Computational and Structural Biotechnology Journal 21 (2023) 5125–5135

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research article

Collaborative effects of 2019-nCoV-Spike mutants on viral infectivity

Senbiao Fang [a,1], Chuqi Lei [a,1], Meng Li [b,1], Yongfan Ming [a], Liren Liu [c], Xuming Zhou [b,*], Min Li [a,*]

[a] Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, Changsha 410083, China
[b] Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
[c] Department of Molecular Pharmacology, Tianjin Medical University Cancer Institute & Hospital; National Clinical Research Center for Cancer; Key Laboratory of Cancer Prevention and Therapy, Tianjin; Tianjin's Clinical Research Center for Cancer, Tianjin 300060, China

ARTICLE INFO

Keywords:
COVID-19
Molecular dynamics simulations
Multi-mutant
Infection assay
Spike

ABSTRACT

Background: The emerging mutants of the 2019-nCoV coronavirus are posing unprecedented challenges to the pandemic prevention. A thorough, understanding of the mutational characterization responsible for the pathogenic mechanisms of mutations in 2019-nCoV-Spike is indispensable for developing effective drugs and new vaccines.

Methods: We employed computational methods and viral infection assays to examine the interaction pattern and binding affinity between ACE2 and both single- and multi-mutants of the Spike proteins.

Results: Using data from the CNCB-NGDC databank and analysis of the 2019-nCoV-Spike/ACE2 interface crystal structure, we identified 31 amino acids that may significantly contribute to viral infectivity. Subsequently, we performed molecular dynamics simulations for 589 single-mutants that emerged from the nonsynonymous substitutions of the aforementioned 31 residues. Ultimately, we discovered 8 single-mutants that exhibited significantly higher binding affinities (<−65.00 kcal/mol) to ACE2 compared with the wild-type Spike protein (−55.07 kcal/mol). The random combination of these 8 single-mutants yielded 184 multi-mutants, of which 60 multi-mutants exhibit markedly enhanced binding affinities (<−65.00 kcal/mol). Moreover, the binding free energy analyses of all 773 mutants (including 589 single- and 184 multi-mutants) revealed that Y449R and S494R had a synergistic effect on the binding affinity with other mutants, which were confirmed by virus infection assays of six randomly selected multi-mutants. More importantly, the findings of virus infection assay further validated a strong association between the binding free energy of Spike/ACE2 complex and the viral infectivity.

Conclusions: These findings will greatly contribute to the future surveillance of viruses and rational design of therapeutics.

1. Introduction

Since the first case of Coronavirus disease 2019 (COVID-19), over 260 million people have been infected and 6.34 million people died directly linked to the infection, causing a serious global public health emergency [1]. The fight against COVID-19 has seen remarkable advancements in the development of therapeutics. Notably, pharmaceutical giants such as Moderna, Merck, Pfizer, BioNTech, and Sinovac have made several vaccines available on the market. Furthermore, Molnupiravir and PaxlOVID, oral drugs, have been authorized for emergency treatment of COVID-19 patients [2–7]. However, the emerged 2019-nCoV variants compromised the therapeutic efficacy by enhancing the infectivity and transmissibility of the virus. In February 2020, a SARS-CoV-2-Spike mutant (D614G) was identified that significantly enhanced the entry of the virus with enhanced ACE2-binding affinity [8]. In September 2020, the Alpha variant N501Y (B.1.1.7) was first detected in southeast England, with a transmissibility rate roughly 1.5 times higher than the earlier SARS-CoV-2 variants, leasing to an increased risk of death by 1.6-fold [9]. The other major variants, such as Beta (B.1.351), Gamma (P.1), Zeta (P.2), Epsilon (B.1.429), Eta (B.1.525), Iota (B.1.526–1) and Iota (B.1.526–2), also show increased infectivity to human and monkey cell lines [10,11]. Notably, the

* Corresponding authors.
E-mail addresses: zhouxuming@ioz.ac.cn (X. Zhou), limin@mail.csu.edu.cn (M. Li).
[1] These authors contributed equally.

https://doi.org/10.1016/j.csbj.2023.10.030
Received 6 July 2023; Received in revised form 14 October 2023; Accepted 16 October 2023
Available online 17 October 2023

B.1.617double-mutant strain (E484Q-L452R) discovered in India has been found to possess a considerable ability to escape neutralizing antibodies, consequently diminishing the efficacy of SARS-CoV-2 vaccines [12]. A triple-mutant Beta (N501Y-K417N-E484K) strain exhibits an evidently higher infectivity in ACE2-expressing mouse cells [13]. A plethora of evidence demonstrates that new variants, particularly those with multiple mutations emerge rapidly, marking a new era in the COVID-19 epidemic [14]. Therefore, understanding mutational characterization of the emerging mutants, especially the multi-mutants, will help to elucidate their pathogenic mechanisms, thereby facilitating the development of effective drugs and new vaccines.

The 2019-nCoV-Spike glycoprotein plays key roles in virus attachment, host cell entry, thus may serve as an ideal target for developing vaccines and potential drugs [15]. The origin and evolution studies on SARS-CoV-2 have identified a number of mutations on the Spike region with higher binding activities to human ACE2 receptor [16]. To delineate the impact of Spike mutant, particularly the collective effects of multi-mutants, on viral infectivity, we initially categorized the key residues on the 2019-nCoV-Spike protein that might influence the viral infectivity via structural modeling. We then investigated the interaction pattern and binding affinity between ACE2 and potential Spike single- and multi-mutants using *in silico* approaches and virus infection assays. Our findings strongly support the assertion that these mutations could interact synergistically to magnify the viral infectivity, and thus providing helpful insights into the in-depth understanding of 2019-nCoV coronavirus.

2. Material and methods

2.1. Principle for selecting single mutation sites from CNCB-NGDC databank

The following steps were followed to assembled the single-mutants meeting that meets the 3 H-principal criteria:

2.1.1. High hazardous effect

Within the CNCB-NGDC databank, a classification system has been implemented to categorize the potential hazard effects of mutation sites based upon explicit defining criteria for hazardous size, resulting in four distinct categories: High > Moderate > Modifier > Low.

To maximize mutations' influence, amino acid sites with low rankings were excluded and remaining three classes of mutations were assigned.

2.1.2. High mutation frequency

The databank also grouped all mutation sites into I∼V series considering mutation frequency on high quality sequences and the distribution of variations. The lower the frequency grade, the higher impacts of mutations. We selected mutation sites within the range of I-IV grades.

2.1.3. High distribution

Regional and temporal variability can indicate the specific distributions of mutations in diverse regions and countries. A higher greater of regional variance suggests that variant strains were collected from various locations and have a broader spatial distribution. The greater temporal variance indicates the long-time existence of the variants.

2.2. Single-mutants preparation

The crystal structure of 2019-nCoV-Spike trimer with two RBDs (PDB ID: 7CAI) was used as the initial wild-type structure, and the monomer 2019-nCoV-Spike/ACE2 was extracted using UCSF Chimera [21]. Mutated systems were constructed based on the wild-type 2019-nCoV--Spike/ACE2 complex using Mutagenesis Wizard tool on PyMol software [22]. Removal of crystal water molecules and addition of hydrogen atoms to protein residues for each mutant were accomplished through the deployment of the leap module. Standard protonation states were evaluated using Pro-pKa. 17 sodium ions Na^+ were added for neutralizing the complex system. The complex was then solvated in a cubic periodic box of explicit TIP3P water model that extended a minimum 10 Å distance from protein atoms. To prevent potential clashes between atoms and excessive conformational changes, all heavy atoms were subjected to positional restraints using a constant force of 1500 kcal/mol.

