

S2, respectively. The RMSD plots showed that both RBD and hACE2 proteins from each system reached equilibria in the equilibrium phase (20 ns). The stable RMSD values were around 1~2.5 Å for most systems. The low RMSD values of RBD and hACE2 indicated the complex stability of the MD systems is satisfactory during the simulation time.

Representative MD structure, which has the closest RMSD value compared to the average MD structure RMSD value, was selected for all (16) of the MD systems. As the prototype RBD/hACE2 complex was built based on crystal wild-type RBD/hACE2 complex structure, we overlaid the two structures to see difference between the two structures. Interestingly, the representative prototype RBD/hACE2 MD structure was very similar to the crystal structure, especially for the RBD/hACE2 binding interface (**Figure 1A**). The good overlap between two structures indicate the reliability of our computational model.

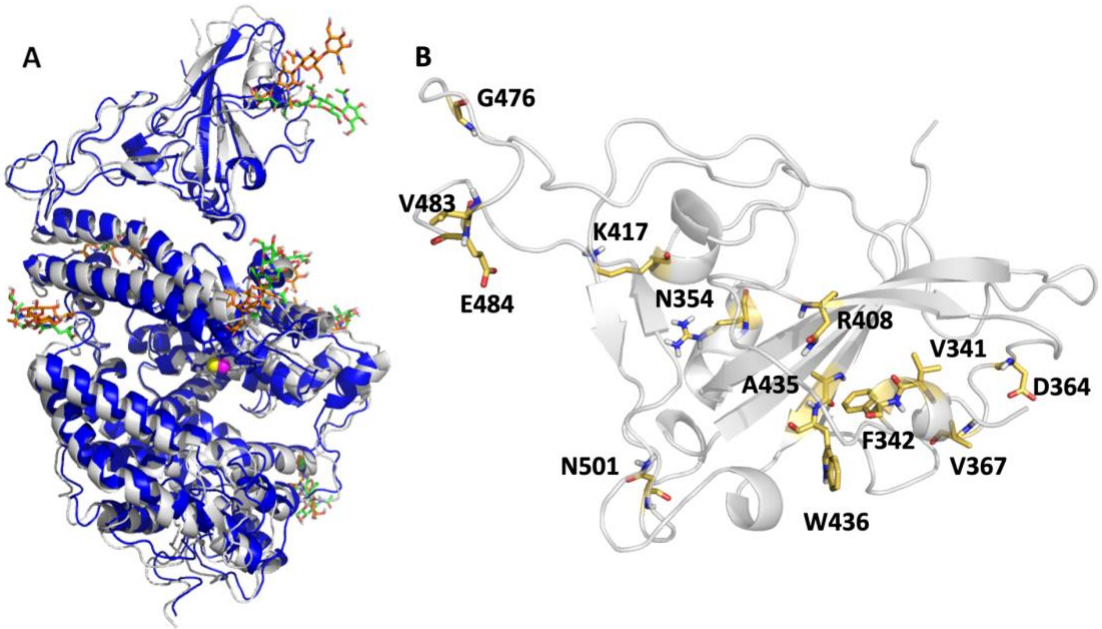


Figure 1. Overlay of crystal and MD representative structure for the wild type RBD/hACE2 (Panel A) and the 13 RBD mutation sites (Panel B).

Glycosylated residues are shown in greenish sticks for the X-ray and brownish sticks for representative MD structures. The residues undergo mutations are shown as greenish sticks.

For each mutation position we studied, the key interactions in prototype and in the mutant types are shown in **Figure S3 and S4**, respectively. As shown in the **Figure 1B** and **Figure S3A**, 9 mutation sites are closed to RBM including Gly476 (giving G476S mutant), Arg408 (giving R408I mutant), Val484 (giving V483A mutant), and 3 N501Y related mutations sites (Asn501, Glu484, and Lys417). Though Gly476, Arg408, Val484, and Glu484 were closed to the interface, no polar interaction of these three residues with the hACE2 residue was observed. As for Asn501 and Lys417 mutation site, polar interactions were observed indicating the structural importance of residues located at these positions. Residues Asn354, Asp364, Val367, Arg408, Ala435 and Trp436 had strong interactions with nearby residues, whilst residues Val341, Phe342, Gly476 and Val483 interacted with surrounded residues weakly. The interaction with nearby residues may be altered due to mutation. For example, Asn354 in the prototype can form multiple hydrogen bonds with Ala348 and Ser399; after Asn354 changed to Asp354, its interaction with Ala348 disappeared, which lead to a binding affinity drop. For V483A case, no polar contact with nearby residues has been observed before or after the mutation. And no significant binding affinity alteration showed up for this mutant. By comparing the interaction with nearby residues before and after mutations, we found that the greater interaction change of the mutated residues with its nearby residues, the greater change of the RBD-ACE2 binding affinity. To be specific, V341I, A435S, W436R, G476S, and V483I, most of the interactions formed before remained after the mutation, and only slight binding affinity change was observed for these types. For cases of F342L, N354D, N354D/D364Y, D364Y, V367F, and R408I, former interaction disappeared and new strong interaction formed, and drastic binding affinity changes were observed (**Figure 2**).

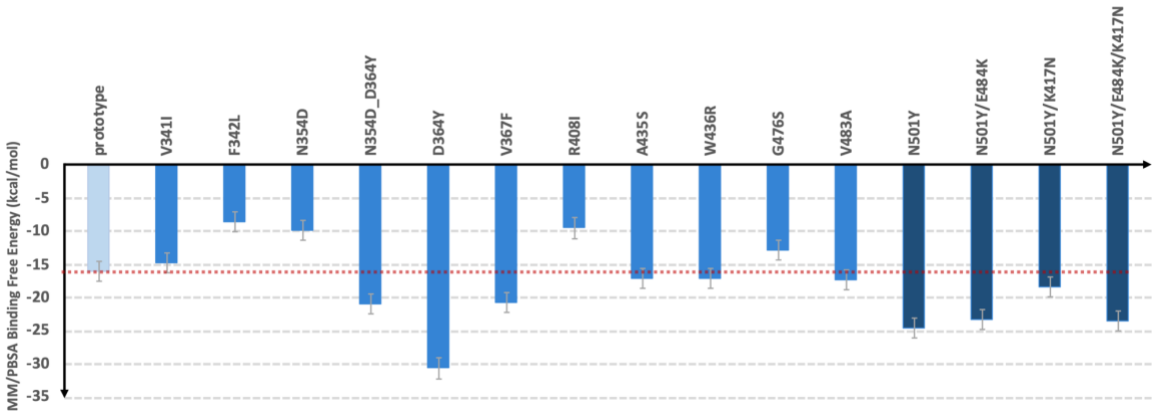


Figure 2. The comparison of binding free energy of each mutant and the prototype SARS-CoV-2 RBDs with hACE2.

Overall, the residues which had strong interactions with the nearby residues were more likely to alter the binding free energy. This may explain why some of the mutants showed the altered binding affinity, but the other ones had comparable affinity as the wild type. By comparing the changes of conformation and the interaction pattern before and after the mutation, we can estimate the strength of the perturbation induced by a mutation at the mutation site.

3.1.2 Binding Free energies of RBDs/hACE2

The binding affinity between prototype SARS-CoV-2 spike and human ACE2, or a mutant SARS-CoV-2 spike protein and human ACE2 are represented by the binding free energy (ΔG) between the RBDs and the hACE2. To obtain the ΔG values, we conducted Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) calculations for the MD snapshots sampled in the product phase of the RBDs-hACE2 simulations. MM-PBSA is a popular end-point free energy calculation method[64]. We estimated the entropic contribution using a method called WSAS [59], as such, we call this binding free energy calculation method MM-PBSA-WSAS. The results of MM-

PBSA-WSAS calculations were summarized in the **Table 1**, and the comparison between mutants and the prototype was shown in **Figure 2**.

Table 1. Results of calculated MM-PBSA energy terms and binding free energies of prototype and mutant RBD/hACE2 systems.

Mutant Type	ΔE_{vdw}	ΔE_{eel}	ΔG_p^{sol}	ΔG_p^{sol}	$T\Delta S$	$\Delta G_{MM\text{-}PBSA}$
Wild Type	-91.54 ± 0.22	-615.23 ± 0.90	667.81 ± 1.10	-10.10 ± 0.01	-33.02 ± 0.04	-16.04 ± 0.05
V341I	-91.85 ± 0.85	-584,95 ± 5.67	640.29 ± 5.38	-9.91 ± 0.06	-31.64 ± 0.26	-14.77 ± 0.83
F342L	-93.96 ± 0.56	-564.05 ± 6.68	627.49 ± 5.90	-10.43 ± 0.04	-32.37 ± 0.22	-8.59 ± 0.85
N354D	-93.20 ± 0.63	-415.49 ± 3.40	476.86 ± 4.00	-10.36 ± 0.06	-32.29 ± 0.14	-9.90 ± 1.17
N354D/D364Y	-93.19 ± 0.81	-584.45 ± 5.67	634.35 ± 4.27	-9.94 ± 0.08	-32.39 ± 0.07	-20.84 ± 1.41
D364Y	-96.07 ± 0.24	-754.67 ± 1.89	797.13 ± 2.15	-10.61 ± 0.06	-36.26 ± 0.03	-30.61 ± 0.47
V367F	-97.29 ± 0.43	-604.85 ± 3.1	657.86 ± 2.4	-10.45 ± 0.03	-34.08 ± 0.21	-20.66 ± 1.3
R408I	-88.52 ± 0.93	-410.55 ± 2.13	468.19 ± 2.77	-10.12 ± 0.19	-31.48 ± 0.26	-9.53 ± 0.55
A435S	-101.03±0.65	-554.74 ± 3.41	614.91 ± 3.82	-11.03 ± 0.06	-34.83 ± 0.16	-17.07 ± 0.35
W436R	-93.92 ± 0.71	-831.27 ± 5.51	884.57 ± 4.39	-10.41 ± 0.04	-33.92 ± 0.17	-17.12 ± 1.51
G476S	-94.04 ± 0.67	-585.17 ± 3.49	643.48 ± 4.13	-10.55 ± 0.04	-33.50 ± 0.18	-12.79 ± 1.22
V483A	-91.31 ± 0.58	-622.57 ± 6.54	673.45 ± 6.48	-10.69 ± 0.04	-33.82 ± 0.15	-17.30 ± 1.57
N501Y	-96.43 ± 0.17	-638.99 ± 0.31	686.94 ± 0.62	-10.18 ± 0.01	-34.17 ± 0.03	-24.48 ± 0.19
N501Y/E484K	-96.24 ± 0.15	-1031.75±1.55	1081.47±1.28	-10.48 ± 0.01	-33.75 ± 0.04	-23.24 ± 0.28
N501Y/K417N	-96.12 ± 0.20	-344.10 ± 0.66	398.97 ± 0.57	-10.24 ± 0.01	-33.22 ± 0.05	-18.27 ± 0.33
N501Y/E484K/ K417N	-102.07±0.36	-822.37 ± 0.78	876.33 ± 1.14	-10.80 ± 0.03	-35.42 ± 0.04	-23.50 ± 0.68

Compared to -16.04 kcal/mol binding free energy of the prototype SARS-CoV-2 RBD system, three mutant types, N354D/D364Y, D364Y, and V367F, showed significant lower ΔG values indicating the significant higher binding affinities with hACE2. Two of these three mutants, N354D/D364Y and V367F, have drawn extensive attention as they have been reported from multiple countries, indicating its enhanced binding affinity in the real world[65]. Four mutant types showed comparable binding affinity with the wild type, which are V341I, A435S, W436R, and V483A. Four mutant types showed lower binding affinity than the prototype system, which are F342L, N354D, R408I, and G476S. From other publication[66], R408I was once reported to have a lower binding affinity, which is consistent with our study. In the case of N354D/D364Y, D364Y mutation significantly enhanced the binding affinity, while N354D lowered the binding affinity, indicating the mutation on Asp364 contributed more to the increment of the mutant N354D/D364Y

binding affinity. It is pointed out that the single D364Y mutant has not been observed in any country. With this observation, we have one theory that the extremely high binding affinity is not beneficial to the SARS-CoV-2 cell entry. If the binding of RBD with the hACE2 is too high, it will be hard for the spike conformational change which is necessary for the virus cell entry. The predicted binding free energies were compared with experimental data from multiple reports applying different experimental techniques. **Figure 3** shows how well the binding free energies by MM-PBSA-WSAS are consistent with the experimental values reported by AcroBiosystems[67]. The original experimental values were listed in **Table S1** in the **Supporting Information** (**SI**).

