

# Interaction Analysis of the Spike Protein of Delta and Omicron Variants of SARS-CoV-2 with hACE2 and Eight Monoclonal Antibodies Using the Fragment Molecular Orbital Method

Sungbo Hwang, Seung-Hwa Baek, and Daeui Park\*



Cite This: *J. Chem. Inf. Model.* 2022, 62, 1771–1782



Read Online

ACCESS |



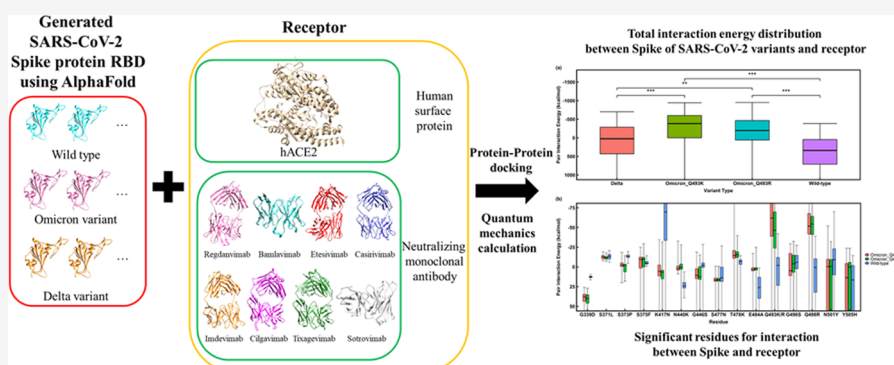
Metrics & More



Article Recommendations



Supporting Information



**ABSTRACT:** In the past 2 years, since the emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), multiple SARS-CoV-2 variants have emerged. Whenever a new variant emerges, considerable time is required to analyze the binding affinity of the viral surface proteins to human angiotensin-converting enzyme 2 (hACE2) and monoclonal antibodies. To efficiently predict the binding affinities associated with hACE2 and monoclonal antibodies in a short time, herein, we propose a method applying statistical analysis to simulations performed using molecular and quantum mechanics. This method efficiently predicted the trend of binding affinity for the binding of the spike protein of each variant of SARS-CoV-2 to hACE2 and individually to eight commercial monoclonal antibodies. Additionally, this method accurately predicted interaction energy changes in the crystal structure for 10 of 13 mutated residues in the Omicron variant, showing a significant change in the interaction energy of hACE2. S375F was found to be a mutation that majorly changed the binding affinity of the spike protein to hACE2 and the eight monoclonal antibodies. Our proposed analysis method enables the prediction of the binding affinity of new variants to hACE2 or to monoclonal antibodies in a shorter time compared to that utilized by the experimental method.

## 1. INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) first emerged in China and has rapidly propagated, leading to a pandemic.<sup>1</sup> The SARS-CoV-2 genome encodes various proteins referred to as spike (S), envelope (E), membrane (M), and nucleocapsid (N). Additionally, six open reading frames have been predicted to correspond to hypothetical proteins with no known function.<sup>2</sup> Each known protein has a distinct function. In particular, the S protein mediates viral invasion into host cells via human angiotensin-converting enzyme 2 (hACE2), and the E protein is involved in the formation and release of viral particles.<sup>3</sup> The S protein, owing to its presence on the surface of the virus, plays a key role in mediating recognition of the host receptor, hACE2, and viral entry through the cell membrane. The S protein is composed of two subunits, S1 and S2.<sup>4</sup> The S1 subunit contains a receptor binding domain (RBD) that recognizes and binds to the host receptor, whereas the S2 subunit mediates viral

membrane fusion.<sup>5</sup> Therefore, it is imperative to investigate how S protein mutations affect viral transmission and infection.