2.3. Molecular dynamics simulation

The Amber (version 16) Molecular Dynamics package [23] was used to perform molecular dynamics simulation for each complex, applying AMBER FF99SB force field. Each mutant system underwent 5000 steps of steepest descent, and 5000 steps of conjugate gradient minimization (EM). The system was then heated to a temperature of 298 K for a period of 2000 ps under constant volume periodic boundary conditions (NVT), with weak positional restraints applied to the protein atoms. Constant temperature was maintained at 298 K with the NPT ensemble, and approximately 1 ns temperature simulation was carried out to equilibrate the system. Finally, at least > 6 ns of production simulation was performed for each wild-type and each mutant system under the same conditions. A cutoff value of 10 Å was implemented for non-bonded interactions and long-range electronic interactions calculations, which were carried out using the Particle Mesh Ewald (PME) method. Ultimately, a minimum of 100 snapshots were obtained from the > 6 ns MD simulation trajectory to determine the final average structure of each mutant and wild system. The free binding energy between 2019-nCoV--Spike and ACE2 were computed for each snapshot and then averaged using MMPBSA module in AMBER software.

2.4. Structure preparation of 184 multiple-mutants and MD simulation

Similar to the method mentioned above, mutated process of 3D multiple mutants' structures were executed in Mutagenesis Wizard tool on PyMol software. We performed a minimum of 6 ns molecular dynamics simulations to get insights into the thermodynamic and structural implications of tertiary changes associated with mutations. During MD simulations, a constant 1500 kcal/mol force was applied to all heavy atoms on both Spike and ACE2 proteins. Root Mean Square Deviation (RMSD) of heavy atoms for each multiple-mutant was checked until the MD simulation reached equilibrium state. Binding free energies of 2019-nCoV-Spike/ACE2 mutants were extracted for difference analysis.

2.5. Cell lines and plasmids

The cell lines HEK-293FT, HEK-293FT-hACE2, A549, and Huh-7 were cultured in Dulbecco's modified Eagle medium (DMEM, Gibco, Cat #C11995500BT) which was supplemented with 20 mM N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid (HEPES, GIBCO, Cat#15630080), 1% of Penicillin Streptomycin solution (GIBCO, Cat #15140163) and 10% fetal bovine serum (FBS, Gibco, Cat #10091148). The cells were cultured at 37 ℃ with 5% CO_2, and their passage were performed using 0.25% EDTA trypsin every 2–3 days.

The codon-optimized full-length coding sequences of wild type or mutated Spike gene of SARS-CoV-2 (Wuhan-Hu-1 strain, GenBank NC_045512.2) [24] were synthesized and respectively cloned into pcDNA3.1-myc-HisA vector. Plasmids were transformed into DH5α *E. coli* and extracted using QIAGEN Plasmid Maxi Kit (QIAGEN, Cat #12163).

2.6. Production and titration of pseudotyped viruses

The procedures for incorporating pseudotyped viruses with wild type (WT) and mutated Spike protein were described previously [25]. An

amount of 20 μg of DNA plasmid expressing the spike protein was transfected into HEK-293FT cells (60%−80% confluence) using Lipofectamine 3000. After 24 h of transfer, the cells were infected with G*△G-VSV (VSV-G pseudotyped virus, MOI=3). Following a 2-hour incubation period, the supernatant of the cell was discarded. Then, 10 mL fresh complete DMEM medium was added and cultured for one day. Finally, SARS-CoV-2 pseudotyped viruses with supernatants were harvested, filtered and stored at − 80 ℃ in 1 mL aliquots until further use.

The titration and operation of the pseudotyped virus, as well as the benchmark for the 50% tissue culture infectious dose (TCID50) of the SARS-CoV-2 pseudo-virus, were performed, as previously described. Adjusted 293FT-hACE2 cells to 5×10^3 cells/well in 96-well plates. This procedure can avoid inconsistencies resulting from the following repeated freezing-thawing cycles. For titration of the pseudotyped virus (wild-type spike and 6 multiple-mutants of spike protein), a gradually decrease of 2-fold, 4-fold, 8-fold, 15-fold, and 30-fold initial dilution with five replicates was made in 96-well culture plates. The last column was assigned as the cells control without pseudotyped virus. Subsequently, the 96-well plates were seeded with 293FT-hACE2 cells adjusted to 5×10^3 cells/well. After 24 h incubation, the above cells were collected for further flow cytometry assays. The median tissue culture infectious dose (TCID50) defined as the dilution of a virus required to infect 50% of a given cell culture, was calculated using the Reed and Muench method [26].

2.7. Quantification of pseudotyped virus particles using RT-PCR

Before qualification, we used a 25% sucrose gradient ultracentrifugation at 50,000 g for 4 h to purify all the pseudotyped viruses [25]. The 200 μL purified pseudotyped viruses suspensions of wild-type and multiple-mutants were used for extracting viral RNA through Trizol regentt (Life invitrogen, Cat# 15596018). Following this, the viral RNA for either the wild-type or multiple-mutant was used as the template for subsequent reverse transcription, utilizing the GoScript™ Reverse Transcription System (Promega, Cat# A5001). The real-time PCR experiments were implemented for virus quantification. The LightCycler 480 SYBR Green I Master (Roche, Cat #04887352001) was used for PCR experiments. The Spike protein gene of 2019-nCoV coronavirus (wild--type and 6 multiple-mutants) was cloned into the vector pCDNA3.1-myc-HisA as a plasmid standard. Finally, the viral copy number was calculated accordingly.

2.8. Infection assay

Pseudotyped viral particle numbers within the same range were obtained and normalized to the same level using quantitative RT-PCR module. 100 μL pseudotyped virus were interpolated to the A549 or Huh7 cells in 48-well cell culture plate after normalization. The plates were then incubated at 37 ℃ in a humidified atmosphere with 5% CO_2. The cells were trypsin-digested 24 h post infection, and flow cytometry were performed to account the rate of infected cells. Each group contained 3–5 replicates.

2.9. Quantification and statistical analysis

Plotting and statistical analysis were conducted by GraphPad Prism 8 software. The mean and standard error of mean (SEM) were used to describe the variability of infective assay values within the sample. One-way ANOVA and Holm-Sidak's multiple comparisons test was used to analyze the differences between groups. High percentage of infected cells was considered to be enhanced infectivity.

3. Results

3.1. Identification of 31 key amino acids in 2019-nCov-Spike protein

Firstly, the SARS-CoV-2 Spike trimer with RBD domain (RCSB PDB ID: 7CAI) adjusted at a low resolution of 3.49 Å was adopted to identify key amino acids at 2019-nCoV-Spike/ACE2 binding interface. The monomer structure. containing the Spike region and receptor ACE2 was extracted (Fig. 1 A, B) [17] and amino acids located within a 5.0 Å distance from each protein were regarded as protein-protein binding interface (Fig. 1 C). An intricate interaction network connecting 2019-nCoV-Spike and the ACE2 protein was identified (Fig. 1 C), showing that 18 amino acids from the ACE2 and 16 corresponding residues from 2019-nCoV-Spike can establish favorable contacts with each other. The complex is stabilized by non-bonded hydrogen bonds and salt bridge interactions, predominantly through the contribution of polar amino acids, particularly 13 polar residues from ACE2, interact with 7 residues from Spike protein to form a network of polar interactions. Among those, the strongest salt bridge is formed between the residue K417 (Spike) and opposite D30 (ACE2). In addition, a typical hydrogen bond exists as the amino acid Q493 (Spike) orientated itself simultaneously to residues K31 and E35 (ACE2). Besides hydrophilic interactions, a handful hydrophobic interactions also played an indispensable role in protein-protein interactions, including π-π stacking of the aromatic residues F486 and Y505 (Spike) with the neighboring Y83 and Y41, respectively (ACE2). According to the hydrogen bonds and the complementarity of hydrophobic/hydrophilic properties on 2019-nCoV-Spike/ACE2 binding interface, 16 amino acids (K417, G446, Y449, Y453, L455, F456, F486, N487, Y489, Q493, G496, Q498, T500, N501, G502, and Y505) in Spike protein were identified as key residues.