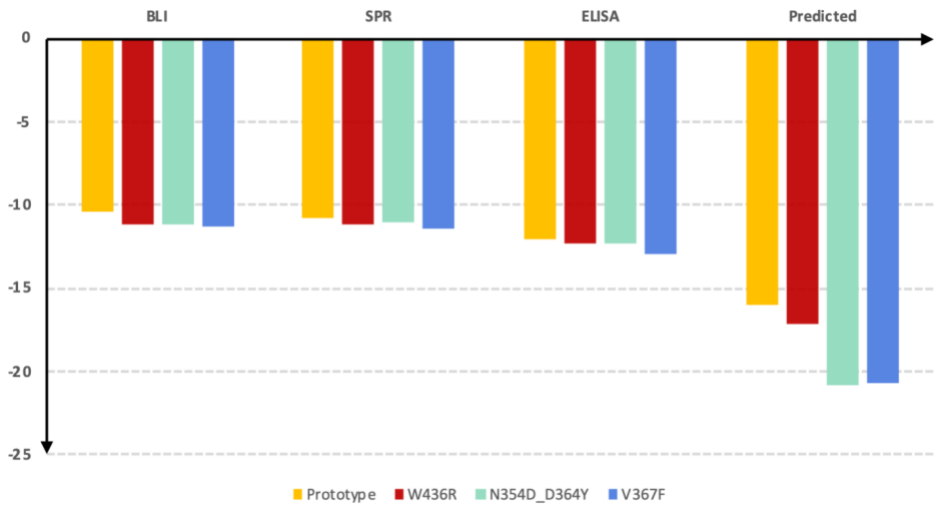


Figure 3. The comparison of experimental results and simulation results.

BLI is Bio-Layer Interferometry method; SPR is Surface plasmon resonance method; ELISA is enzyme-linked immunosorbent assay. $\Delta G = RTlnK_D$ is used to convert K_D to binding free energy for BLI and SPR methods. $\Delta G = RTlnEC_{50}$ is used to convert EC50 to binding free energy for ELISA method.

The K_D and EC_{50} values were converted into ΔG values using equation $\Delta G = -RTlnK_D$ or $\Delta G = -RTln\mathrm{EC50}$. The MM/PBSA binding free energy calculation results suggested that all the three mutants (N354D/D364Y, V367F, and W436R) can enhance the binding between spike

RBD and hACE2, which agrees with the experimental data. The predicted binding affinities in our study were about 0.5-fold higher than the experimental values, with binding affinity of W436R was slightly higher than that of the prototype, and other two types, V367F and N354D/D364Y, exhibiting more higher binding affinities (**Table 1** and **Table S1**). In summary, our MM-PBSA-WSAS method correctly captured the trend of the binding affinity increment of the aforementioned three mutants than the prototypes measured by experiment, albeit it overestimated the absolute values of the binding free energies. Compared to a limited number of experimental binding affinity data, our *in silico* method can distinguish the difference of mutation effects for a broad set of mutations, and offers a clearer view on the direction of future studies.

3.1.3 RBDs/hACE2 complex energies

Unlike the first principal method, molecular mechanics energies are not comparable for different molecules. In Rosetta force field, a score12 term was introduced to account for different amino acid type in protein design. This term reflects the energy of an amino acid in unfolded state [68,69]. This score12 has a narrow range (< 2 kcal/mol) for different amino acid types. Considering the larger variation of the MM-PBSA-WSAS complexation energies between different mutants, the reference energy correction applied in the Rosetta force field was not made in this work. The free energy of a complex mainly reflects the stability of the complex. The MM-PBSA-WSAS free energies of the prototype and 15 mutants were listed in **Table 2** and shown in **Figure 4**.

Table 2. Complex energies of RBDs/hACE2.

The system name refers to the RBDs/hACE2 complex. The unit of complex free energies is kcal/mol.

System	Wild Type	V341I	F342L	N354D	N354D /D364Y	D364Y	V367F	R4081
Complex Energies (kcal/mol)	-27800	-27784	-27792	-27774	-27730	-27742	-27834	-27606
sd	95	290	91	87	86	94	284	296
System	**A435S**	**W436R**	**G476S**	**V483A**	**N501Y**	**N501Y /E484K**	**N501Y /K417N**	**N501Y /E484K /E417N**
Complex Energies (kcal/mol)	**-27791**	-27988	-27743	-27808	-27723	-27682	-27780	-27732
sd	**89**	299	89	303	90	87	91	87

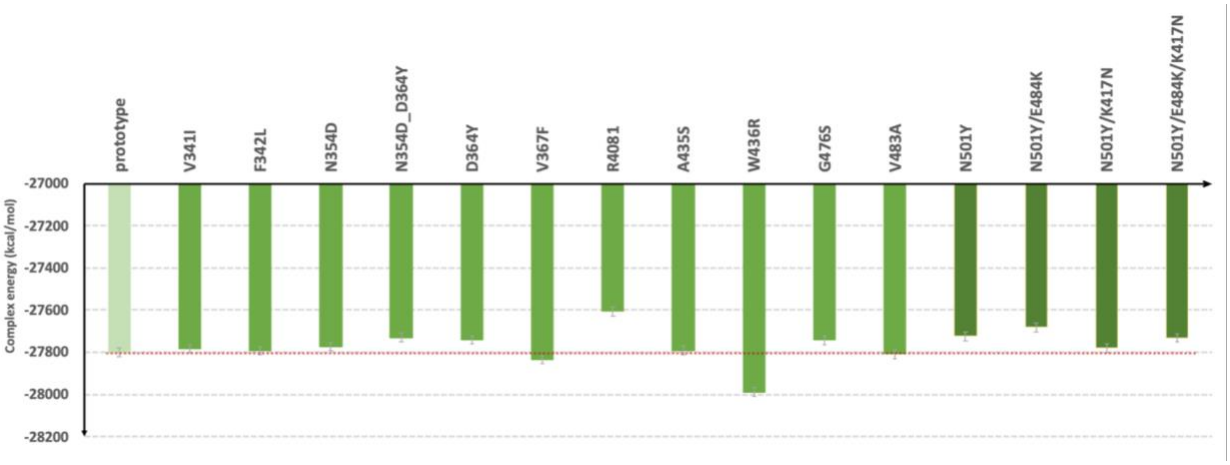


Figure 4. The complex energies comparison of each RBD/hACE2 system.

Among all the first-wave mutant systems, most of the RBD/hACE2 complexes had comparable free energies with that of the prototype except for R408I and W436R. R408I exhibited a significantly higher complex free energy than the prototype, while W436R had a significantly lower complex free energy than the prototype. This observation is understandable as R408I has a charged amino acid (Arg408) being replaced with a neutral one (Ile408), and the latter mutant has a neutral amino acid (Trp436) being replaced with a charged one (Arg436). For the R408I case, in

the prototype Arg408 forming a salt-bridge with the Aap405, moreover, it can also interact with Gln414 by forming a side chain-side chai hydrogen-bond (**Figure S3G**). From Arg to Ile, the basic amino turned into one hydrophobic amino acid, and the mutation disrupted the interactions with the polar amino acids, explaining why a single mutation has such a significant effect on the whole complex free energy. This change in complex might not have an effect on the binding affinity, but it can affect the stability of the protein complex. The unfavorable complex energy and less binding affinity make R408I less risky compared to other mutants. On the other hand, as for W436R, the Trp436 formed two hydrogen bonds with Arg509 in the wild type (**Figure S3I**); after it was mutated into Arg, four hydrogen bonds formed between Arg436 and Arg509/Ser373 (**Figure S4H**). For this reason, the W436R mutant achieves the best complex energy than the other RBD/hACE2 proteins, suggesting its high structural stability. Also, the mutant's binding affinity was only slightly higher than the prototype according to the MM-PBSA-WSAS calculation. The infection of SARS-CoV-2 need binding of spike protein with hACE2 as well as the unbinding of spike protein/hACE2, which means that the strong complex stability is not beneficial for the cell entry. Under this circumstance, W436R may not be an infective mutant. For the case of RBD/hACE2 binding, a high complex stability may impair the infectivity. N501Y, N501Y/E484K, N501Y/K417N, and N501Y/E484K/K417N exhibit significantly higher complex energy indicating lower complex stability. Also, these mutants have properly enhanced binding affinity (**Figure 2**), which equipped them with high infectivity according to our theory. Combined with the truth of prevalence of these N501Y related mutants, it further validates our theory that the relatively low RBD/hACE2 complex energy may help enhance infectivity. N354D/D364Y and V367F are two possibly more infectious mutants among the first-wave mutants. N354D/D364Y

has a properly enhanced binding affinity and lower complexation stability; V367F has comparable complex stability with the prototype complex but has a properly enhanced binding affinity.

3.1.4 Binding Pattern of Prototype SARS-CoV-2 RBDs/hACE2

Through the MM-PBSA-WSAS calculation, the binding free energy for the prototype SARS-CoV-2 spike protein with human ACE2 protein is -16.04 kcal/mol, which is closed to -12.16 kcal/mol reported[22]. As MM-GBSA-WSAS method tends to overestimate the absolute values of binding free energies, it's reasonable that the predicted binding free energies are more negative than experimental values. We conducted binding free energy decomposition for the prototype and mutants of RBD/hACE2 using the MM-GBSA method since its computational cost is much lower than MM-PBSA-WSAS. Hotspot residues were identified according to its interaction energy with the protein binder, $\Delta\Delta G_{inter}$. The $\Delta\Delta G_{inter}$ of a RBD residue is its interaction energy with hACE2, while $\Delta\Delta G_{inter}$ of a hACE2 residue is its interaction energy with the RBD domain of the SARS-CoV-2 spike protein. The hotspot residues for the wild type were listed in **Tables S2** and illustrated in **Figure 5**. We used a color scheme to indicate the interaction strength measured by $\Delta\Delta G_{inter}$, and it showed that most strong interactions occurred to the binding interface residues which are colored in red and blue. The hotspot residues that strongly interact with their protein binder were labeled and shown in **Figure 5B-5F**. Termination of the strong contacts may directly interfere with the binding between two proteins. The residues positions of the selected hotspots were compared with RBM sequence and shown in the **Figure 6**. RBM sequence was from 438-506 and most of the significant hotspot residues were within RBM.

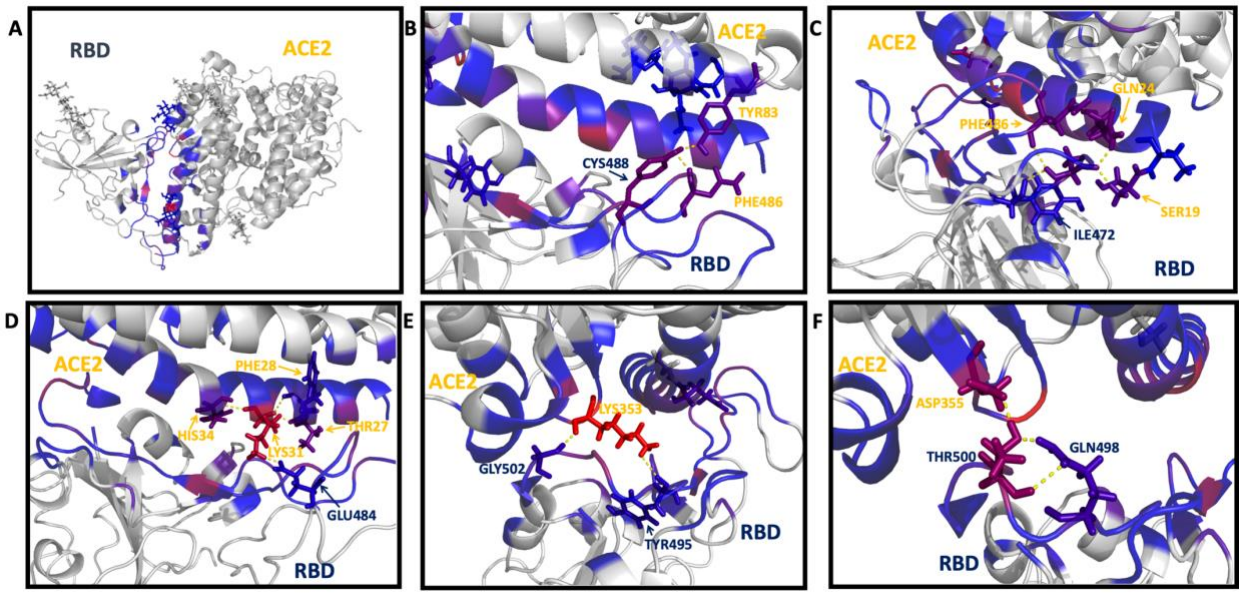


Figure 5. The binding hotspot of prototype SARS-CoV-2 RBD with the human ACE2.