In the past 2 years, since the emergence of the pandemic, the number of cumulative deaths from coronavirus disease (COVID-19) caused by SARS-CoV-2 has surpassed five million.<sup>6</sup> Along with the large death toll, many variants of SARS-CoV-2 have emerged. The first new SARS-CoV-2 variants detected were announced in December 2020 by the COVID-19 Genomics U.K. Consortium and were named  $\alpha$  and  $\beta$  by the World Health Organization (WHO). These variants had 17 mutations and 70% greater transmission ability

Received: January 26, 2022

Published: March 21, 2022



than the wild type.<sup>7</sup> The S protein mutations in the  $\alpha$  variant are  $\Delta$ H69/ $\Delta$ V70,  $\Delta$ 144/144, N501Y, A570D, D614G, P681H, T716I, S982A, and D1118H. The N501Y mutation being located in the RBD causes the S protein to have increased binding affinity to hACE2 and a consequent increase in the transmission ability of the virus.<sup>8</sup> The  $\alpha$  and  $\beta$  variants thus became the dominant variants in most countries, and many researchers consequently began to monitor for new variants. In March 2021, another variant named Delta, which had 20 mutations, was reported for the first time. The Delta variant rapidly became dominant over the  $\alpha$  and  $\beta$  variants. The mutations of the S protein in the Delta variant are T19R, E156G,  $\Delta$ 157/158, L452R, T478K, D614G, P681R, and D950N. The L452R and T478K mutations are located in the RBD, and they significantly increase the binding affinity of the S protein to hACE2.<sup>9</sup> Following the emergence of the Delta variant, a new SARS-CoV-2 variant, called Omicron, was first reported by the WHO on November 25, 2021, in South Africa. Recently, there has been a marked increase in the number of Omicron variant breakthrough infections. The Omicron variant is a highly divergent variant having 32 mutations, including substitutions, deletions, and an insertion in the S protein, half of which reside within the RBD. Many of the mutations in the Omicron variant have the potential to increase the infectivity of the virus and confer resistance to therapeutic agents, which are factors that may be associated with immune escape and high transmissibility. Importantly, as all mutations of the Omicron variant are still not precisely known, there is considerable uncertainty regarding the effects of the complete combination of deletions and mutations on the viral behavior and susceptibility to human immunity.