We then retrieved a total of 2681 mutations that encode 215 amino acids in the 2019-nCoV-Spike protein from the 2019 Novel Coronavirus Information databank (CNCB-NGDC, https://bigd.big.ac.cn/nCoV/) [18]. Single-mutants meeting 3 H-principal criteria with high hazardous effect, high mutation frequency and high distribution characteristics were included (Fig. 1D), thereby 18 mutation sites including R357, V367, S371, S375, V382, V407, Q414, K417, N439, N440, L452, A475, S477, T478, E484, Y489, S494, and N501 were selected for subsequent analysis (Fig. 1E). Among these, 3 amino acid sites (K417, Y489, and N501) were located in the binding interface. Thus, a total of 31 amino acids in 2019-nCov-Spike protein were classified as the key residues that may play important roles in the infectivity of 2019-nCov virus.

3.2. Analyses of the binding free energy for 589 single-mutants in Spike

To investigate the change in binding-mode of each 2019-nCoV-Spike mutant with ACE2, all nonsynonymous substitutions in the 31 amino-acid sites were analyzed, resulting in 589 single-mutants, which were then employed for > 6 ns molecular dynamics simulations. The binding free energy between 2019-nCoV-Spike and ACE2 was calculated for each snapshot and averaged using MMPBSA module in AMBER software [23]. Since the binding free energy between wild-type Spike region and human ACE2 was − 55.07 kcal/mol, a free energy value < −60 kcal/mol or > −50 kcal/mol is considered as an enhanced or weakened effect due to the perturbation respectively, whereas a value ranging from − 60 kcal/mol to − 50 kcal/mol represents a neutral result in the present study. The analysis revealed that the binding free energies of all single-mutants varied between − 76.54 kcal/mol to − 31.68 kcal/mol (Fig. 2 and Table S1), which corresponds to our previous findings [19,20]. The protein-protein binding affinities of 258 mutants, which represent 47% of 589 single-mutants, have apparently improved (<−55.07 kcal/mol) when compared to the wild-type Spike (Fig. 2 and Fig. 3B). Three single-mutants S494R, G446Q and E484Q show high binding affinities to ACE2, among which the binding free energy of S494R (−76.54 kcal/mol) is 38.98% higher than that of the wild-type system. As expected, the binding energies of two new

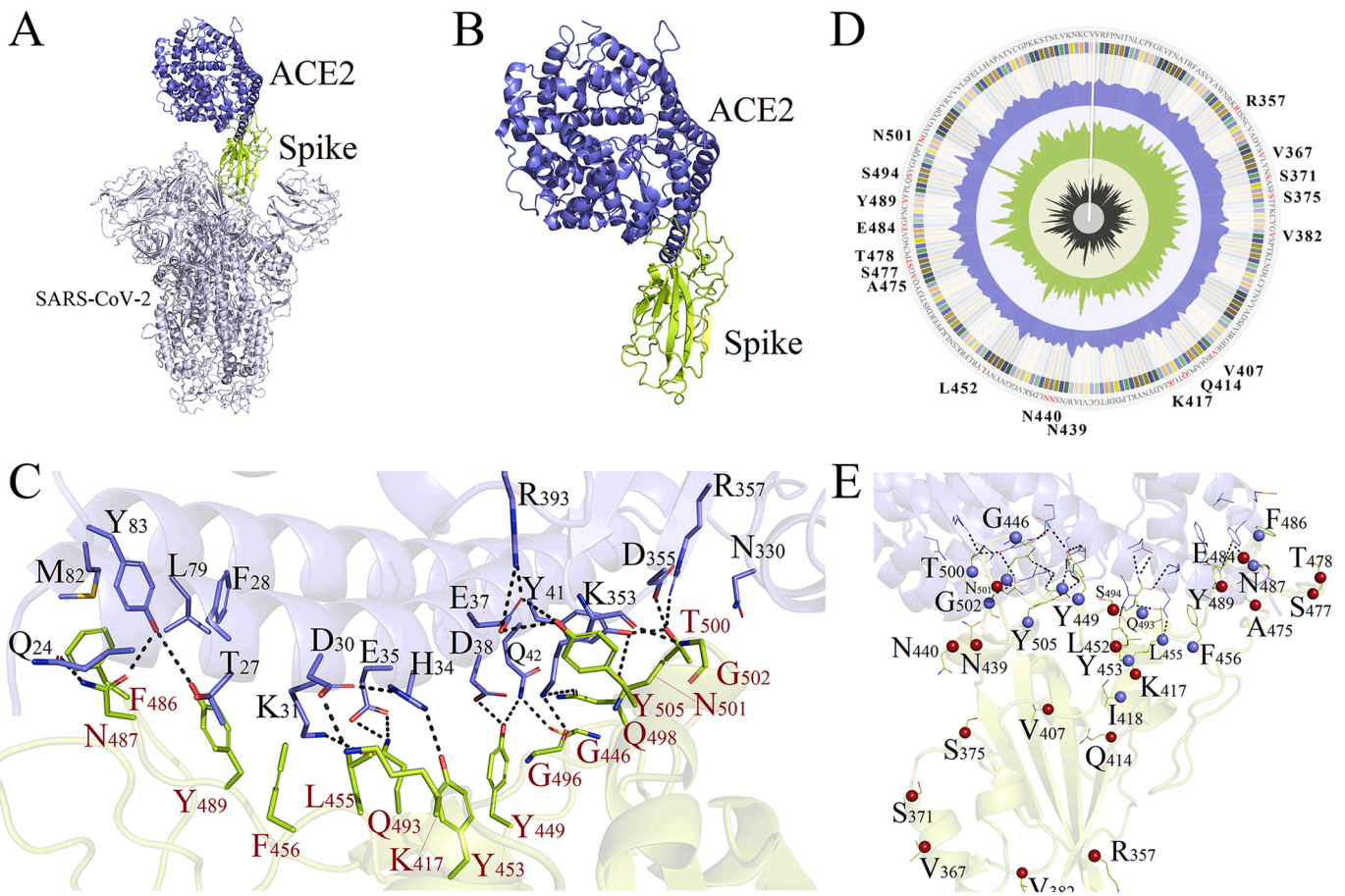

Fig. 1. 31 amino-acid sites with mutational properties on 2019-nCov-Spike region were identified. (A, B) Overall structure of the coronavirus 2019-nCoV-Spike in complex with human receptor ACE2. (C) Schematic of binding interface between 2019-nCoV-Spike and ACE2 receptor. 16 amino-acid sites on the Spike region were selected. Crystal structure of SARS-Cov-2 spike trimer (PDB ID: 7CAI) is shown with top (left and right panels) views. Binding interface of 2019-nCoV-Spike/ACE2 complex was shown as cartoon. Spike regions and ACE2 are colored as follows: SARS-CoV-2 in white, ACE2 in blue, and Spike in cyan. Key residues from ACE2 and Spike receptor were shown as sticks and colored as blue and cyan. The hydrogen bonds are shown in dotted lines. (D) Mutations on 2019-nCoV-Spike protein selected from CNCB-NGDC databank. (E) 31 key residues with mutational properties were identified. 18 key residues with high hazardous effects, high mutation frequency and high distribution were shown as sticks. One total number of 16 amino acid sites directly formed hydrogen bonds with ACE2 receptor were shown as blue sphere.

mutations (E484Q and L452R) were − 68.63 kcal/mol and − 60.33 kcal/mol, respectively, in agreement with the reported enhanced infectivity and lethality of Indian double-mutants (E484Q-L452R) [12]. Among 589 single-mutants, 8 mutations, including G446R, G446Q, Y449R, F456Y, E484N, F486Y, S494R and G496S, that locate at the binding interface of 2019-nCoV-Spike/ACE2 complex show significantly enhanced binding affinities (<−65.00 kcal/mol). Of note, G446R and E484Q have been reported to have a high mutation-frequency of 13.81% and 2.44%, respectively, among all existing mutations.