The binding interface was colored in red and blue to show the interaction strength between the prototype RBD/hACE2 complex. The red color indicates strong interaction; the blue color indicates mild interaction. The grey color indicates no interaction or the negligible interaction. The binding interactions between the interface were labeled with residue names. The residues from RBD are labeled in dark blue, and the residues from ACE2 are labeled in yellow. Panel B shows the interaction between Cys488 (RBD), Tyr83 (ACE2) and Phe486. Panel C shows the interaction of Ile472 (RBD) with Gln24 (ACE2), Ser19 (ACE2), and Phe486 (ACE2). Panel D shows the interaction of Glu484 (RBD) with Thr27 (ACE2), Phe28 (ACE2), Lys31 (ACE2), and His34 (ACE2). Panel E shows the interaction of Lys353 (ACE2) with Tyr495 (RBD) and Gly502 (RBD). Panel F shows the interaction of Asp355 (ACE2) with Gln498 (RBD) and Thr500 (RBD).

```
RBD 336 CPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSN
    395 VYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKY
                                                          438
    454 RLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVV
            RBM                                                506
    513 LSFELLNAPATVCGPKLSTDLIK
```

Figure 6. RBD sequence.

The RBD sequence starts from Cys336 to Lys535. The residues selected in the hotspot were highlighted; the residues starting from Thr438 to Gln 506 in RBM were colored in red to show the overlap between the binding area and the RBM motif.

3.1.5 Key Residues in RBDs/hACE2 Binding

The key residues for RBDs/hACE2 binding were identified by analysis of the MM-GBSA binding free energy decomposition results. The residues with a decomposed binding energy $\Delta\Delta G_{inter} < -0.1$ kcal/mol were selected as key residues and listed in **Tables S3** and **S4**. We generated heatmaps to show the overall contributions of each key residue to the protein-protein binding for both the prototype and mutants (**Figures 7** and **8**). In a heatmap, the selected key residues were listed vertically and their interaction energies in different systems were mapped into colors using a colormap shown below. Two heatmaps, for RBD and hACE2, were generated. In the RBD heatmaps, $\Delta\Delta G_{inter}$ is the interaction energy between an RBD residue with the whole hACE2 protein (**Figure 7**), while in the hACE2 heatmap, $\Delta\Delta G_{inter}$ is the interaction energy between an hACE2 residue with the whole RBD domain (**Figure 8**).


SER19
THR20
ILE21
GLU22
GLU23
GLN24
ALA25
LYS26
THR27
PHE28
LEU29
ASP30
LYS31
PHE32
ASN33
HIS34
GLU35
ALA36
GLU37
ASP38
LEU39
TYR41
GLN42
LEU45
PHE72
GLU75
GLN76
THR78
LEU79
ALA80
GLN81
MET82
TYR83
ASN90
ASN322
THR324
GLN325
GLY326
PHE327
GLU329
ASN330
LEU351
GLY352
LYS353
GLY354
ASP355
PHE356
ARG357
ASP382
MET383
ALA386
ALA387
GLN388
PRO389
PHE390
ARG393
Prototype
V341I
F342L
N354D
N354D/D364Y
D364Y
V367F
R408I
A435S
W436R
G476S
V483A
N501Y
N501Y/E484K
N501Y/K417N
N501Y/E484K/K417N
−20
−16
−12
−8
−4
0

Figure 7. Heatmap of hACE2 residues.

The residue was selected if the energy contribution in RBD/hACE2 binding exceeded -0.1 kcal/mol. The y axis label presents the selected residue names; the x axis presents mutant system name. The bar on the bottom represent the relation between energy contribution and the color: darker color on the heatmap indicates bigger contribution of the residue in the binding process.

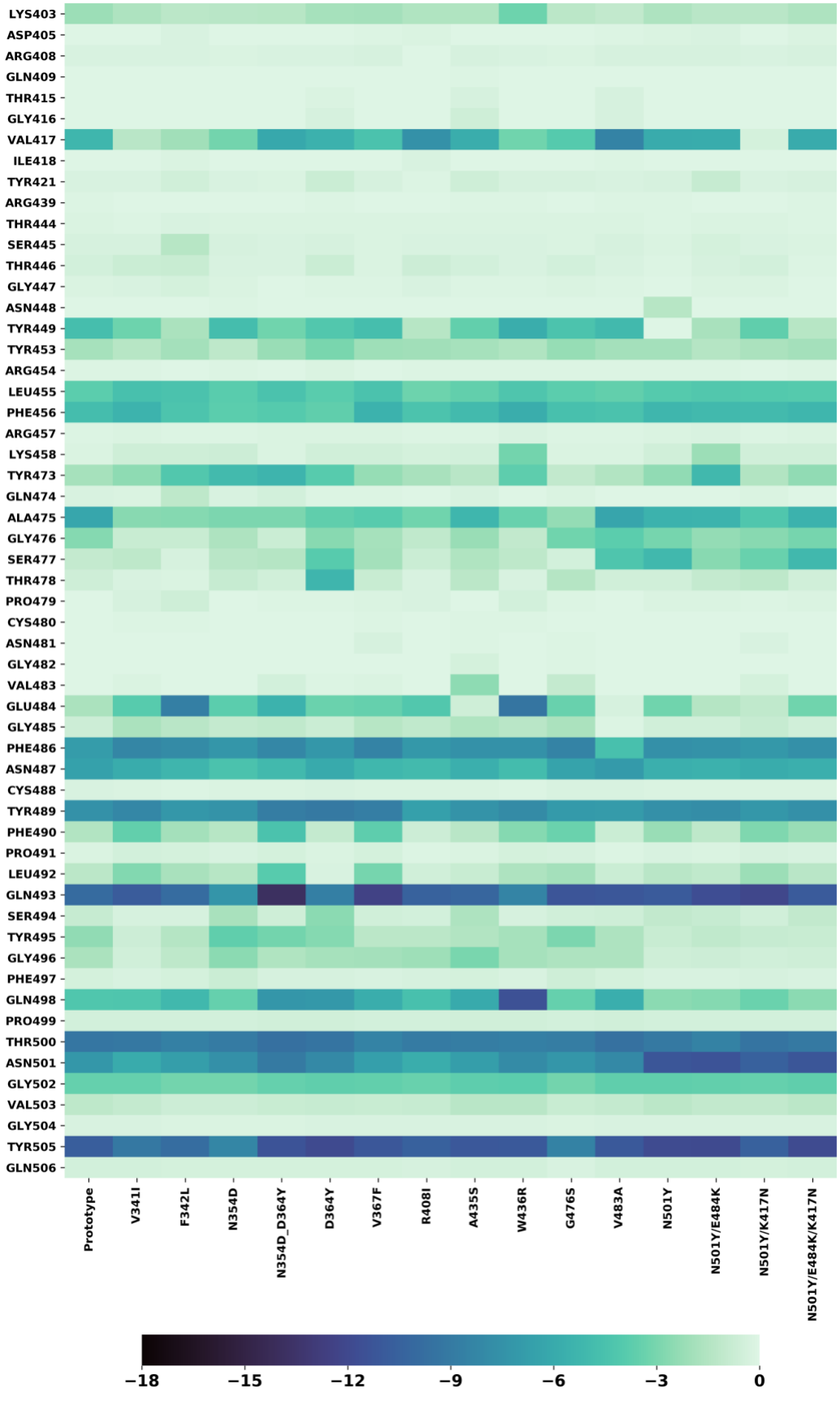

LYS403
ASP405
ARG408
GLN409
THR415
GLY416
VAL417
ILE418
TYR421
ARG439
THR444
SER445
THR446
GLY447
ASN448
TYR449
TYR453
ARG454
LEU455
PHE456
ARG457
LYS458
TYR473
GLN474
ALA475
GLY476
SER477
THR478
PRO479
CYS480
ASN481
GLY482
VAL483
GLU484
GLY485
PHE486
ASN487
CYS488
TYR489
PHE490
PRO491
LEU492
GLN493
SER494
TYR495
GLY496
PHE497
GLN498
PRO499
THR500
ASN501
GLY502
VAL503
GLY504
TYR505
GLN506
Prototype
V341I
F342L
N354D
N354D_D364Y
D364Y
V367F
R408I
A435S
W436R
G476S
V483A
N501Y
N501Y/E484K
N501Y/K417N
N501Y/E484K/K417N
−18
−15
−12
−9
−6
−3
0

Figure 8. Heatmap of RBD residues.

The residue was selected if the energy contribution in RBD/hACE2 binding exceeded -0.1 kcal/mol. The y axis label presents the selected residue names; the x axis presents mutant system names. The bar on the bottom represents the relation between energy contribution and the color: darker color on the heatmap indicates bigger contribution in the binding process.

For most mutants, the residues at the mutant positions did not take a crucial role in RBD/hACE2 binding and were not recognized as the key residues. Gly476, Val483, and Arg408 were selected and shown in the heatmap. Gly476 is making a significant contribution (-0.8 ~ -3.8 kcal/mol, **Table S4**) in protein binding compared to other mutation residues, as it is a part of the protein-protein binding interface (**Figure 1**). Unlike Gly476, the $\Delta\Delta G_{inter}$ of Arg408 is neglectable (~ -0.2 kcal/mol) and the $\Delta\Delta G_{inter}$ of Val483 is only significant for a few mutants including A435S and G476S. Asn501 is one of the most important residues in RBD, as the residues on this position all showed dark color on the heatmap. The ΔΔGinter of Asn501 varies from -5 kcal/mol ~-12 kcal/mol in different systems. Among the N501Y related mutants (N501Y, N501Y/E484K, N501Y/K417N, and N501Y/E484K/K417N), the Tyr501 was significantly heated compared to Asn501 in other mutants. Combined with the truth that the high infectivity of N501Y related mutants, the residue at 501 position can be a promising therapeutic target. We further selected some residues which are key residues in all of the systems. Drugs that target these key residues may block RBD-hACE2 binding for not only the prototype, but also the mutants, and are likely free of the drug resistance problem. Among all RBD key residues, nines have dark color in all system: Leu455, Phe456, Phe486, Asn487, Tyr489, Gln493, Thr500, Asn501, and Tyr505. Similarly, seven hACE2 residues have dark color on the heatmap indicating their essential role in RBD-hACE2 binding: Gln24, Thr27, Asp30, Lys31, His34, Lys353, and Asp355. These key

residues are shown in **Figure 9**. These residues should be considered as hotspots in rational drug and vaccine development.

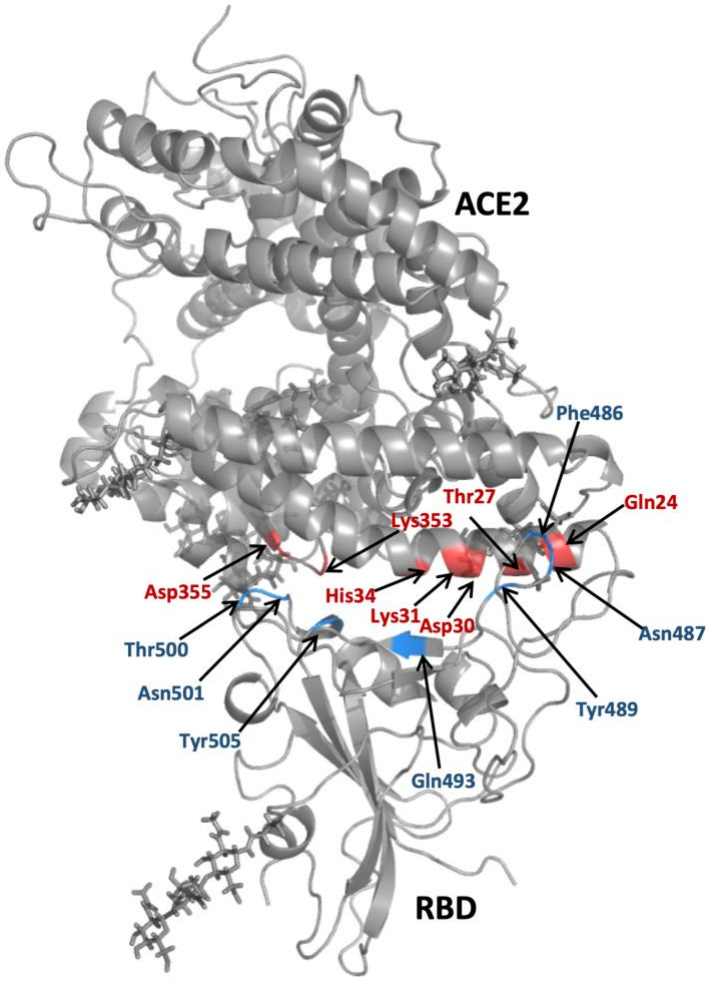


Figure 9. The selected key binding residues.

The red colored residues are the key binding residues selected from hACE2. The blue colored residues are the key binding residues selected from RBD.

We also designed binding pockets for both RBD and hACE2 based on the key binding residues positions, and the results are shown in the **Figure 10**. We designed four binding pockets

for RBD, and two binding pockets for hACE2. These binding pockets can be used to screen potential RBD/hACE2 binding inhibitors.