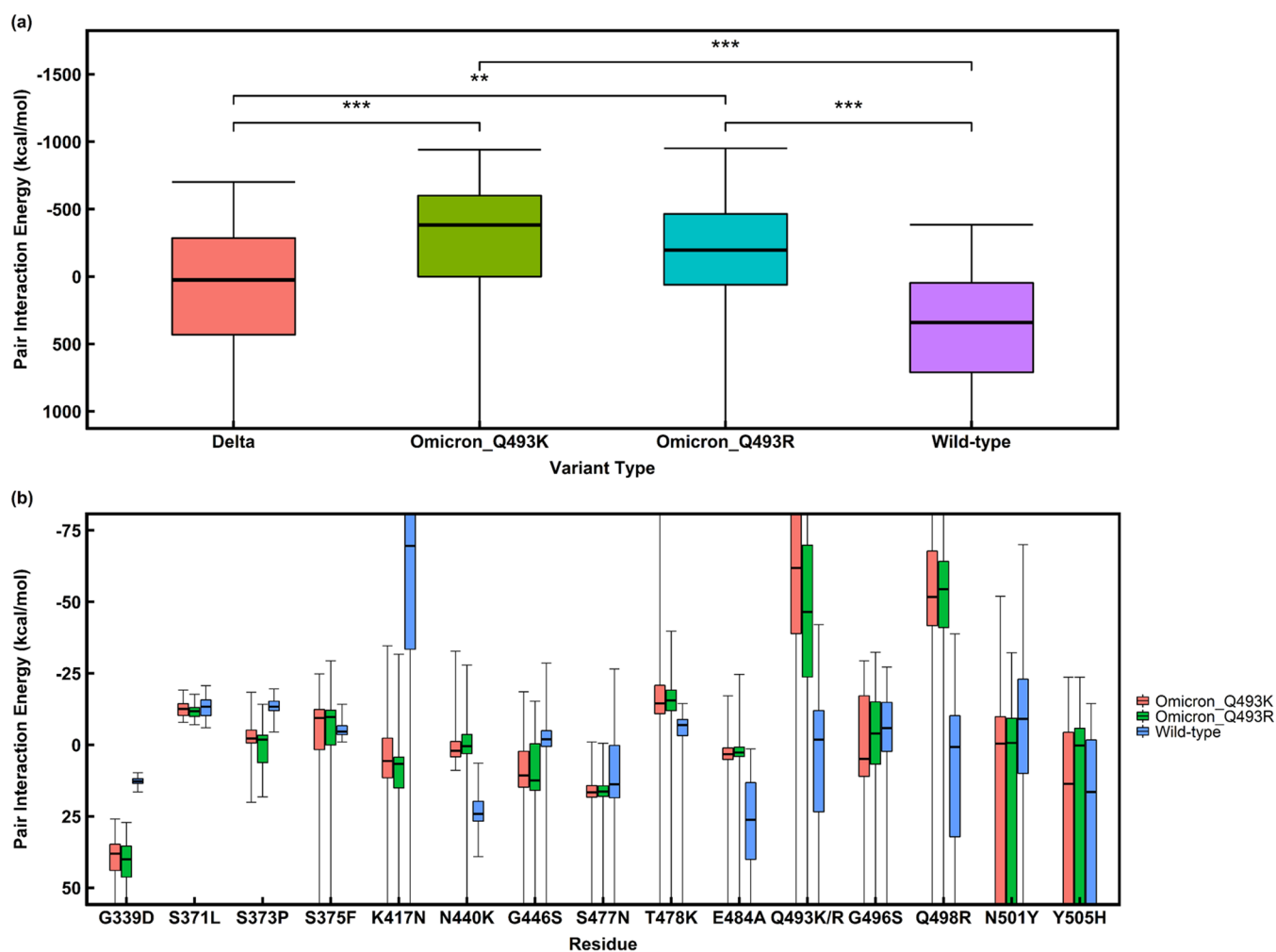
Several studies have been conducted to predict the binding force between SARS-CoV-2 RBD and hACE2 to determine the propagation characteristics of the Omicron variant. Based on docking studies with single-mutant S proteins, the Q493K, N501Y, S371L, S373P, S375F, Q498R, and T478K mutations have been reported to contribute significantly to the high binding affinity of the Omicron S protein to hACE2.<sup>10</sup> This study has revealed that compared to the proteins of the Delta variant, the complete S protein and RBD of the Omicron variant possess a high proportion of hydrophobic amino acids, such as leucine and phenylalanine, which are located within the protein core and are required for its structural stability. However, the exact structural changes caused by the combination of two or more mutations have not been accounted for by this approach. Another study has reported the generation of RBD-hACE2 structures optimized using the AMBER FF14SB force field for S proteins of Delta and Omicron variants, wherein the steric hindrances were calculated based on an ab initio quantum mechanical (QM) model called the fragment bond order.<sup>11</sup> Although the interaction energy of each residue in the RBD or hACE2 was more accurate in the QM model used in this approach than in the molecular mechanical (MM) model used in docking studies, the overall accuracy of the generated structures was dependent on the accuracy of the force field. In addition, the RBD-hACE2 complex structure does not have a fixed position and orientation but moves continuously and changes form to drive the structure toward a state of global optimum energy. The protein–protein binding structure of the RBD-hACE2 complex being a markedly large system makes the approximation of this multitude of structural states to the global optimal energy state rather difficult.

Here, we propose a novel binding affinity prediction method that can facilitate the understanding of the emergence of new variants having unknown structures as well as the existence of various binding poses of S proteins in predicted unknown structures. Three-dimensional structures of proteins were generated using AlphaFold to construct new variants having unknown structures.<sup>12</sup> We obtained conformational variation of binding poses by performing protein–protein docking simulations with constraints on the distance between the residues of the RBD and receptor. These simulations generated structural information of all binding poses located close to the binding site that can be determined by energy calculations. The binding affinity for all generated binding poses was predicted using a QM model called the fragment molecular orbital (FMO).<sup>13</sup> FMO-based calculations have been applied to various large biomolecular systems and used in analyses of hot spot residues between hACE2 and S proteins.<sup>14,15</sup> Our research goal was to use statistical analyses to identify significant differences among the interaction energies of all binding poses as well as the most stable RBD-hACE2 complex structures for each variant. In addition, the difference in interaction energy between the hACE2 residues and S protein residues of the wild type as well as the Omicron variant was analyzed for all binding poses.