3.3. Prediction of the binding affinities for 184 multi-mutants

A total number of 184 multiple-mutants including 28 double-mutants, 50 triple-mutants, 55 tetra-mutants, 36 penta-mutants, 13 hexa-mutants and 2 hepta-mutants were obtained by a random arrangement of the above 8 single-mutants (Fig. 3A and Table S2-S7). We found 54% (99/184) of the multi-mutants show positive effects on their binding modes to ACE2 compared with the wild-type complex system (Fig. 3I). Additionally, the data indicated that as the number of mutation sites increases, a greater proportion of mutants exhibit lower energies (<−55.07 kcal/mol), showing a synergistic improvement among the 8 single-mutants. Among 28 double-mutants, the binding free energies in 18 systems were < −55.07 kcal/mol. The significant low binding energies (<−65.00 kcal/mol) were observed in 5 double-mutant systems, namely G446Q-F456Y, G446R-S494R, Y449R-G496S, E484Q-S494R and E484Q-G496S. In particular, the binding affinity for G446Q-F456Y was − 70.61 kcal/mol, 28.22% lower than that of the wild-type system (Fig. 3C and Table 1).

Multi-mutants appear to exert a more pronounced influence on viral infectivity, as evidenced by the lower energies observed in triple-mutants, tetra-mutants, penta-mutants, and hexa-mutants, which ranged from 66% to 100% of the wild-type system (Fig. 3D-G). Five multi-mutant systems, including F456Y-S494R-G496S, G446Q-Y449R-E484Q-S494R, Y449R-F456Y-F486Y-S494R, Y449R-F456Y-E484Q-F486Y-S494R and G446R-Y449R-E484Q-S494R-G496S, show binding energies < −80.00 kcal/mol. Specifically, the binding energies for 3 hexa-mutants, namely G446R-Y449R-F456Y-E484Q-S494R-G496S, Y449R-F456Y-E484Q-F486Y-S494R-G496S and G446Q-Y449R-F456Y-E484Q-F486Y-G496S were − 70.4191 kcal/mol, − 71.4697 kcal/mol and − 70.3267 kcal/mol (Fig. 3G and Table 1), while 2 hepta-mutants show values at − 72.7691 kcal/mol and − 69.2738 kcal/mol, respectively (Fig. 3H and Table 1).

However, it is worth noted that not all mutations can enhance the binding affinity to ACE2. In contrast, some mutants apparently disturb the structural stability thereby weaken the interaction between Spike and ACE2 with a binding energy > −50.00 kcal/mol. For example, 3/27 double-mutants, 2/50 triple-mutants and 1/55 tetra-mutants have the

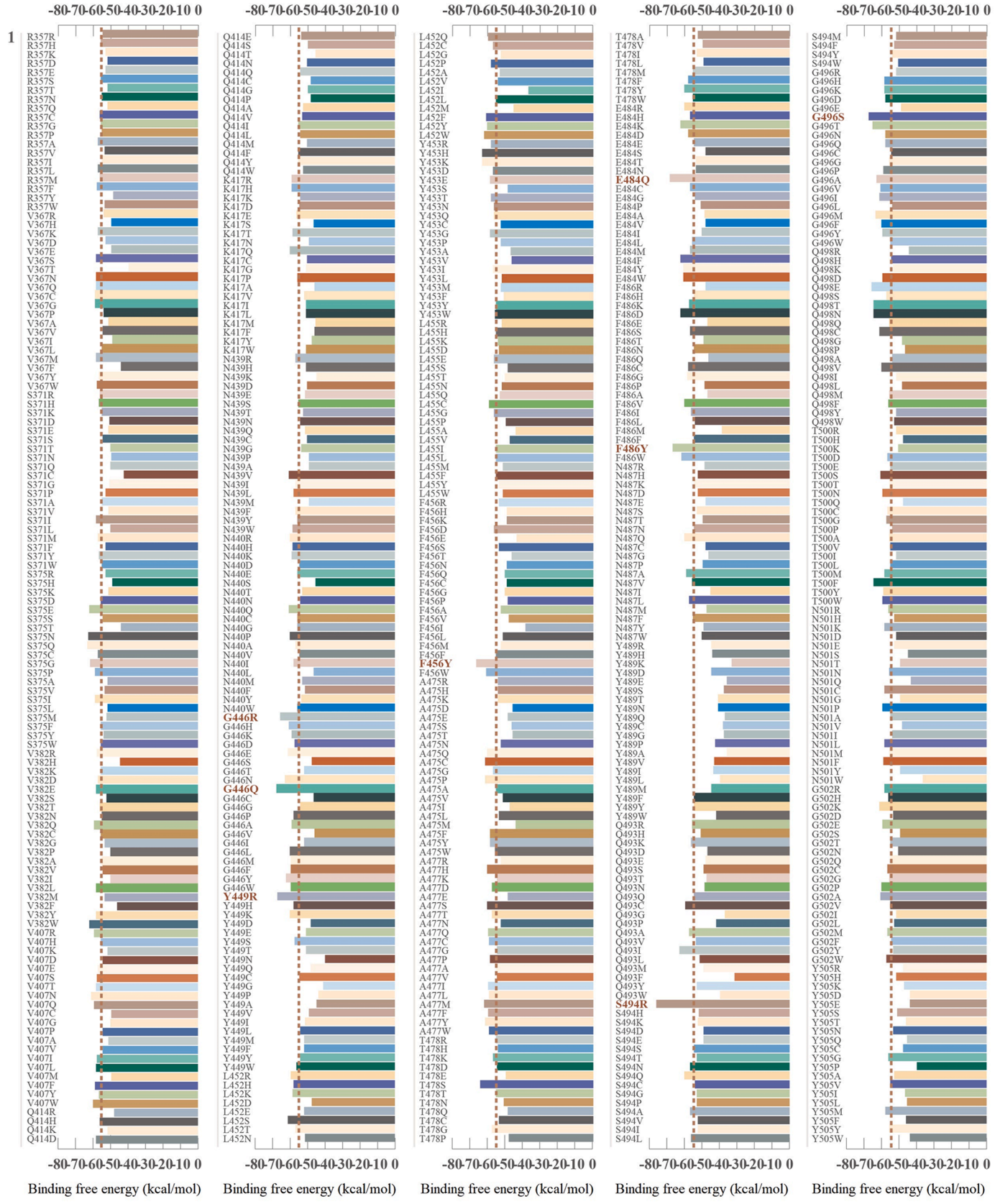

Fig. 2. Binding free energies for 589 single mutants. 1 represent the single-mutants. Mutants with enhanced binding energies (<−65.00 kcal/mol) compared with wild type system were considered significant improvement systems and marked as red.