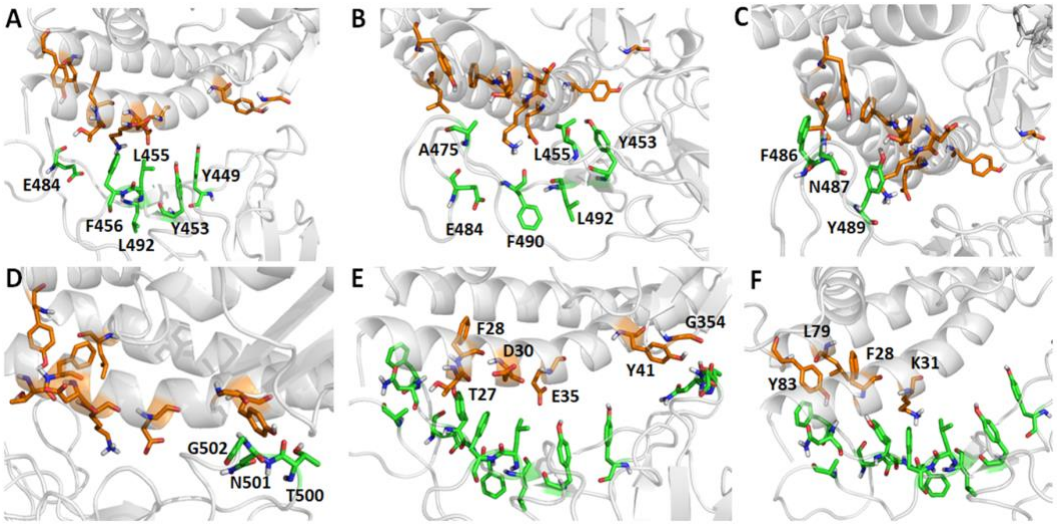


Figure 10. The cluster of binding sites of RBD/hACE2 complex.

A: Cluster 1 for RBD; B: Cluster 2 for RBD; C: cluster 3 for RBD; D: Cluster 4 for RBD; E: Cluster 1 for hACE2; F: Cluster 2 for hACE2.

3.2 Drug Screening for Spike Protein RBD

In this section, we used molecular docking to find potential therapeutic small molecules for RBD. First, we performed docking study, then the docking scores were compared with experimental results to find out potential molecules. The comparison of docking scores and experimental data helps to understand the underlying mechanism of inhibitory efficacy of the small molecules.

3.2.1 Molecular docking using NPC drug library

Drug screening was conducted using one binding pocket designed for RBD (Figure 10). Cluster A was selected to do the screening study to select RBD/hACE2 complex binding inhibitors. The docking study was carried out using Glide docking. Ligands collected from the NPC library were prepared to write out all the possible conformations. In the docking study, at most two docking scores were output for each conformation, and the conformation with the best docking score was kept for each ligand. The best docking score of each ligand was also compared to the AlphaLisa bioassay test result[70] and shown in Table 3. The AlphaLisa aimed to find small molecules that can interrupt the RBD/hACE2 binding, which is also the objective of our screening study. To illustrate the essence of the docking study, Table 3 only presents the top 100 molecules (sorted by docking score). The rest part of docking/experimental results was deposited in the attached file. The comparison between docking score and bioassay test helps us to find out the lying mechanism of the potential drug candidate. In bioassay study, the molecules were only screened for efficacy of interrupting binding of Spike/hACE2, whereas the mechanism remains unknown. Consistency of good docking score and the efficacy in bioassay test indicates the therapeutic potential of the small molecule, and also suggest that the binding pocket used in docking study maybe the therapeutic target of the molecule.

Table 3. Docking resutls and the bioassay results for screened compounds.

No.	Drug Name	Docking	AC50 (uM)	No.	Drug Name	Docking	AC50 (uM)
1	NCGC00167473-02	-9.1	NA	**51**	NCGC00016288-06	-5.9	NA
2	NCGC00390238-01	-7.3	NA	**52**	NCGC00532495-01	-5.9	NA
3	NCGC00015693-13	-7.3	0.40	**53**	NCGC00178638-13	-5.9	NA
4	NCGC00179611-05	-7.1	NA	**54**	NCGC00274067-01	-5.9	2.67
5	NCGC00179612-03	-6.9	NA	**55**	NCGC00168746-02	-5.8	12.59

6	NCGC00178839-08	-6.9	NA	**56**	NCGC00015692-08	-5.8	NA
7	NCGC00532509-01	-6.9	NA	**57**	NCGC00159460-01	-5.8	NA
8	NCGC00159339-10	-6.8	NA	**58**	NCGC00262929-01	-5.8	11.92
9	NCGC00165736-04	-6.7	NA	**59**	NCGC00482790-01	-5.8	NA
10	NCGC00389393-01	-6.7	17.90	**60**	NCGC00482904-01	-5.8	NA
11	NCGC00319019-01	-6.7	15.00	**61**	NCGC00387997-02	-5.8	NA
12	NCGC00485883-01	-6.6	26.68	**62**	NCGC00159462-04	-5.8	26.68
13	NCGC00179506-05	-6.6	NA	**63**	NCGC00163572-01	-5.8	3.57
14	NCGC00166210-02	-6.6	NA	**64**	NCGC00015760-12	-5.8	NA
15	NCGC00093350-05	-6.6	NA	**65**	NCGC00163355-05	-5.8	NA
16	NCGC00091469-04	-6.5	10.62	**66**	NCGC00015094-22	-5.8	NA
17	NCGC00095055-08	-6.4	NA	**67**	NCGC00016306-08	-5.7	NA
18	NCGC00356145-10	-6.4	22.54	**68**	NCGC00164560-17	-5.7	NA
19	NCGC00250404-15	-6.4	NA	**69**	NCGC00164560-05	-5.7	NA
20	NCGC00178734-13	-6.4	NA	**70**	NCGC00386239-01	-5.7	21.19
21	NCGC00262961-01	-6.4	21.19	**71**	NCGC00015074-17	-5.7	NA
22	NCGC00483033-01	-6.4	NA	**72**	NCGC00179501-03	-5.7	11.92
23	NCGC00182027-01	-6.3	10.62	**73**	NCGC00386335-01	-5.7	7.13
24	NCGC00181135-02	-6.3	NA	**74**	NCGC00167429-04	-5.7	NA
25	NCGC00178734-06	-6.3	NA	**75**	NCGC00015034-21	-5.7	NA
26	NCGC00016403-11	-6.3	NA	**76**	NCGC00091414-07	-5.7	NA
27	NCGC00379057-02	-6.3	7.13	**77**	NCGC00179367-03	-5.7	NA
28	NCGC00178852-05	-6.2	NA	**78**	NCGC00095143-07	-5.7	NA
29	NCGC00483012-01	-6.2	NA	**79**	NCGC00095143-05	-5.7	NA
30	NCGC00167982-04	-6.2	5.32	**80**	NCGC00016285-05	-5.7	NA
31	NCGC00186028-02	-6.2	NA	**81**	NCGC00183283-01	-5.7	NA
32	NCGC00091469-15	-6.2	8.00	**82**	NCGC00378687-03	-5.7	22.54
33	NCGC00389765-01	-6.2	NA	**83**	NCGC00015819-15	-5.7	NA
34	NCGC00183867-06	-6.1	7.52	**84**	NCGC00263218-07	-5.7	NA
35	NCGC00183867-02	-6.1	11.30	**85**	NCGC00179135-02	-5.7	2.01
36	NCGC00379018-02	-6.1	NA	**86**	NCGC00183095-01	-5.7	29.93
37	NCGC00015821-10	-6.1	NA	**87**	NCGC00345451-03	-5.7	NA

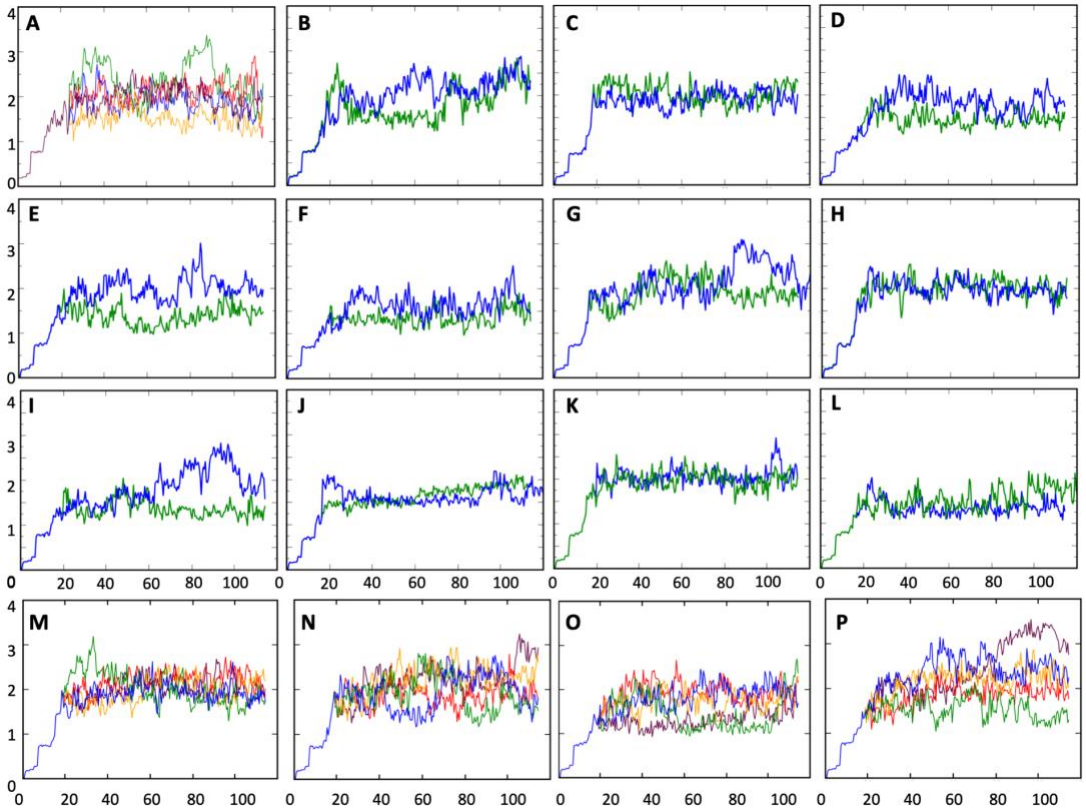
38	NCGC00483046-01	-6.1	3.36	**88**	NCGC00016480-10	-5.7	2.01
39	NCGC00094087-05	-6.1	NA	**89**	NCGC00167480-02	-5.7	NA
40	NCGC00485882-01	-6.1	NA	**90**	NCGC00379005-02	-5.7	NA
41	NCGC00179313-05	-6.1	NA	**91**	NCGC00480787-02	-5.7	10.62
42	NCGC00188433-02	-6.1	NA	**92**	NCGC00379085-03	-5.7	0.80
43	NCGC00182070-03	-6.0	NA	**93**	NCGC00025349-05	-5.6	16.83
44	NCGC00183095-07	-6.0	5.66	**94**	NCGC00386194-01	-5.6	4.74
45	NCGC00386253-01	-6.0	NA	**95**	NCGC00015513-11	-5.6	NA
46	NCGC00018167-05	-6.0	NA	**96**	NCGC00016401-09	-5.6	NA
47	NCGC00390678-03	-6.0	NA	**97**	NCGC00093846-07	-5.6	NA
48	NCGC00178526-03	-6.0	NA	**98**	NCGC00389453-01	-5.6	13.37
49	NCGC00166322-02	-5.9	NA	**99**	NCGC00160628-18	-5.6	NA
50	NCGC00482848-01	-5.9	NA	**100**	NCGC00263918-08	-5.6	23.78

4.0 Conclusion

This study applied a series of computational methods to explore the binding pattern of SARS-CoV-2 RBD with the hACE2, and to find promising small molecules. Molecular mechanics models of SARS-CoV-2 RBD/hACE2 prototype and mutants were built after a set of force field parameterizations of modified resides. Through MD Simulations and MM-PBSA calculations, the binding free energies of the prototype and mutant type RBDs binding to hACE2 were predicted. Among the first-wave Spike protein mutants, V367F and N354D/D364Y were predicted to have significantly higher binding affinities, and the prediction was validated by experiment and the virus epidemiology. The other four N501Y related mutants, which are also known as second-wave mutants, showed enhanced binding affinity according to the calculation results. The enhanced binding affinity of N501Y series is consistent with the worldwide prevalence of these mutants. We came up with one theory from this observation: only the properly enhanced binding affinity benefits the viral infectivity. If the binding affinity is too strong, like the D364Y mutant, it may eventually hinder the cell entry. Moreover, the complex free energy was applied to predict structural stability altered by mutagenesis. Among the first-wave mutants, R408I and W436R were found to significantly decrease and enhance the structural stability, respectively. All of the N501Y series mutants have shown decreased structural stability. For this observation, we came up with one theory that strong complex stability may impair viral infectivity, as it'll be hard for cell entry after the spike/hACE2 binding. Last, we conducted MM-GBSA free energy decomposition analysis using the snapshots collected during MD simulations. A set of hotspot residues were identified for both the prototype and mutants. After cluster analysis, we identified a set of potential binding sites where the key residues from both the essential and expanded sets are located. Those

binding sites may be applied to develop inhibitors of the RBD/hACE2 binding through virtual screenings. The identified key residues can also provide guidance on vaccine development for the spike protein. This research work also performed a docking study using compounds from NPC drug library, and the docking scores were compared to experimental results. In sum, this study provides insights into binding profile to help understand the RBD/hACE2 binding mechanism and mutations and also gave guidance in SARS-CoV-2 vaccine development or inhibitor design.