## 2. MATERIALS AND METHODS

**2.1. Preparing Protein Structures.** The omicron variant has 15 mutations in the RBD domain of the S protein: G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493K/R, G496S, Q498R, N501Y, and Y505H. In previous studies, viral particles of the Omicron variant were further subclassified into two types: those containing the Q493K and Q493R mutations. The structures of complexes formed by the binding of hACE2 to the RBD of the wild type (PDB ID:6M0J),<sup>16</sup> Delta variant (PDB ID:7V8B),<sup>17</sup> and Omicron variant containing the Q493R mutation (PDB ID:7T9L),<sup>18</sup> and individually to eight different monoclonal antibodies (Table S1), were downloaded from the Protein Data Bank (<https://www.rcsb.org/>). The RBD structure comprised a substructure of the S protein between the Thr333 and Gly526 residues. The amino acid sequences of RBDs required to predict the RBD structures of the Omicron variant by means of AlphaFold were generated by changing the mutant sequence for each mutant using data and tools from the online resource GISAID.<sup>19,20</sup> The RBD structures of the Delta and Omicron variants were generated based on 10 structural models prepared using AlphaFold. The predicted RBD structures of the wild-type and Delta variants were generated based on the sequence of each type and compared with the experimental structures. The complete database provided by AlphaFold was used for structural predictions and analysis. The maximum template release date was defined as 2020-05-14. The RBD structures generated using AlphaFold had hydrogen atoms bound to heavy atoms, and the hydrogen atoms in hACE2 and commercial antibody structures were added using Chimera 1.14.<sup>21</sup> The root-mean-square distance (RMSD) was also calculated using Chimera 1.14.

**2.2. Protein–Protein Docking Simulation.** To generate structures of complexes formed by the predicted RBDs and hACE2, protein–protein docking simulations were performed using PIPER.<sup>22</sup> PIPER uses the method of rigid docking simulations, which is a docking algorithm that does not involve changing the RBD domain structure, and has a low calculation



**Figure 1.** Distribution of pair interaction energy (a) between hACE2 and four types of spike proteins and (b) between each residue of hACE2 and the spike protein of the wild type and the two Omicron variants. Asterisks indicate statistically significant variation (Student's *t*-test, \*\**P*-value < 0.01, \*\*\**P*-value < 0.001).

cost when using Fourier transformations. The binding site was defined as the constrained distance between the residues of the RBD and those of the receptors, namely, hACE2 and the eight monoclonal antibodies. Residue pair sets were defined as the residues of the RBD and the receptor having a distance less than 5.0 Å between them. Changes in binding poses were analyzed by considering binding within a range of 5.0 Å away from the region defined by the calculated distance between the RBD and the receptor in the crystal structure of the complex. Structures of complexes with atoms close to less than the van der Waals distance were filtered out to eliminate complex structures having excessive proximity between the RBD and the receptor. Up to 10 binding poses for each receptor and RBD type were selected based on the docking score.

**2.3. Molecular Dynamics.** Considering the flexibility of hACE2, molecular dynamics simulations were performed to generate binding poses involving hACE2 binding to the three predicted types of RBD structures. The initial binding poses were generated by superimposing the predicted RBD structure of each variant on the experimentally obtained hACE2-RBD crystal structure of the wild type. One predicted RBD structure of the wild type overlapped with hACE2 and hence was excluded from the molecular dynamics simulations. Molecular dynamics simulations were performed by means of NAMD

version 2.12<sup>23</sup> using the CHARMM36m force field.<sup>24</sup> The files of the predicted complex structures were converted from the Protein Data Bank format (.pdb) to the structure (.psf) and coordinate file (.pdb) formats using CHARMM-GUI.<sup>25,26</sup> The water box was prepared using visual molecular dynamics (VMD) version 1.9.4.<sup>27</sup> Energy minimization for the superimposed structure of the Omicron variant was performed for 500 ps, with a time step of 1 fs. The simulation was performed with a constant temperature set at 310 K. Energy minimization for the superimposed structure of the RBDs was performed for a maximum of 60 ns. Complexes categorized as top 10 energy minimum structures were defined as candidate complex structures. The RMSD of the structure of each complex was defined in terms of its alignment to the initial complex structure.