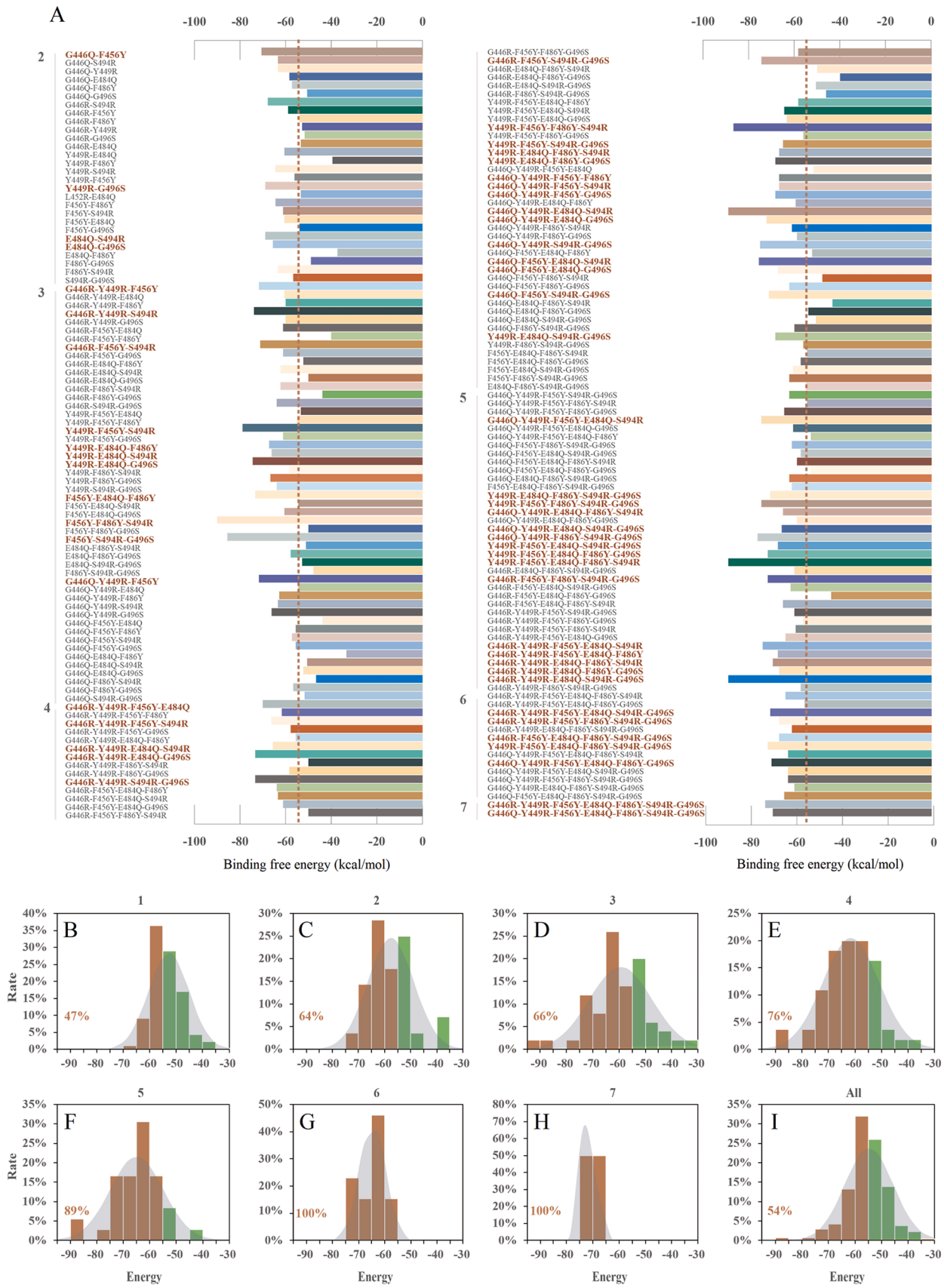

Fig. 3. Binding free energies for 184 multiple mutants. 1, 2, 3, 4, 5, 6 and 7 represent the single-, double-, triple-, tetra-, penta-, hexa- and hepta- mutants. Mutants with enhanced binding energies (<−65.00 kcal/mol) compared with wild type system were considered significant improvement systems and marked as red. Most of the multiple-mutants on the 2019-nCoV-Spike protein could enhance the binding ability between the 2019-nCoV coronavirus and the human protein ACE2.

Table 1
Multiple-mutants improving binding affinity (<−65.00 kcal/mol) compared with wild 2019-nCov-Spike/ACE2 complex. 6 multiple-mutants marked as red were randomly selected for virus infection assays.

Multiple - mutants improving binding affinity (< - 65.00 kcal/mol) compared with wild 2019 - nCov - Spike/ACE2 complex.6 multiple- mutants marked as red were randomly selected for virus infection assays.					
G446Q-F456Y	-70.61	G446R-Y449R-E484Q-S494R	-65.47	Y449R-F456Y-F486Y-S494R-G496S	-74.29
G446R-S494R	**-67.98**	G446R-Y449R-E484Q-G496S	-73.14	G446Q-Y449R-E484Q-F486Y-S494R	-65.26
Y449R-G496S	-68.62	G446R-Y449R-S494R-G496S	-73.10	G446Q-Y449R-E484Q-S494R-G496S	-65.75
E484Q-S494R	-68.64	G446R-F456Y-S494R-G496S	-74.55	G446Q-Y449R-F486Y-S494R-G496S	-76.30
E484Q-G496S	-65.38	Y449R-F456Y-F486Y-S494R	-86.53	Y449R-F456Y-E484Q-S494R-G496S	-67.33
G446R-Y449R-F456Y	-71.85	Y449R-F456Y-S494R-G496S	-65.14	Y449R-F456Y-E484Q-F486Y-G496S	-71.76
G446R-Y449R-S494R	-73.80	Y449R-E484Q-F486Y-S494R	-66.77	Y449R-F456Y-E484Q-F486Y-S494R	-89.15
G446R-F456Y-S494R	-71.24	Y449R-E484Q-F486Y-G496S	-68.21	G446R-F456Y-F486Y-S494R-G496S	-71.56
Y449R-F456Y-S494R	-78.81	G446Q-Y449R-F456Y-F486Y	-66.91	G446R-Y449R-F456Y-E484Q-S494R	-74.07
Y449R-E484Q-F486Y	-67.35	G446Q-Y449R-F456Y-S494R	-66.79	G446R-Y449R-F456Y-E484Q-F486Y	-67.06
Y449R-E484Q-S494R	-66.36	G446Q-Y449R-F456Y-G496S	-68.55	G446R-Y449R-E484Q-F486Y-S494R	-69.57
Y449R-E484Q-G496S	-74.29	G446Q-Y449R-E484Q-S494R	-89.15	G446R-Y449R-E484Q-F486Y-G496S	-66.82
Y449R-F486Y-G496S	-66.49	**G446Q-Y449R-E484Q-G496S**	**-72.08**	G446R-Y449R-E484Q-S494R-G496S	-89.16
F456Y-E484Q-F486Y	-73.09	G446Q-Y449R-S494R-G496S	-75.22	**G446R-Y449R-F456Y-E484Q-S494R-G496S**	**-70.41**
F456Y-F486Y-S494R	-90.01	G446Q-F456Y-E484Q-S494R	-75.36	G446R-Y449R-F456Y-F486Y-S494R-G496S	-66.79
F456Y-S494R-G496S	**-85.79**	G446Q-F456Y-E484Q-G496S	-67.09	G446R-F456Y-E484Q-F486Y-S494R-G496S	-66.64
G446Q-Y449R-F456Y	-71.85	G446Q-F456Y-S494R-G496S	-71.26	Y449R-F456Y-E484Q-F486Y-S494R-G496S	-71.46
G446Q-Y449R-G496S	-66.4	Y449R-E484Q-S494R-G496S	-68.12	G446Q-Y449R-F456Y-E484Q-F486Y-G496S	-70.32
G446R-Y449R-F456Y-E484Q	-70.09	**G446Q-Y449R-F456Y-E484Q-S494R**	**-74.64**	**G446R-Y449R-F456Y-E484Q-F486Y-S494R-G496S**	**-72.7691**
G446R-Y449R-F456Y-S494R	-65.88	Y449R-E484Q-F486Y-S494R-G496S	-70.53	G446Q-Y449R-F456Y-E484Q-F486Y-S494R-G496S	-69.2738

binding energy value ranging from − 40.00 kcal/mol to − 30.00 kcal/mol, which indicate a disruptive binding ability.