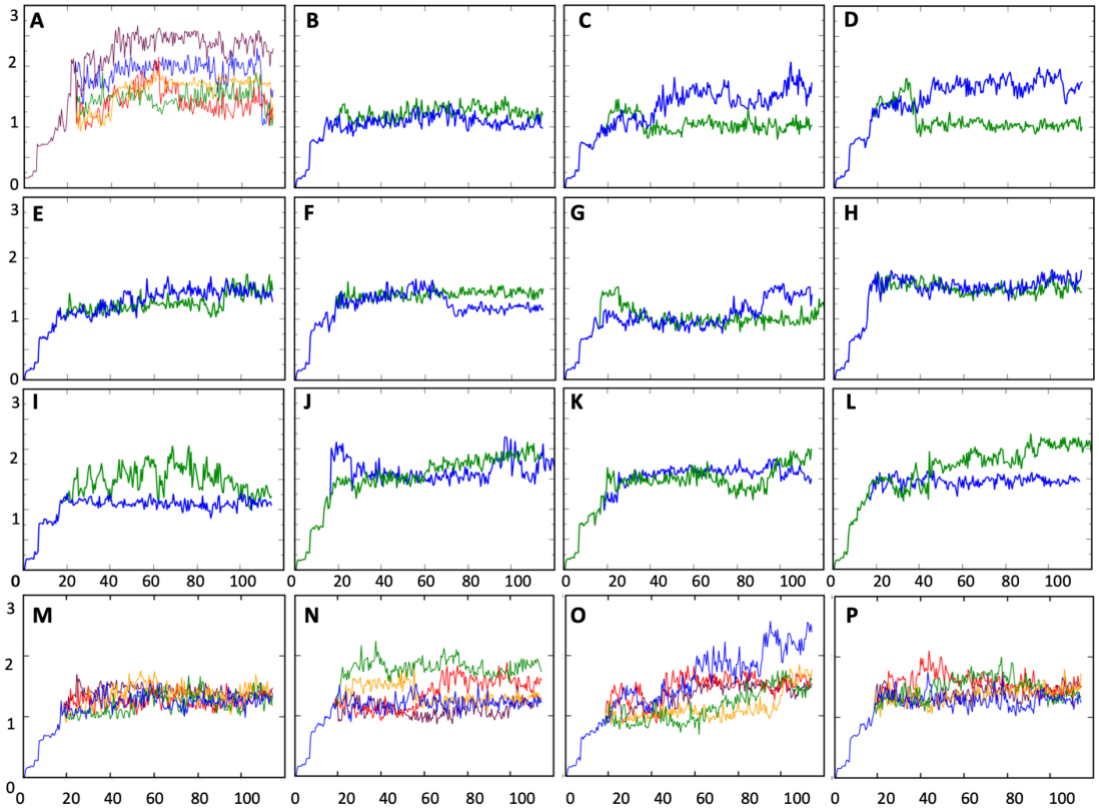
Appendix A Figures



Appendix Figure 1. RMSD~Time plots for hACE2s

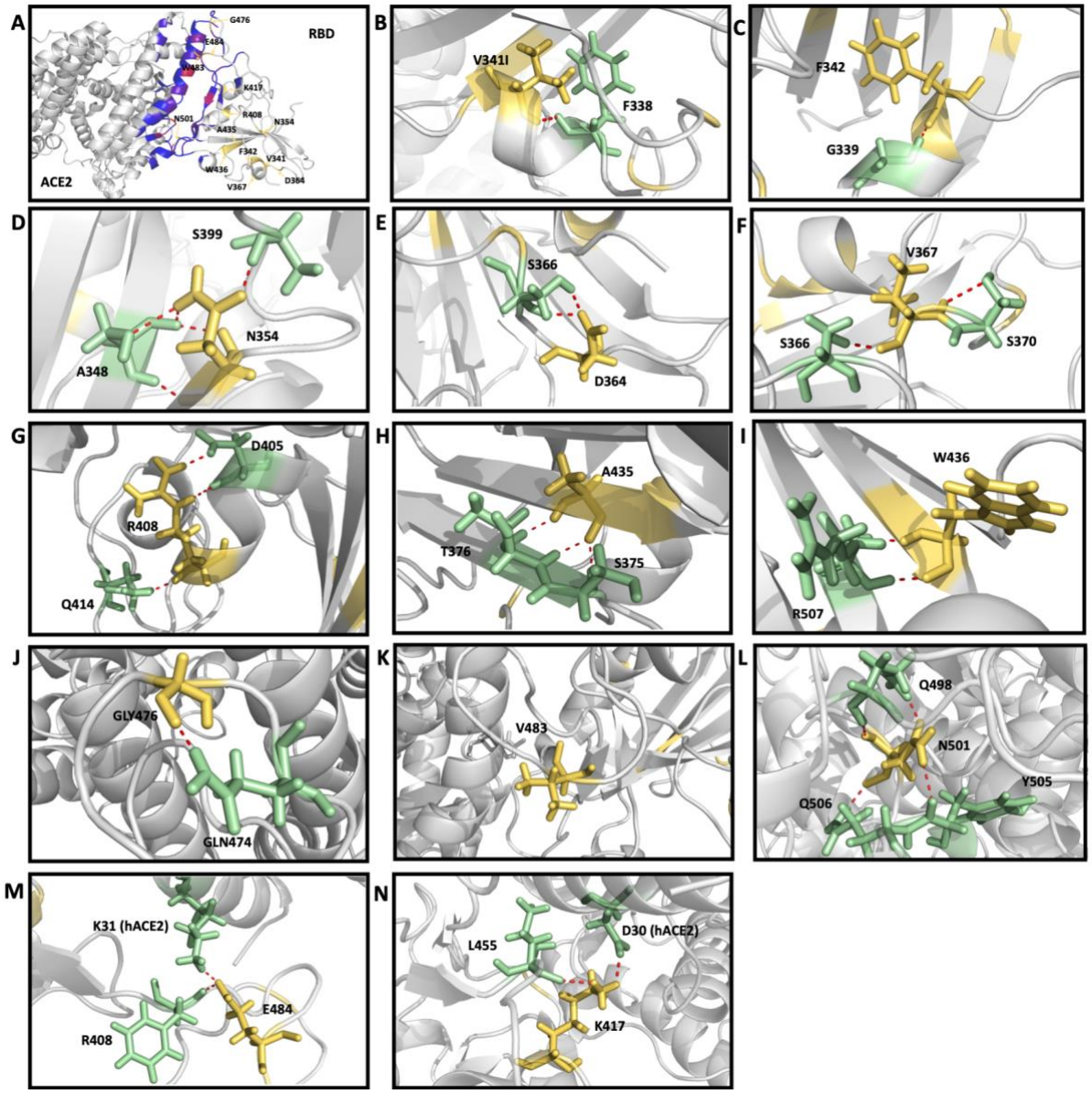
The y axis stands for RMSD (Å), the x axis stands for simulation time (ns). The different colors represent the different simulation runs. Panel A~L are for ACE2 from different system. A: prototype hACE2; B: V341I hACE2; C: F342L hACE2; D: N354D hACE2; E: N354D/D364Y hACE2; F: D364Y hACE2; G: V367F hACE2; H: R408I hACE2; I: A435S hACE2; J: W436R hACE2; K: G476S hACE2; L: V483A hACE2; M: N501Y hACE2; N: N501Y/E484K hACE2; O: N501Y/K417N hACE2; P: N501Y/E484K/K417N hACE2.

To create alternate Appendix Figure/Table caption labels, begin by inserting a caption as you would for any other figure or table. Then use the new label button to specify a new label for the caption. You will need to either create a separate list of figures/tables for these labels or use the ETD formatting guide for creating custom Appendix labels.



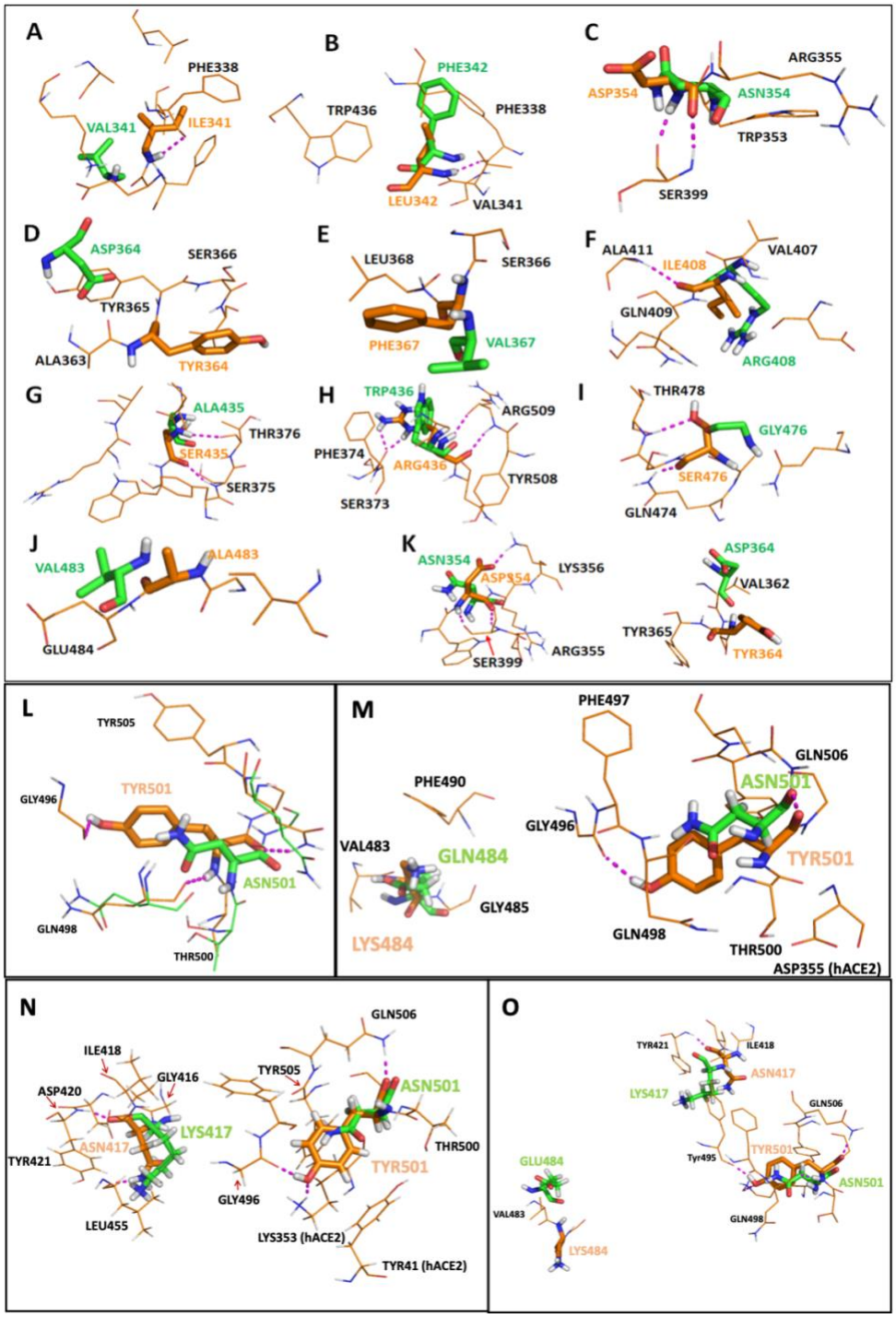
Appendix Figure 2. RMSD~Time plots for RBDs.

The y axis stands for RMSD (Å), the x axis stands for simulation time (ns). The different colors represent different simulation runs. Panel A~L are for RBDs from different system. A: prototype RBD; B: V341I RBD; C: F342L RBD; D: N354D RBD; E: N354D/D364Y RBD; F: D364Y RBD; G: V367F RBD; H: R408I RBD; I: A435S RBD; J: W436R RBD; K: G476S RBD; L: V483A RBD; M: N501Y RBD; N: N501Y/E484K RBD; O: N501Y/K417N RBD; P: N501Y/E484K/K417N RBD.



Appendix Figure 3. Key interaction of mutated residues revealed by wild type crystal structure.