**2.4. FMO Calculation and Pair Interaction Energy Decomposition Analysis (PIEDA).** To calculate the interaction energy between RBDs and the receptors, FMO calculations were performed using the predicted structures of complexes via protein–protein docking simulations or molecular dynamics simulations. FMO-based calculations are reported to yield predicted values having a high correlation with the experimental data.<sup>28</sup> All FMO calculations in this study were performed using the FMO-DFTB3 method.<sup>29</sup> All

input files were prepared in compliance with the hybrid orbital projection fragmentation scheme.<sup>30</sup> One fragment was defined as two cysteine residues that were within a distance of 2.15 Å from each other and were composed of disulfide bonds. The polarizable continuum model was used considering that the binding of the RBDs to the receptors occurs in a solution state.<sup>31</sup> The interaction energy between the RBD and receptor is defined by the following equation

$$\Delta E_{\text{inter}} = \Delta E_{\text{RBD-receptor}} - (\Delta E_{\text{RBD}} + \Delta E_{\text{receptor}})$$

where  $\Delta E_{\text{inter}}$  and  $\Delta E_{\text{RBD-receptor}}$  are the pair interaction energy (PIE) between the RBD and receptor and the total interaction energy for the complex structure, respectively.  $\Delta E_{\text{RBD}}$  and  $\Delta E_{\text{receptor}}$  are the total interaction energies for the RBD and receptor, respectively. The binding affinity was not predicted solely based on the minimum energy structure due to the flexibility of the complex structures. Instead, it was defined as the median value of PIE of all structures of each variant obtained using protein–protein docking simulations or molecular dynamics. When performing FMO calculation, the binding pose between the two monomer proteins with a very close distance cannot be calculated as energy divergence because of some very large repulsive interactions between two proteins. In this case, the total interaction energy will be a highly positive total interaction energy compared to other bonding structures. The average value cannot properly reflect the distribution of most normal binding poses of the very close distance. Therefore, the PIE distribution of the binding poses as a single statistical value, and the median value was considered more appropriate than the average value.

The PIE distribution of all binding poses of RBD with the receptors was compared using the Kruskal–Wallis rank-sum test to obtain a statistical difference in the median values of PIE among all variants. In addition, the Wilcoxon rank-sum test was used to compare the median of interaction energy of all binding poses of variants to obtain the statistical differences in values of PIE between the residues of the wild type and the Omicron variant by comparing the median of interaction energy of all binding poses of variants.

The statistical significance of binding affinity was described as the *P*-value calculated by the Kruskal–Wallis rank-sum test or the Wilcoxon rank-sum test. The *P*-value indicates the statistical significance that two distributions of binding affinity between the wild type and variant are the same or not. Since a smaller *P*-value (*P* < 0.05) can indicate a high probability that the two distributions are different, the binding affinity can be analyzed for all possible binding poses of the S protein with receptors not specific to one binding pose. We determined that the difference of median PIEs between the wild type and variant of the S protein has statistical significance when the *P*-value is less than 0.05. In addition, we determined whether the median PIE for the 15 mutant residues of the Omicron variant was significantly based on the *P*-value between the wild type and the Omicron variant. After selecting the residues with a significant median PIE difference based on the *P*-value, the residue with a strong interaction energy for hACE2 and the residues with a weak interaction energy for monoclonal antibodies were selected and analyzed.

To analyze interaction types of the residue of the S protein, we performed pair interaction energy decomposition analysis (PIEDA) with FMO calculation. In PIEDA, the pair interaction energy of a set of residue pairs consisting of S protein residues and hACE2 receptor residues could be

analyzed by dividing them into five interaction energy types. The five interaction energy types are electrostatic energy, exchange energy (also called steric repulsion), charge transfer energy, dispersion energy, and solvation energy. To determine the significantly different interaction energy types in 13 mutation positions of the Omicron variant, the same method as the previous PIE distribution comparison was used.

FMO calculation and PIEDA were performed using the version June 30, 2020, R1 GAMESS.<sup>32</sup> Statistical analysis for PIE was performed using R, version 4.1.0. PIE distribution plots were generated using the ggplot2 library in R, version 4.1.0.