3.4. Enhanced viral infectivity of 6 multi-mutants

To validate the *in silico* predictions, we randomly selected 6 multi-mutants (ranging from double- to hepta-mutants), including G446R-S494R, F456Y-S494R-G496S, G446Q-Y449R-E484Q-G496S, G446Q-Y449R-F456Y-E484Q-S494R, G446R-Y449R-F456Y-E484Q-S494R-G496S and G446R-Y449R-F456Y-E484Q-F486Y-S494R-G496S, from the 60 multi-mutants (5 double-mutants, 13 triple-mutants, 20 tetra-mutants, 15 penta-mutants, 5 hexa-mutants and 2 hepta-mutants) showing significant enhanced binding affinities (<−65.00 kcal/mol) to ACE2 for virus infection assays (Table 1). Two cell lines, Huh-7 and A549, were infected with pseudotyped viruses containing the wild-type 2019-nCoV-Spike or the 6 multi-mutants (Fig. 4A). Although both cell lines were susceptible to infection by the 2019-nCoV-Spike pseudotyped virus, the infection efficiency of Huh-7 cells was higher than that of A549 cells. Our findings showed now discernible variation in viral infectivity, as demonstrated by the proportion of GFP^+ population, across all the tested samples at a concentration of 5 ug/mL in both cell lines. Nonetheless, the augmented infectivity capabilities were detected in all 6 mutants, as exemplified by the microscopic observations of the GFP^+ population of the F456Y-S494R-G496S mutant and the wild-type control (Fig. 4B & C). Among the 6 multi-mutants,the triple F456Y-S494R-G496S mutant display the highest level of viral infectivity to Huh-7 cells, with an approximately 18% increase comparted to the wild-type, whereas the Tetra-, Penta-, Hexa- and Hepta-mutants showed a relatively high level of infectivity, ranging from 5% to 16% increase at different concentrations of the pseudo-viruses in Huh-7 cells (Fig. 4C). Notably, these results are perfectly consistent with our previous molecular dynamics simulation data in that the F456Y-S494R-G496S mutation shows the lowest binding free energy (−85.79 kcal/mol). Moreover, the increase of the viral infectivity by addition of these pseudo-viral mutants shows a dose-dependent manner, and this infectivity-promoting effect could be amplified in multiple rounds of pseudo-viral infection (Fig. 4D). All of these virus infection data support the reliability of our previous *in silico* results and demonstrate that the selected multi-mutants exhibit enhanced viral infectivity.

3.5. Mutants confer structural flexibility to Spike protein

Following confirmation of the enhanced infectivity of the multi-mutants, we proceeded to investigate the difference in binding mode between the wild-type and each multi-mutant. The Spike structure with G446R-S494R mutations exhibits slight fluctuation due to strong salt-bridging interaction (R494-E35), resulting in increased accessibility for the mutant to bind with ACE2 (Fig. 5). The modeling of binding free energies yielded values of − 72.08 kcal/mol for tetra-mutant, − 74.64 kcal/mol for penta-mutant, − 70.41 kcal/mol for hexa-mutant and − 72.76 kcal/mol for hepta-mutant, which are consistent with our virus infection assay results showing a comparable cell infective ability for tetra-, penta-, hexa- and hepta-mutants. An in-depth assessment of the binding mode of these 4 multi-mutants revealed marked structural alterations near the mutation sites, with increased polar contacts observed at the 2019-nCoV-Spike/ACE2 binding interface. Moreover, salt-bridging bonds including R494-E35 and R449-D38 exist in all 6 selected mutated systems.

For the F456Y-S494R-G496S triple-mutant, substitution from G496 to S496 or S494 to R494 on the Spike introduces a larger side chain at the interface, which provides a more stable interaction with ACE2. The Glutamic acid residue at position D88 located at ACE2 can form the hydrogen bond with a serine residue at position S496 on Spike. Furthermore, a new salt-bridging bond is formed between E35 (ACE2) and the arginine residue at R494 on F456Y-S494R-G496S Spike mutant. These results explain how this triple-mutant is endowed with high viral infectivity.

It is noteworthy that among 184 multi-mutants, the mutation Y449R demonstrated the utmost frequency, with its binding free energy ranging from − 95.00 kcal/mol to − 65.00 kcal/mol. Conversely, it was the least common in the binding energy range of − 50.00 kcal/mol to − 30.00 kcal/mol. As for residue S494 in wild-type Spike protein, the

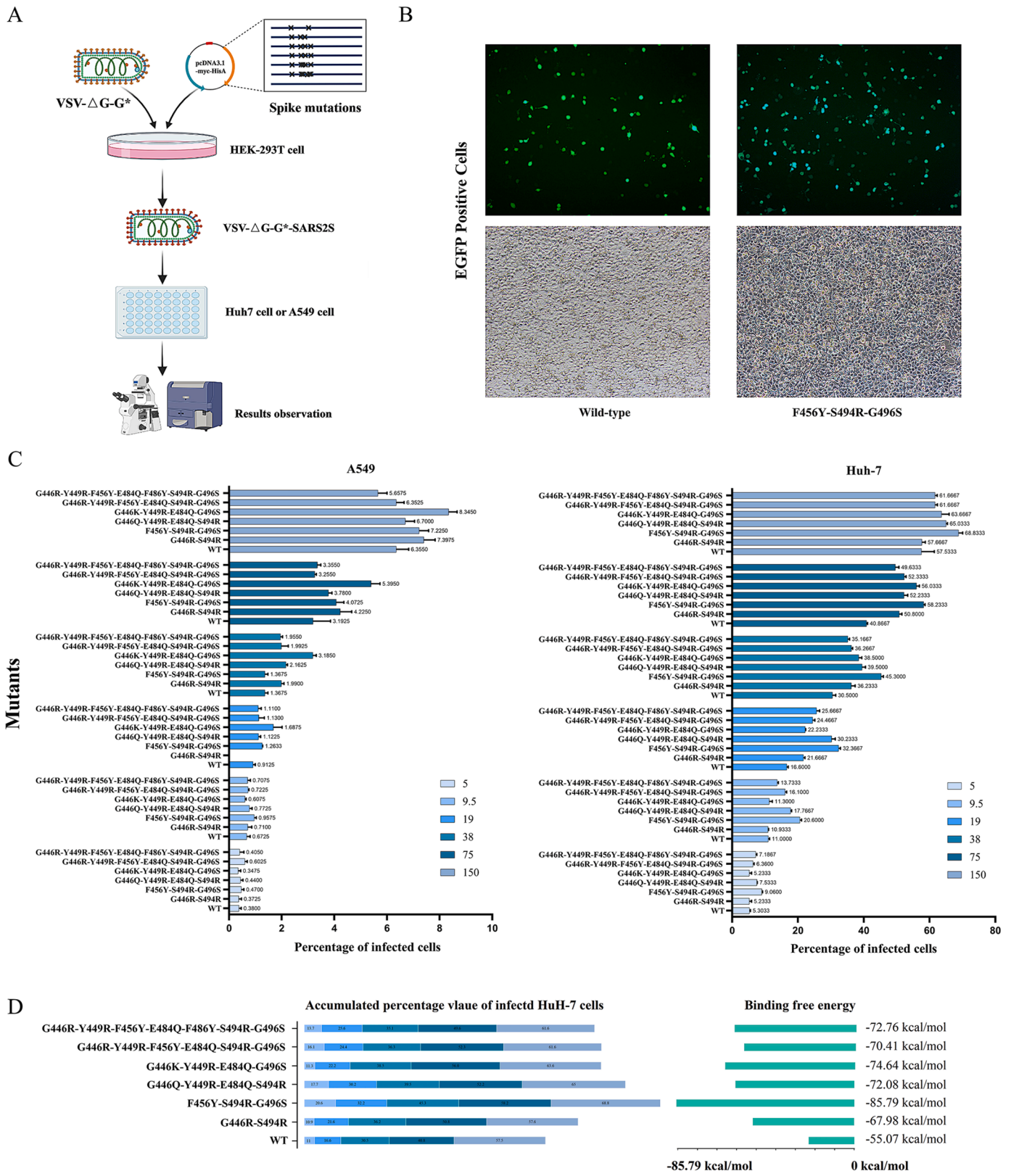

Fig. 4. Infectivity of multiple-mutant on 2019-nCoV-Spik in different doses. 2019-nCoV-Spike mediated cell entry is highly dependent on virus concentrations. Data from two experiments are shown as a percentage of cell entry of SARS-S-pseudotyped viruses into Huh-7 and A549 cells expressing ACE2 only. Accumulated percentage values of infected Huh-7 cells were counted, and corresponding binding free energies for each multiple-mutant was used for comparing analysis.