The overall mutation sites of RBD (colored in yellow) and the hotspot residues of ACE2 (colored in blue and red using the scheme of Figure 5) is shown in Panel A. The details of interactions at each mutation site is shown in the rest of the panels with the mutation residue colored in yellow, the interacting residues in green and dashed lines indicating hydrogen-bonds. B: Val341, C: Phe342, D: Asn354, E: Asp364, F: Val367, G: Arg408, H: Ala435, I: Trp436, J: Gly476, K: Val483, L: Asn501, M; Glu484, N: Lys417.



Appendix Figure 4. Key interactions of the mutated residues revealed by MD simulations.

For each panel, the mutation residue is shown as brownish sticks and labeled with brown text, the surrounding key residues are colored as brownish lines and labeled with black text. The original residue is shown as greenish sticks and labeled with green text. Hydrogen bonds are shown as magenta dashed lines. A: V341I; B: F342L;

C: N354D; D: D364Y; E: V367F; F: R408I; G: A435S; H: W436R; I: G476S; J: V483A; K: N354D/D364Y; L: N501Y; M: N501Y/E484K; N: N501Y/K417N; O: N501Y/E484K/K417N.

Appendix B Tables

Appendix Table 1. The experimental values of prototype and three mutant type SARS-CoV-2 with ACE2. The conversion between experimental values and the binding free energy uses the equation: $\Delta G = RTlnK_d$. BLI refers to Bio-Layer Interferometry; SPR refers to Surface plasmon resonance; ELISA refers to enzyme-linked immunosorbent assay.

	Experimental Methods	Experimental Kd/EC_{50} value (nM)	Experimental binding free energy (kcal/mol)	Predicted binding free energy (kcal/mol)
Wild Type SARS-CoV-2	BLI (Kd)	24.4[67]	-10.38	-16.04
	SPR (Kd)	13.1[67]	-10.74	
	ELISA (EC_{50})	1.47[67]	-12.04	
	ELISA* (EC_{50})	1.2[13]	-12.16	
V367F	BLI (Kd)	5.5[67]	-11.26	-20.66
	SPR (Kd)	4.33[67]	-11.40	
	ELISA (EC_{50})	0.31[67]	-12.96	
W436R	BLI (Kd)	6.85[67]	-11.13	-17.12
	SPR (Kd)	6.96[67]	-11.12	
	ELISA (EC_{50})	0.89[67]	-12.34	
N354D/D364Y	BLI (Kd)	6.33[67]	-11.18	-20.84
	SPR (Kd)	7.57[67]	-11.07	
	ELISA (EC_{50})	0.97[67]	-12.29	

Appendix Table 2. The residues selected from the prototype RBD/ACE2 system using $\Delta\Delta G_{int}$ < -0.1. The unit of $\Delta\Delta G_{inter}$ is kcal/mol. These residues are colored in the hotspot in red or blue represent binding intensity.

Residue ID	$\Delta\Delta G_{inter}$	Residue ID	$\Delta\Delta G_{inter}$	Residue ID	$\Delta\Delta G_{inter}$	Residue ID	$\Delta\Delta G_{inter}$
Human ACE2							
SER19	-5.44	LEU29	-0.17	TYR41	-2.81	LEU351	-0.23
THR20	-0.12	ASP30	-7.33	GLN42	-0.84	GLY352	-0.73
ILE21	-0.24	LYS31	-12.35	LEU45	-0.57	LYS353	-15.42
GLU22	-0.13	PHE32	-0.24	GLN76	-0.23	GLY354	-2.61
GLU23	-1.26	ASN33	-0.20	LEU79	-2.65	ASP355	-9.38
GLN24	-7.01	HIS34	-7.10	ALA80	-0.12	PHE356	-0.48
ALA25	-0.16	GLU35	-4.32	MET82	-2.44	ALA386	-0.70
LYS26	-0.19	GLU37	-4.21	TYR83	-5.46	ALA387	-0.34
THR27	-6.60	ASP38	-5.00	ASN90	-0.24	PHE390	-0.12

PHE28	-2.86	LEU39	-0.13	ASN322	-0.16	ARG393	-0.86
Spike RBD Domain							
LYS403	-2.14	LEU455	-3.77	GLY485	-0.65	GLY496	-1.73
ARG408	-0.28	PHE456	-4.79	PHE486	-6.86	PHE497	-0.42
VAL417	-5.24	ARG457	-0.11	ASN487	-6.55	GLN498	-4.21
TYR421	-0.25	LYS458	-0.15	CYS488	-0.22	PRO499	-0.49
THR444	-0.16	TYR473	-1.89	TYR489	-7.55	THR500	-9.23
SER445	-0.33	GLN474	-0.22	PHE490	-1.51	ASN501	-7.18
THR446	-0.42	ALA475	-6.20	PRO491	-0.20	GLY502	-3.51
GLY447	-0.15	GLY476	-2.70	LEU492	-1.22	VAL503	-1.15
TYR449	-4.77	SER477	-0.97	GLN493	-9.98	GLY504	-0.23
TYR453	-1.90	THR478	-0.58	SER494	-0.88	TYR505	-10.75
ARG454	-0.11	GLU484	-1.73	TYR495	-2.42	GLN506	-0.47

Appendix Table 3. The residues in ACE2 selected form each system.

The residues are selected if the binding free energy contribution is lower than -0.1 kcal/mol. If a residue is not selected in the corresponding system, "0" was entered in the corresponding table cell.

	WT	V341I	F342L	N354D	D364Y	N354 D/ D364Y	V367F	R408I	A435S	W436R	G476S	V483A	N501Y	N501Y/E 484K	N501Y/ K417N	N501Y/E 484K/K4 17N
SER19	-5.44	-1.35	-2.80	-1.06	-0.19	-0.73	-1.07	-0.24	-4.18	-0.71	-1.41	-1.70	-1.86	-0.99	-1.71	-1.86
THR20	-0.12	-0.23	-0.24	-1.43	-0.31	-0.15	-1.68	0	-0.30	-0.18	-0.20	-3.34	-3.67	-0.12	-2.72	-3.67
ILE21	-0.24	-0.23	-0.15	-0.26	-1.92	-0.67	-0.31	-0.32	-0.64	-0.50	-0.78	-0.47	-0.45	-1.29	-0.43	-0.45
GLU22	-0.13	-0.13	-0.12	-0.15	-8.46	-0.16	-0.14	-0.14	-0.14	-0.14	-0.13	-0.20	-0.17	-0.23	-0.10	-0.17
GLU23	-1.26	-1.97	-1.31	-1.83	-2.86	-3.72	-1.57	-2.29	-1.52	-6.02	-1.58	-2.25	-2.88	-8.04	-1.57	-2.88
GLN24	-7.01	-6.90	-6.77	-8.07	-8.26	-8.94	-8.55	-6.77	-8.41	-7.41	-6.73	-10.17	-9.30	-10.07	-7.92	-9.30
ALA25	-0.16	-0.19	-0.13	-0.16	-0.19	-0.17	-0.24	-0.13	-0.16	-0.19	-0.15	-0.14	-0.17	-0.21	-0.18	-0.17
LYS26	-0.19	-0.41	-0.63	-0.50	-0.44	0	-0.14	-0.33	-0.29	-0.34	-0.17	0	-0.11	-0.75	-0.55	-0.11
THR27	-6.60	-7.67	-6.99	-7.02	-6.66	-5.75	-7.23	-5.59	-6.26	-7.77	-5.83	-6.40	-6.26	-7.06	-6.32	-6.26
PHE28	-2.86	-2.67	-2.39	-2.76	-2.46	-3.34	-3.17	-2.74	-2.69	-2.63	-2.88	-2.60	-3.22	-3.18	-2.51	-3.22
LEU29	-0.17	-0.23	-0.20	-0.19	-0.18	-0.16	-0.20	-0.12	-0.18	-0.21	-0.18	-0.13	-0.17	-0.18	-0.21	-0.17
ASP30	-7.33	-3.04	-3.41	-6.06	-6.57	-7.93	-6.64	-8.23	-6.57	-5.34	-5.37	-8.74	-8.45	-7.89	-2.32	-8.45
LYS31	-12.35	-18.79	-21.80	-13.22	-9.12	-22.50	-20.12	-13.45	-10.89	-20.06	-18.09	-10.49	-15.43	-10.86	-15.83	-15.43
PHE32	-0.24	-0.27	-0.26	-0.29	-0.25	-0.30	-0.25	-0.22	-0.25	-0.28	-0.23	-0.21	-0.24	-0.24	-0.23	-0.24
ASN33	-0.20	-0.22	-0.18	-0.18	-0.23	-0.21	-0.22	-0.15	-0.22	-0.20	-0.18	-0.17	-0.22	-0.19	-0.24	-0.22
HIS34	-7.10	-6.18	-6.19	-6.47	-11.21	-6.84	-7.52	-6.84	-8.72	-6.22	-6.68	-7.20	-8.01	-7.32	-6.52	-8.01
GLU35	-4.32	-4.17	-3.42	-2.19	-4.42	-4.71	-3.96	-4.00	-4.62	-4.31	-4.26	-4.19	-3.61	-4.36	-4.29	-3.61
ALA36	0	0	0	0	0	-0.10	0	0	0	-0.11	0	0	0.00	0.00	0.00	0.00
GLU37	-4.21	-3.28	-1.74	-1.60	-5.43	-4.60	-4.47	-3.09	-4.84	-6.00	-1.83	-3.64	-5.57	-4.88	-3.97	-5.57
ASP38	-5.00	-3.53	-2.15	-4.45	-3.49	-4.99	-4.98	-2.03	-3.80	-6.61	-3.82	-4.82	-2.01	-2.74	-3.97	-2.01
LEU39	-0.13	-0.13	-0.18	-0.14	-0.13	-0.17	-0.15	-0.11	-0.12	-0.20	-0.15	-0.12	0.00	-0.10	-0.14	0.00