### 3. RESULTS AND DISCUSSION

#### 3.1. Predicted RBD Structures of the Wild-Type, Delta, and Omicron Variants. Assuming the RBD structure

**Table 1. Median Interaction Energy of Residues in the Receptor Binding Domain Interacting with hACE2 Obtained Using Protein–Protein Docking Simulations<sup>a</sup>**

residue	energy with wild type (kcal/mol)	Omicron (Q493K)		Omicron (Q493R)	
		energy	<i>P</i> -value	energy	<i>P</i> -value
G339D	12.7	38.1	$5.4 \times 10^{-32}$	40.0	$4.3 \times 10^{-31}$
S371L <sup>b</sup>	<u>-13.3</u>	-12.5	$2.6 \times 10^{-1}$	<u>-11.8</u>	$2.6 \times 10^{-3}$
S373P	-13.4	-2.2	$6.3 \times 10^{-28}$	-1.9	$5.0 \times 10^{-29}$
<u>S375F</u>	<u>-4.6</u>	<u>-9.4</u>	$5.1 \times 10^{-6}$	<u>-9.7</u>	$6.8 \times 10^{-4}$
K417N	-69.6	5.6	$6.8 \times 10^{-22}$	6.7	$1.9 \times 10^{-22}$
<u>N440K</u>	<u>24.1</u>	<u>2.1</u>	$1.6 \times 10^{-31}$	<u>0.4</u>	$1.5 \times 10^{-29}$
G446S	-2.0	10.7	$1.9 \times 10^{-15}$	12.4	$1.5 \times 10^{-12}$
S477N	13.8	16.6	$1.1 \times 10^{-3}$	16.3	$1.5 \times 10^{-3}$
<u>T478K</u>	<u>-6.9</u>	<u>-14.5</u>	$1.1 \times 10^{-18}$	<u>-15.5</u>	$2.8 \times 10^{-22}$
<u>E484A</u>	<u>26.2</u>	<u>3.3</u>	$2.3 \times 10^{-28}$	<u>2.7</u>	$5.5 \times 10^{-27}$
<u>Q493K/R</u>	<u>-1.9</u>	<u>-61.9</u>	$3.4 \times 10^{-25}$	<u>-46.5</u>	$5.8 \times 10^{-14}$
G496S	-5.9	4.9	$9.0 \times 10^{-3}$	-4.0	$2.8 \times 10^{-1}$
<u>Q498R</u>	<u>0.7</u>	<u>-51.7</u>	$4.8 \times 10^{-21}$	<u>-54.5</u>	$5.8 \times 10^{-19}$
N501Y	-9.1	-0.4	$2.1 \times 10^{-4}$	-0.7	$2.2 \times 10^{-6}$
YS05H	16.4	13.6	$5.9 \times 10^{-1}$	0.3	$1.5 \times 10^{-1}$

<sup>a</sup>Underlined residues and values indicate significantly stronger interactions in the Omicron variants than those in the wild type for which the *P*-value was less than  $5.0 \times 10^{-2}$  and the median interaction energy for the Omicron variant was lower than that for the wild type.

<sup>b</sup>The Omicron variant containing Q493K showed no significantly stronger interaction, but the Omicron variant containing Q493R showed a significantly stronger interaction than that observed in the wild type.

of the Omicron variant to be unknown and a stable RBD structure generated among various possible flexible RBD structures, we used AlphaFold to generate 10 models of each of the four predicted types of RBD structures. The predicted structure superimposed on the crystal structure is described in Figure S1. The average RMSD with the crystal structure of each variant was 4.455 Å for the wild type and 3.655 Å for the Delta variant. The lowest values of RMSD among those obtained with the crystal structures were 2.996 Å for the wild type, 2.932 Å for the Delta variant, and 2.963 Å for the Omicron variant containing the Q493R mutation. The RMSD for the Omicron variant containing Q493K could not be calculated since Gln493 was substituted with a different

**Table 2. Pair Interaction Energy of Residues of hACE2 and Residues of the Receptor Binding Domain in Crystal Structures<sup>a</sup>**