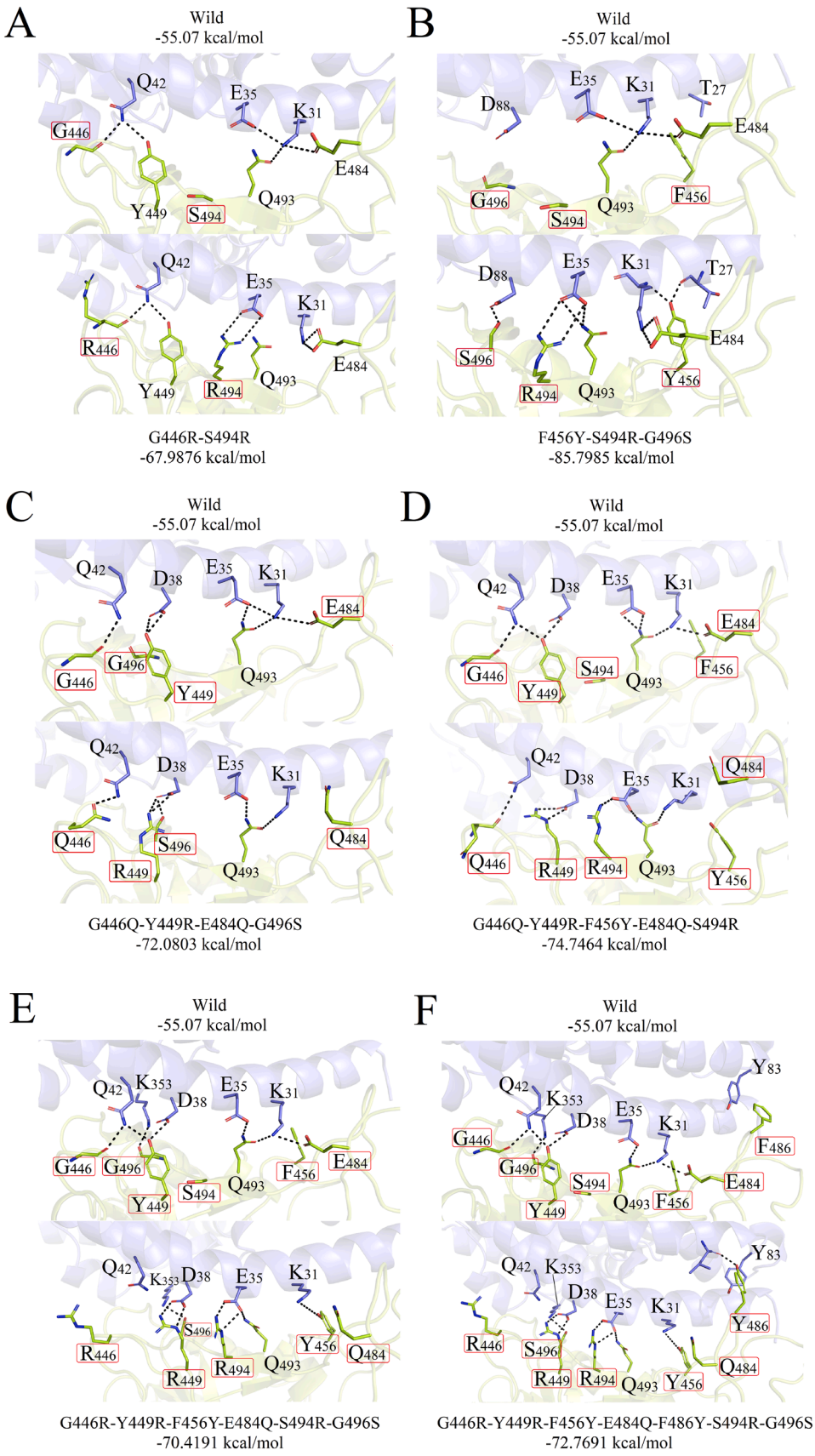

Fig. 5. Structural overlay of 2019-nCoV-Spike with ACE2 in different multiple-mutant complex and key residues improving binding affinity. Structural comparison of mutation sites between each multiple-mutant and wild-type complex. Spike is shown in shallow green color and Spike is in blue. Enlargement of the binding region within mutated sites, and key interacting residues were shown as sticks. Mutated site in each multiple-mutant was circled with red box. The figures were drawn with PyMOL software.

chemical group -C-OH of the side chain is not long enough to make any interaction with ACE2 receptor. Also, residue Y449 can only form weak hydrogen bonds with opposite residues Q42 and D38 (Fig. 6). Considering the binding characteristics of this region, structure of residues with opposite charge (positively charged R, negatively charged D/E) have experienced significantly swings as they are not matching the original physical and chemical characteristics in wild-type complex. Thus, mutations on residues of R449 and R494 could form new strong salt-bridge interactions with D38 and E35 on ACE2, resulting in significantly increased binding affinities.

4. Discussion

COVID-19 has escalated into the most severe global pandemic, and the emergence of 2019-nCoV variants carrying multi-mutations has presented unprecedented challenges for humankind. This study, initially identified 31 amino acid sites in the 2019-nCoV-Spike protein and subsequently calculated the binding free energies of the 589 single-mutants derived from these 31 sites through molecular dynamics simulations. Our findings show that about 47% of the 589 single-mutants could establish stronger interactions with ACE2 protein compared to the wild-type Spike (<−55.06 kcal/mol), and 9.1% of the 589 single-mutants displayed a binding energy value of < −60.00 kcal/mol. We further selected 8 single-mutants with the binding energy < −65.00 kcal/mol, including G446R, G446Q, Y449R, F456Y, E484N, F486Y, S494R and G496S, and obtained 184 multi-mutants by random compositions of these 8 single-mutants, which resulted in 28 double-, 50 triple-, 55 tetra-, 36 penta-, 13 hexa- and 2 hepta-mutants. Upon evaluating the binding behavior of these 184 multi-mutants, we discovered that approximately 54% of them possessed enhanced binding capabilities with ACE2 receptor. It is noteworthy that the double-, triple-, tetra-, penta-, hexa-, and hepta-mutants exhibit lower binding energies with ACE2 (<−55.07 kcal/mol) compared to the wild-type Spike, with percentages of the mutants displaying lower binding energies are 64%, 66%, 76%, 89%, 100%, and 100%, respectively, thereby showing a favorable influence of the mutant sites on ACE2 binding affinity. Furthermore, comparison of the binding energies for all 589 single-mutants and 184 multiple-mutants indicated that residues Y449R and S494R may have promoting effects to other mutants on the enhancement of binding affinity to ACE2 receptor.

To verify the reliability of the enhanced binding affinity of the multi-mutants predicted by *in silico* simulations, we randomly designated one multi-mutant for each mutation type, through double- to hepta-mutants, consequently, identified 6 multi-mutants, which were then employed for conducting infection assays using pseudotyped viruses. In two cell lines, all six multiple-mutants exhibited increased infectivity rates when infected with different doses of pseudo-viruses. Besides being dose-dependence, a strong linear correlation was observed between the estimated binding free energies and viral infectivity for these 6 multiple-mutants, demonstrating the high reliability of our *in silico* simulation results, and inferring that the 60 putative multi-mutants with lower binding energies (<−65.00 kcal/mol) may possess robust infective capabilities towards the host. The *in silico* and in vitro results have shown a close correlation between the 2019-nCoV-Spike mutants and the viral infectivity, hinting that these putative multi-mutants with high viral infectivity should receive more attention during future surveillance of 2019-nCoV. Hopefully, our data may provide novel insights into the in-depth understanding of the Spike mutants to aid in the development of therapeutics for COVID-19.

Funding

This work was supported in part by the National Natural Science Foundation of China (No. 61832019 and No.32070528), Hunan Provincial Science and Technology Program (No. 2021RC4008) and the key program of Chinese Academy of Sciences (KJZD-SW-L11)This work is supported in part by the High Performance Computing Center of Central South University.