TYR41	-2.81	-3.04	-4.84	-3.57	-3.64	-3.72	-2.31	-2.13	-3.04	-3.12	-3.36	-4.25	-3.53	-4.10	-4.33	-3.53
GLN42	-0.84	-0.60	-2.36	-0.94	-2.95	-0.86	-0.40	-0.82	-1.01	-3.20	-1.67	-2.13	-0.47	-1.06	-1.74	-0.47
LEU45	-0.57	-0.49	-0.77	-0.59	-0.95	-0.80	-0.58	-0.41	-0.66	-0.80	-0.67	-0.92	-0.50	-0.75	-1.06	-0.50
PHE72	0	-0.13	-0.11	-0.13	0	-0.45	-0.12	-0.11	0	-0.12	0	0	0.00	0.00	-0.12	0.00
GLU75	0	-0.13	-0.15	-0.15	-0.20	-0.24	-0.24	0	-0.53	-0.18	-0.25	0	-0.13	-1.32	-0.25	-0.13
GLN76	-0.23	-0.52	-0.43	-0.69	-0.77	-1.97	-0.40	-0.34	-0.40	-0.38	-0.53	-0.14	-0.24	-0.25	-0.47	-0.24
THR78	0	0	0	0	0	0	0	0	-0.48	0	-0.14	0	0.00	0.00	-0.12	0.00
LEU79	-2.65	-3.40	-3.32	-3.08	-2.76	-3.30	-3.71	-2.90	-4.11	-3.02	-3.59	-1.54	-3.10	-2.90	-3.07	-3.10
ALA80	-0.12	-0.12	-0.11	0	0	-0.13	-0.15	-0.10	-0.11	-0.11	-0.15	0	-0.12	-0.10	-0.13	-0.12
GLN 81	0	0	0	0	0	0	-0.10	0	0	0	0	0	0.00	0.00	0.00	0.00
MET82	-2.44	-3.17	-2.81	-2.21	-2.67	-3.01	-3.18	-2.65	-3.49	-3.14	-3.16	-2.04	-2.41	-2.18	-2.86	-2.41
TYR83	-5.46	-4.76	-4.17	-2.99	-6.88	-5.89	-6.03	-3.80	-4.28	-4.13	-6.02	-4.47	-5.64	-4.61	-5.16	-5.64
ASN90	-0.24	-1.06	-0.63	-0.24	-2.34	-0.19	-0.38	-0.51	-2.86	-0.18	-0.22	-2.88	-0.36	-0.61	-0.15	-0.36
ASN322	-0.16	-0.34	0	0	0	-0.11	0.00	0	0	-0.18	0	0	0.00	0.00	-0.28	0.00
THR324	-0.89	-0.69	-0.35	-0.37	-0.38	-0.53	-0.56	-0.78	-1.32	-1.43	-0.53	-0.46	-0.70	-0.43	-0.53	-0.70
GLN325	-0.53	-0.53	-0.43	-0.51	-0.62	-0.23	-0.22	-0.19	-0.48	-0.30	-0.37	-0.75	-0.84	-0.67	-0.39	-0.84
GLY326	-0.35	-0.34	-0.27	-0.31	-0.35	-0.30	-0.27	-0.29	-0.37	-0.22	-0.32	-0.35	-0.42	-0.32	-0.32	-0.42
PHE327	-0.10	0	0	0	0	-0.10	-0.10	0	-0.10	0	0	0	0.00	-0.20	0.00	0.00
GLU329	-0.19	-0.20	-0.17	-0.19	-0.24	-0.19	-0.19	-0.18	-0.18	-0.16	-0.19	-0.20	-0.22	-1.72	-0.21	-0.22
ASN330	-1.88	-1.85	-1.44	-1.65	-2.27	-1.88	-1.63	-1.61	-1.74	-1.54	-1.75	-2.11	-2.09	-0.11	-2.16	-2.09
LEU351	-0.23	-0.27	-0.27	-0.22	-0.22	-0.22	-0.20	-0.21	-0.23	-0.20	-0.20	-0.24	-0.25	-0.23	-0.26	-0.25
GLY352	-0.73	-0.59	-0.79	-1.02	-0.62	-0.75	-0.59	-0.62	-0.58	-0.59	-0.82	-0.70	-0.63	-0.63	-0.67	-0.63
LYS353	-15.42	-12.26	-13.45	-17.47	-18.48	-18.61	-16.78	-13.98	-16.37	-21.21	-16.10	-15.25	-13.79	-14.65	-14.24	-13.79
GLY354	-2.61	-2.35	-2.49	-2.65	-2.61	-2.65	-2.74	-2.68	-2.65	-2.50	-2.65	-2.64	-2.48	-2.53	-2.51	-2.48
ASP355	-9.38	-9.32	-7.73	-8.11	-7.62	-9.52	-8.15	-9.78	-8.28	-7.47	-7.72	-8.44	-8.65	-7.65	-7.94	-8.65
PHE356	-0.48	-0.37	-0.43	-0.43	-0.40	-0.48	-0.53	-0.53	-0.43	-0.44	-0.41	-0.43	-0.46	-0.44	-0.37	-0.46
ARG357	0	0	-0.23	-0.19	-0.34	-0.10	-0.12	0	-0.25	-0.39	-0.16	-0.19	-0.23	-0.25	-0.31	-0.23
ASP382	0	0	0	0	0	0	0	0	0	0	0	0	0.00	-0.10	0.00	0.00
MET383	-0.11	0	0	0	0	0	-0.13	-0.12	-0.13	-0.10	-0.11	0	0.00	0.00	0.00	0.00
ALA386	-0.70	-0.32	-1.38	-0.90	-0.18	-0.53	-0.66	-1.27	-0.37	-0.46	-0.86	-0.57	-0.57	-0.61	-0.34	-0.57
ALA387	-0.34	-0.29	-0.91	-0.38	-0.16	-0.31	-0.39	-0.51	-0.21	-0.30	-0.35	-0.37	-0.47	-0.55	-0.22	-0.47
GLN388	0	0	-0.11	0	0	0	0	-0.11	0	0	0	0	0.00	0.00	0.00	0.00
PRO389	0	0	-0.21	0	0	0	-0.12	-0.15	0	-0.11	0	0	0.00	-0.11	0.00	0.00
PHE390	-0.12	0	-0.10	0	-0.10	-0.12	-0.13	-0.11	-0.13	-0.12	0	-0.10	-0.12	-0.11	0.00	-0.12
ARG393	-0.86	-0.42	-0.93	-0.33	-0.44	-0.66	-0.95	-1.05	-0.57	-0.76	-0.32	-0.83	-0.86	-0.74	-0.59	-0.86

Appendix Table 4. The residues in RBD selected form each system.

The residues are selected if the binding energy contribution is lower than -0.1 kcal/mol. If a residue is not selected in the corresponding system, "0" was entered in the corresponding table cell.

	WT	V341I	F342L	N354D	D364Y	N354D/D364Y	V367F	R408I	A435S	W436R	G476S	V483A	N501Y	N501Y/E484K	N501Y/K417N	N501Y/E484K/K417N
LYS403	-2.14	-1.62	-1.25	-1.30	-1.88	-1.37	-1.91	-1.55	-1.62	-3.28	-1.21	-1.03	-1.66	-1.31	-1.32	-1.66
ASP405	0	0	-0.22	0	0	0	-0.12	-0.19	0	0	0	-0.10	-0.20	-0.22	0.00	-0.20
ARG408	-0.28	-0.28	-0.26	-0.19	-0.22	-0.22	-0.31	0	-0.32	-0.24	-0.16	-0.30	-0.33	-0.34	-0.22	-0.33
GLN409	0	0	0	0	0	0	0	0	-0.12	0	0	0	0.00	0.00	0.00	0.00
THR415	0	0	0	0	-0.17	0	0	0	-0.35	0	0	-0.35	0.00	0.00	0.00	0.00
GLY416	0	0	0	0	-0.34	0	0	0	-0.62	0	0	-0.34	0.00	0.00	0.00	0.00
VAL417	-5.24	-1.33	-1.99	-3.11	-5.60	-6.15	-4.46	-7.53	-5.80	-3.18	-3.97	-8.44	-5.96	-5.90	-0.38	-5.96
ILE418	0	0	-0.16	0	0	0	0	-0.27	0	0	0	-0.13	0.00	0.00	0.00	0.00
TYR421	-0.25	-0.24	-0.50	-0.21	-0.73	-0.17	-0.32	-0.14	-0.59	-0.33	-0.30	-0.27	-0.29	-0.86	-0.23	-0.29
ARG439	-0.10	0	0	0	-0.11	0	0	0	-0.11	-0.10	0	-0.10	-0.11	0.00	0.00	-0.11
THR444	-0.16	-0.14	-0.22	-0.17	-0.18	-0.15	-0.18	-0.18	-0.18	-0.21	-0.17	-0.17	-0.14	-0.15	-0.15	-0.14
SER445	-0.33	-0.33	-1.39	-0.29	-0.30	-0.22	-0.18	-0.23	-0.24	-0.25	-0.16	-0.30	-0.15	-0.38	-0.24	-0.15
THR446	-0.42	-0.75	-0.83	-0.25	-0.74	-0.23	-0.21	-0.65	-0.47	-0.27	-0.43	-0.27	-0.12	-0.29	-0.47	-0.12
GLY447	-0.15	-0.23	-0.36	-0.15	-0.11	0	-0.11	-0.22	-0.13	-0.12	-0.15	-0.13	0.00	-0.16	-0.18	0.00
ASN448	0	0	0	-0.10	0	0	0	0	0	0	0	0	-1.40	0.00	0.00	0.00
TYR449	-4.77	-3.27	-1.69	-4.81	-4.15	-3.23	-4.71	-1.36	-3.56	-5.79	-4.36	-5.04	0.00	-1.77	-3.60	-1.40
TYR453	-1.90	-1.40	-1.95	-1.14	-2.97	-2.21	-2.09	-2.03	-1.86	-1.55	-2.28	-1.92	-1.96	-1.46	-1.74	-1.96
ARG454	-0.11	-0.12	0	-0.12	-0.16	0	0	-0.11	-0.12	-0.12	-0.10	-0.11	-0.12	-0.12	0.00	-0.12
LEU455	-3.77	-4.60	-4.44	-3.85	-3.80	-4.45	-4.43	-3.30	-3.53	-4.28	-3.78	-3.55	-3.95	-4.13	-4.03	-3.95
PHE456	-4.79	-5.48	-4.30	-3.77	-3.60	-3.99	-5.54	-4.40	-4.96	-5.81	-4.63	-4.45	-5.21	-5.02	-4.95	-5.21
ARG457	-0.11	-0.18	-0.12	-0.13	-0.17	-0.14	-0.17	-0.13	-0.14	-0.22	-0.12	-0.11	-0.16	-0.23	-0.12	-0.16
LYS458	-0.15	-0.62	-0.58	-0.70	-0.54	-0.19	-0.56	-0.42	-0.46	-3.12	-0.17	-0.16	-0.52	-2.12	-0.53	-0.52
TYR473	-1.89	-2.51	-4.09	-4.99	-3.90	-5.35	-2.28	-1.77	-1.32	-3.71	-1.04	-1.51	-2.41	-5.07	-1.48	-2.41
GLN474	-0.22	-0.14	-1.15	-0.25	-0.14	-0.46	-0.13	0	-0.24	0	-0.14	0	0.00	0.00	-0.14	0.00
ALA475	-6.20	-2.61	-2.73	-2.93	-3.64	-2.95	-3.88	-3.20	-5.25	-3.42	-2.33	-6.27	-5.53	-5.42	-4.19	-5.53
GLY476	-2.70	-0.83	-0.82	-1.67	-2.64	-0.74	-1.87	-1.11	-2.24	-1.05	-3.22	-3.77	-3.06	-2.36	-2.61	-3.06
SER477	-0.97	-1.15	-0.35	-1.29	-3.89	-1.46	-1.91	-0.72	-1.58	-1.19	-0.48	-4.25	-5.11	-2.60	-3.42	-5.11
THR478	-0.58	-0.22	-0.19	-0.85	-5.31	-0.55	-0.80	-0.22	-1.20	-0.26	-1.42	-0.52	-0.55	-0.93	-1.19	-0.55
PRO479	0	-0.29	-0.63	0	-0.12	-0.13	-0.15	-0.24	0	-0.49	-0.11	0	-0.17	-0.17	-0.13	-0.17
CYS480	0	-0.11	-0.13	0	0	0	-0.12	0	-0.13	-0.11	0	0	0.00	0.00	0.00	0.00
ASN481	0	0	0	0	0	0	-0.30	0	-0.12	0	-0.13	0	0.00	0.00	-0.27	0.00
GLY482	0	0	0	0	0	0	0	0	-0.41	0	-0.11	0	0.00	0.00	0.00	0.00
VAL483	0	-0.20	0	0	-0.12	-0.45	-0.18	0	-2.50	0	-0.93	0	0.00	0.00	-0.39	0.00
GLU484	-1.73	-3.91	-8.72	-3.76	-3.36	-5.51	-3.50	-4.09	-0.59	-9.29	-3.39	-0.33	-3.21	-1.42	-1.12	-3.21
GLY485	-0.65	-1.71	-1.30	-0.95	-0.74	-1.10	-1.37	-1.07	-1.55	-1.30	-1.71	-0.13	-0.56	-0.54	-0.88	-0.56
PHE486	-6.86	-8.27	-7.96	-7.24	-7.23	-8.20	-8.46	-7.14	-7.58	-7.59	-8.44	-4.63	-7.61	-7.48	-7.13	-7.61
ASN487	-6.55	-5.87	-5.28	-4.46	-6.01	-5.02	-5.17	-4.96	-5.66	-4.88	-6.42	-7.03	-5.65	-5.56	-5.86	-5.65

CYS488	-0.22	-0.18	-0.11	-0.16	-0.24	-0.19	-0.19	-0.19	-0.18	-0.13	-0.22	-0.27	-0.19	-0.20	-0.20	-0.19
TYR489	-7.55	-8.17	-7.17	-7.42	-9.10	-8.83	-8.74	-6.63	-7.41	-7.97	-6.98	-6.92	-7.61	-7.83	-7.18	-7.61
PHE490	-1.51	-3.52	-1.94	-1.38	-0.93	-4.47	-3.71	-0.68	-1.30	-2.70	-3.33	-0.71	-2.21	-1.17	-2.87	-2.21
PRO491	-0.20	-0.45	-0.42	-0.26	-0.17	-0.35	-0.38	-0.11	-0.19	-0.36	-0.29	-0.10	-0.25	-0.16	-0.31	-0.25
LEU492	-1.22	-2.79	-1.80	-1.38	-0.28	-3.90	-3.04	-0.55	-0.84	-1.64	-1.90	-0.72	-1.32	-1.09	-2.16	-1.32
GLN493	-9.98	-10.87	-9.91	-7.33	-8.66	-13.90	-12.43	-10.52	-10.22	-8.41	-11.25	-11.15	-10.95	-11.74	-12.10	-10.95
SER494	-0.88	-0.34	-0.31	-1.78	-2.57	-0.57	-0.55	-0.43	-1.68	-0.31	-0.53	-0.63	-0.99	-0.85	-0.54	-0.99
TYR495	-2.42	-0.68	-1.41	-3.62	-2.71	-3.10	-1.26	-1.25	-1.48	-1.88	-2.94	-1.64	-0.84	-1.10	-0.88	-0.84
GLY496	-1.73	-0.52	-1.14	-2.55	-1.87	-1.57	-1.90	-2.09	-2.99	-1.83	-1.68	-1.62	-0.57	-0.65	-0.51	-0.57
PHE497	-0.42	-0.22	-0.37	-0.71	-0.35	-0.30	-0.28	-0.35	-0.38	-0.23	-0.63	-0.40	-0.24	-0.27	-0.31	-0.24
GLN498	-4.21	-4.26	-5.13	-3.49	-7.17	-7.26	-5.81	-4.63	-5.98	-11.48	-3.45	-5.73	-2.55	-2.70	-3.37	-2.55
PRO499	-0.49	-0.47	-0.46	-0.47	-0.55	-0.54	-0.47	-0.47	-0.50	-0.52	-0.46	-0.53	-0.48	-0.43	-0.47	-0.48
THR500	-9.23	-9.16	-8.65	-8.95	-9.38	-9.66	-8.45	-8.97	-8.84	-8.72	-8.81	-9.57	-9.14	-8.57	-9.40	-9.14
ASN501	-7.18	-5.96	-6.77	-7.61	-8.04	-9.07	-6.71	-5.81	-6.76	-7.97	-7.25	-8.02	-11.21	-11.38	-10.58	-11.21
GLY502	-3.51	-3.45	-3.15	-3.12	-3.64	-3.46	-3.55	-3.40	-3.70	-3.78	-3.15	-3.59	-3.66	-3.58	-3.48	-3.66
VAL503	-1.15	-0.95	-0.65	-0.69	-0.85	-0.80	-0.79	-0.85	-1.29	-1.32	-0.84	-0.93	-1.26	-1.00	-0.99	-1.26
GLY504	-0.23	-0.23	-0.20	-0.17	-0.25	-0.21	-0.25	-0.21	-0.23	-0.23	-0.17	-0.23	-0.24	-0.27	-0.23	-0.24
TYR505	-10.75	-9.20	-9.83	-8.24	-11.92	-11.39	-11.16	-10.56	-11.08	-10.99	-8.53	-10.97	-11.84	-11.91	-10.58	-11.84
GLN506	-0.47	-0.48	-0.36	-0.37	-0.40	-0.37	-0.36	-0.35	-0.48	-0.49	-0.25	-0.45	-0.47	-0.42	-0.39	-0.47

Bibliography

1 Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* **382**, 1199-1207, doi:10.1056/NEJMoa2001316 (2020).