Declaration of Competing Interest

All the authors (Senbiao Fang, Chuqi Lei, Meng Li, Yongfan Ming,

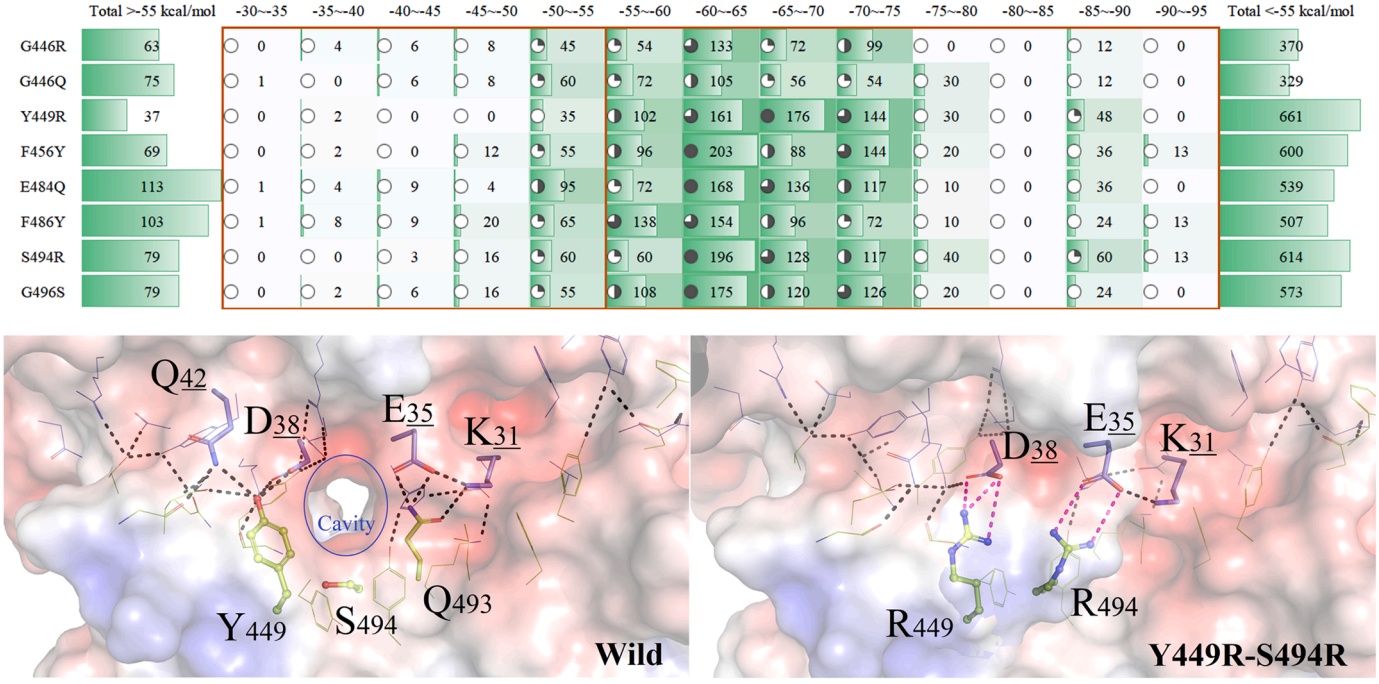

Fig. 6. Heatmap of single mutations that can improve the binding affinities in multiple mutants. Heatmap of single mutations that can improve the binding affinities in multiple mutants. The presence of hydrogen bonds was analyzed based on a geometry criterion where a given H-bond was considered to be formed or disrupted at a cutoff distance of 3.6 Å between the donor and acceptor atom and a cutoff bond angle of 120°.

Liren Liu, Xuming Zhou and Min Li) declare that they have no competing interests.

Data availability

The data underlying this manuscript are available in the article and in its online supplementary material.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2023.10.030.

References

[1] Xu Z, Shi L, Wang Y, Zhang Z, Huang L, Zhang C, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. Lancet Respir Med 2020;8:420–2.
[2] Madhi SA, Baillie V, Cutland CL, Voysey M, Koen AL, Fairlie L, et al. Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 variant. N Engl J Med 2021;384:1885–98.
[3] Ramasamy MN, Minassian AM, Ewer KJ, Flaxman AL, Folegatti PM, Owens DR, et al. Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial. Lancet 2020;396:1979–93.
[4] Xia S, Duan K, Zhang Y, Zhao D, Zhang H, Xie Z, et al. Effect of an inactivated vaccine against SARS-CoV-2 on safety and immunogenicity outcomes: interim analysis of 2 randomized clinical trials. JAMA 2020;324:951–60.
[5] Zhang Y, Zeng G, Pan H, Li C, Hu Y, Chu K, et al. Safety, tolerability, and immunogenicity of an inactivated SARS-CoV-2 vaccine in healthy adults aged 18-59 years: a randomised, double-blind, placebo-controlled, phase 1/2 clinical trial. Lancet Infect Dis 2021;21:181–92.
[6] Ledford H. COVID antiviral pills: what scientists still want to know. Nature 2021;599:358–9.
[7] Willyard Cassandra W. How antiviral pill molnupiravir shot ahead in the COVID drug hunt. Oct 8 Nature 2021. https://doi.org/10.1038/d41586-021-02783-1.
[8] Hou YJ, Chiba S, Halfmann P, Ehre C, Kuroda M, Dinnon 3rd KH, et al. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. Science 2020;370:1464–8.
[9] Supasa P, Zhou D, Dejnirattisai W, Liu C, Mentzer AJ, Ginn HM, et al. Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. Cell 2021;184:2201–11.
[10] Zhang L, Cui Z, Li Q, Wang B, Yu Y, Wu J, et al. Ten emerging SARS-CoV-2 spike variants exhibit variable infectivity, animal tropism, and antibody neutralization. Commun Biol 2021;4:1196.
[11] Shorthouse D, Hall BA. SARS-CoV-2 variants are selecting for Spike protein mutations that increase protein stability. J Chem Inf Model 2021;61:4152–5.
[12] Zhou Z, Du P, Yu M, Baptista-Hon DT, Miao M, Xiang AP, et al. Assessment of infectivity and the impact on neutralizing activity of immune sera of the COVID-19 variant, CAL.20C. Signal Transduct Target Ther 2021;6:285.
[13] Kirby T. New variant of SARS-CoV-2 in UK causes surge of COVID-19. Lancet Respir Med 2021;9:20–1.
[14] Starr TN, Greaney AJ, Addetia A, Hannon WW, Choudhary MC, Dingens AS, et al. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. Science 2021;371:850–4.
[15] Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 2020;581:215–20.
[16] Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol 2021;19:409–24.
[17] Lv Z, Deng YQ, Ye Q, Cao L, Sun C, Fan C, et al. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. Science 2020;18:1505–9.
[18] Gong Z, Zhu JW, Li CP, Jiang S, Ma L, Tang B, et al. An online coronavirus analysis platform from the National Genomics Data Center. Zool Res 2020;41:705–8.
[19] Fang S, Zheng R, Lei C, Wang J, Zheng R, Li M. Key residues influencing binding affinities of 2019-nCoV with ACE2 in different species. Brief Bioinform 2020;22:963–75.
[20] Fang S, Zheng R, Lei C, Wang J, Zhou R, Li M. In silico prediction of new mutations that can improve the binding abilities between 2019-nCoV coronavirus and human ACE2. IEEE/ACM Trans Comput Biol Bioinform 2021;19:1694–702.
[21] Pettersen EF. UCSF Chimera-a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605–12.
[22] Zhu K, Day T, Warshaviak D, Murrett C, Friesner R, Pearlman D. Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. Proteins 2014;82:1646–55.
[23] Götz AW, Williamson MJ, Xu D, Poole D, Grand SL, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. Generalized Born. J Chem Theory Comput 2012;8:1542–55.
[24] Wu F, Zhao S, Yu B, Chen Y, Wang W, Song Z, et al. A new coronavirus associated with human respiratory disease in China. Nature 2020;579:265–9.
[25] Nie J, Li Q, Wu J, Zhao C, Hao H, Liu H, et al. Establishment and validation of a pseudovirus neutralization assay for SARS-CoV-2. Emerg Microbes Infect 2020;9:680–6.
[26] Reed LJ, Muench H. A simple method of estimating fifty-percent endpoints. Am J Epidemiol 1938;27:493–7.