2 Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497-506, doi:10.1016/s0140-6736(20)30183-5 (2020).

3 Wang, W., Tang, J. & Wei, F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J Med Virol* **92**, 441-447, doi:10.1002/jmv.25689 (2020).

4 Wu, Y. C., Chen, C. S. & Chan, Y. J. The outbreak of COVID-19: An overview. *J Chin Med Assoc* **83**, 217-220, doi:10.1097/jcma.0000000000000270 (2020).

5 Ksiazek, T. G. *et al.* A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**, 1953-1966, doi:10.1056/NEJMoa030781 (2003).

6 Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. & Fouchier, R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* **367**, 1814-1820, doi:10.1056/NEJMoa1211721 (2012).

7 Mohd, H. A., Al-Tawfiq, J. A. & Memish, Z. A. Middle East Respiratory Syndrome Coronavirus (MERS-CoV) origin and animal reservoir. *Virol J* **13**, 87, doi:10.1186/s12985-016-0544-0 (2016).

8 Tang, B. *et al.* An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov). *Infect Dis Model* **5**, 248-255, doi:10.1016/j.idm.2020.02.001 (2020).

9 Guo, Y.-R. *et al.* The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak – an update on the status. *Military Medical Research* **7**, 11, doi:10.1186/s40779-020-00240-0 (2020).

10 Chen, Y., Liu, Q. & Guo, D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J Med Virol* **92**, 418-423, doi:10.1002/jmv.25681 (2020).

11 Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454, doi:10.1038/nature02145 (2003).

12 Wang, N. *et al.* Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Res* **23**, 986-993, doi:10.1038/cr.2013.92 (2013).

13 Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e286, doi:10.1016/j.cell.2020.02.058 (2020).

14 Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e278, doi:10.1016/j.cell.2020.02.052 (2020).

15 Tortorici, M. A. & Veesler, D. Structural insights into coronavirus entry. *Adv Virus Res* **105**, 93-116, doi:10.1016/bs.aivir.2019.08.002 (2019).

16 Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* **382**, 727-733, doi:10.1056/NEJMoa2001017 (2020).

17 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273, doi:10.1038/s41586-020-2012-7 (2020).

18 Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565-574, doi:10.1016/s0140-6736(20)30251-8 (2020).
19 Wan, Y., Shang, J., Graham, R., Baric, R. S. & Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *Journal of Virology* **94**, e00127-00120, doi:10.1128/JVI.00127-20 (2020).
20 Chen, Y., Guo, Y., Pan, Y. & Zhao, Z. J. Structure analysis of the receptor binding of 2019-nCoV. *Biochemical and Biophysical Research Communications* **525**, 135-140, doi:https://doi.org/10.1016/j.bbrc.2020.02.071 (2020).
21 Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiology* **5**, 562-569, doi:10.1038/s41564-020-0688-y (2020).
22 Wang, Q. *et al.* Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **181**, 894-904.e899, doi:10.1016/j.cell.2020.03.045 (2020).
23 Crackower, M. A. *et al.* Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature* **417**, 822-828, doi:10.1038/nature00786 (2002).
24 Yagil, Y. & Yagil, C. Hypothesis. *Hypertension* **41**, 871-873, doi:doi:10.1161/01.HYP.0000063886.71596.C8 (2003).
25 Shang, J. *et al.* Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A* **117**, 11727-11734, doi:10.1073/pnas.2003138117 (2020).
26 Tegally, H. *et al.* Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv*, 2020.2012.2021.20248640, doi:10.1101/2020.12.21.20248640 (2020).
27 Interim: Implications of the Emerging SARS-CoV-2 Variant VOC 202012/01. (2021).
28 Kupferschmidt, K. Mutant coronavirus in the United Kingdom sets off alarms, but its importance remains unclear. *European News*, doi:10.1126/science.abg2626 (2020).
29 National Center for Immunization and Respiratory Diseases (NCIRD), D. o. V. D. US COVID-19 Cases Caused by Variants. (2021).
30 Rambaut, A. *et al.* A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology* **5**, 1403-1407, doi:10.1038/s41564-020-0770-5 (2020).
31 Hu, J. *et al.* Emerging SARS-CoV-2 variants reduce neutralization sensitivity to convalescent sera and monoclonal antibodies. *bioRxiv*, doi:https://doi.org/10.1101/2021.01.22.427749 (2021).
32 Mahase, E. Covid-19: Novavax vaccine efficacy is 86% against UK variant and 60% against South African variant. *BMJ* **372**, n296, doi:10.1136/bmj.n296 (2021).
33 Hollingsworth, S. A. & Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **99**, 1129-1143, doi:10.1016/j.neuron.2018.08.011 (2018).
34 Mortier, J. *et al.* The impact of molecular dynamics on drug design: applications for the characterization of ligand–macromolecule complexes. *Drug Discovery Today* **20**, 686-702, doi:https://doi.org/10.1016/j.drudis.2015.01.003 (2015).
35 Ferreira, L. G., Dos Santos, R. N., Oliva, G. & Andricopulo, A. D. Molecular docking and structure-based drug design strategies. *Molecules* **20**, 13384-13421, doi:10.3390/molecules200713384 (2015).
36 Rees, D. C., Congreve, M., Murray, C. W. & Carr, R. Fragment-based lead discovery. *Nat Rev Drug Discov* **3**, 660-672, doi:10.1038/nrd1467 (2004).

37 Taylor, R. D., Jewsbury, P. J. & Essex, J. W. A review of protein-small molecule docking methods. *Journal of Computer-Aided Molecular Design* **16**, 151-166, doi:10.1023/A:1020155510718 (2002).
38 London, N., Raveh, B. & Schueler-Furman, O. Peptide docking and structure-based characterization of peptide binding: from knowledge to know-how. *Current Opinion in Structural Biology* **23**, 894-902, doi:https://doi.org/10.1016/j.sbi.2013.07.006 (2013).
39 Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *Journal of the American Chemical Society* **125**, 1731-1737, doi:10.1021/ja026939x (2003).
40 Yan, R. *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444-1448, doi:10.1126/science.abb2762 (2020).
41 Berman, H. M. *et al.* The Protein Data Bank. *Nucleic acids research* **28**, 235-242, doi:gkd090 [pii] (2000).
42 Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **25**, 247-260, doi:10.1016/j.jmgm.2005.12.005 (2006).
43 Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry* **97**, 10269-10280, doi:10.1021/j100142a004 (1993).
44 Gaussian 16 Rev. C.01 (Wallingford, CT, 2016).
45 Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696-3713, doi:10.1021/acs.jctc.5b00255 (2015).
46 Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-1174, doi:10.1002/jcc.20035 (2004).
47 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926-935, doi:10.1063/1.445869 (1983).
48 Case, D. A. *et al.* AMBER 2018. *University of California, San Francisco* (2018).
49 Larini, L., Mannella, R. & Leporini, D. Langevin stabilization of molecular-dynamics simulations of polymers by means of quasisymplectic algorithms. *J Chem Phys* **126**, 104101, doi:10.1063/1.2464095 (2007).
50 Darden, T., Perera, L., Li, L. & Pedersen, L. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* **7**, R55-60 (1999).
51 Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **9**, 3084-3095, doi:10.1021/ct400341p (2013).
52 Gohlke, H. & Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J Comput Chem* **25**, 238-250, doi:10.1002/jcc.10379 (2004).
53 Kollman, P. A. *et al.* Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* **33**, 889-897, doi:10.1021/ar000033j (2000).

54 Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling* **51**, 69-82, doi:10.1021/ci100275a (2011).
55 Rocchia, W., Alexov, E. & Honig, B. Extending the Applicability of the Nonlinear Poisson−Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions. *The Journal of Physical Chemistry B* **105**, 6507-6514, doi:10.1021/jp010454y (2001).
56 Kuhn, B., Gerber, P., Schulz-Gasch, T. & Stahl, M. Validation and Use of the MM-PBSA Approach for Drug Discovery. *Journal of Medicinal Chemistry* **48**, 4040-4048, doi:10.1021/jm049081q (2005).
57 Swanson, J. M. J., Henchman, R. H. & McCammon, J. A. Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. *Biophysical Journal* **86**, 67-74, doi:10.1016/S0006-3495(04)74084-9 (2004).
58 Wang, E. *et al.* End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chemical Reviews* **119**, 9478-9508, doi:10.1021/acs.chemrev.9b00055 (2019).
59 Wang, J. & Hou, T. Develop and test a solvent accessible surface area-based model in conformational entropy calculations. *J Chem Inf Model* **52**, 1199-1212, doi:10.1021/ci300064d (2012).
60 Junmei, W., Tingjun, H. & Xiaojie, X. Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models. *Current Computer-Aided Drug Design* **2**, 287-306, doi:http://dx.doi.org/10.2174/157340906778226454 (2006).
61 Knight, J. L. & Brooks, C. L., 3rd. Surveying implicit solvent models for estimating small molecule absolute hydration free energies. *J Comput Chem* **32**, 2909-2923, doi:10.1002/jcc.21876 (2011).
62 Hawkins, G. D., Cramer, C. J. & Truhlar, D. G. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *The Journal of Physical Chemistry* **100**, 19824-19839, doi:10.1021/jp961710n (1996).
63 Hawkins, G. D., Cramer, C. J. & Truhlar, D. G. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *Journal of Physical Chemistry* **100**, 19824-19839, doi:DOI 10.1021/jp961710n (1996).
64 Wang, E. *et al.* Assessing the performance of the MM/PBSA and MM/GBSA methods. 10. Impacts of enhanced sampling and variable dielectric model on protein-protein Interactions. *Phys Chem Chem Phys* **21**, 18958-18969, doi:10.1039/c9cp04096j (2019).
65 Ou, J. *et al.* Emergence of RBD mutations in circulating SARS-CoV-2 strains enhancing the structural stability and human ACE2 receptor affinity of the spike protein. *bioRxiv*, 2020.2003.2015.991844, doi:10.1101/2020.03.15.991844 (2020).
66 Jia, Y. *et al.* Analysis of the mutation dynamics of SARS-CoV-2 reveals the spread history and emergence of RBD mutant with lower ACE2 binding affinity. *bioRxiv*, 2020.2004.2009.034942, doi:10.1101/2020.04.09.034942 (2020).
67 ACROBIosystems. Infectivity of SARS-CoV-2 with Spike Protein Mutations. News-Medical. *https://www.news-medical.net/whitepaper/20200618/Infectivity-of-SARS-CoV-2-with-Spike-Protein-Mutations.aspx*. (2020).

68 Leaver-Fay, A. *et al.* Scientific benchmarks for guiding macromolecular energy function improvement. *Methods Enzymol* **523**, 109-143, doi:10.1016/B978-0-12-394292-0.00006-0 (2013).
69 Park, H. *et al.* Simultaneous Optimization of Biomolecular Energy Functions on Features from Small Molecules and Macromolecules. *J Chem Theory Comput* **12**, 6201-6212, doi:10.1021/acs.jctc.6b00819 (2016).
70 Sciences, N. C. f. A. T. Spike-ACE2 protein-protein interaction (AlphaLISA). *Natioanl Institutes of Health* (2020).