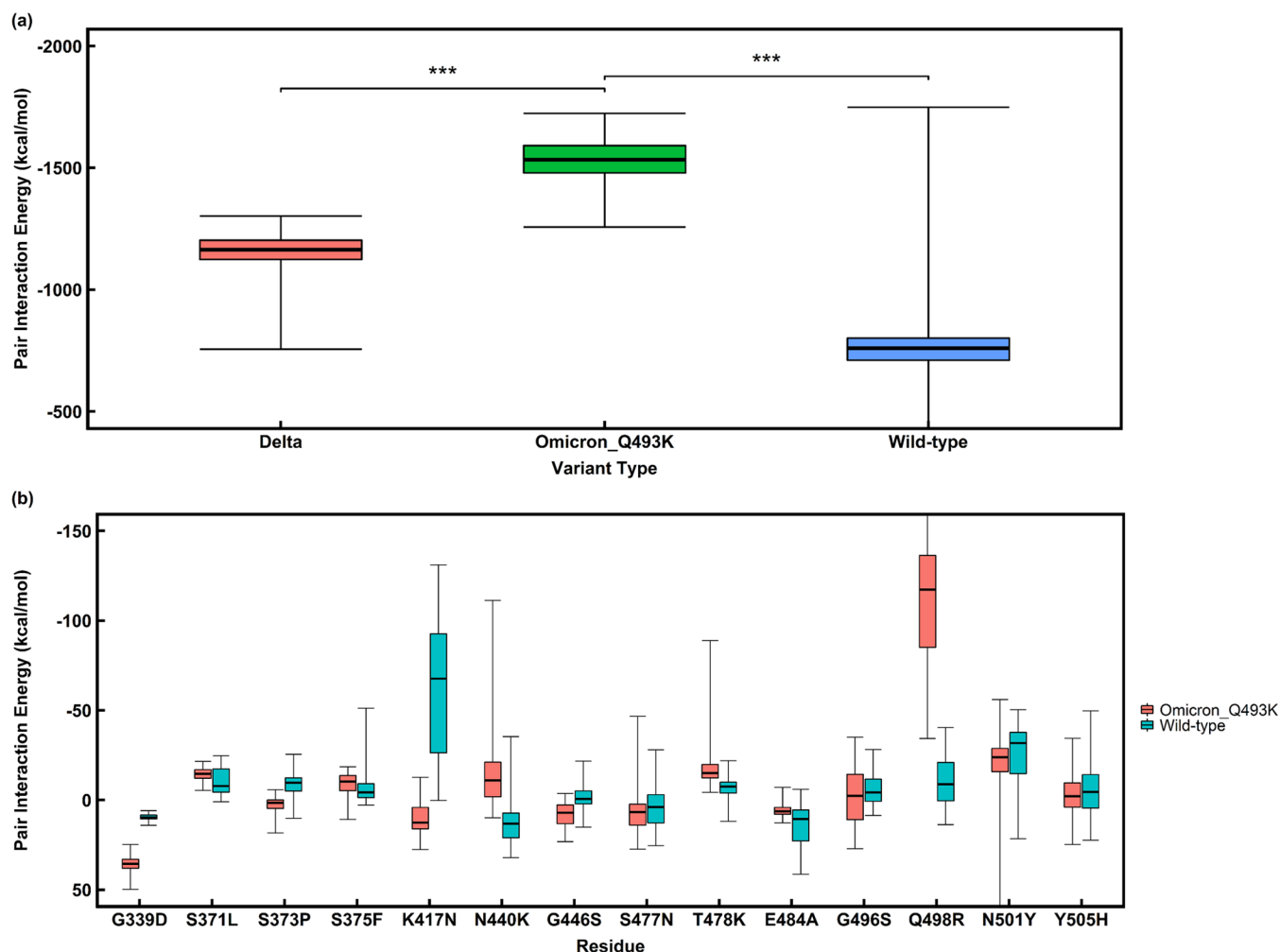
residue	pair interaction energy (kcal/mol) with hACE2	
	wild type	Omicron (Q493R)
G339D	15.9	36.0
S371L	-21.9	-12.0
S373P	-11.2	4.9
<u>S375F</u>	<u>-4.8</u>	<u>-7.9</u>
K417N	-135.8	7.9
<u>N440K</u>	<u>21.0</u>	<u>-7.0</u>
G446S	-5.6	18.7
<u>S477N</u>	<u>15.3</u>	<u>14.3</u>
<u>T478K</u>	<u>-5.1</u>	<u>-20.8</u>
<u>E484A</u>	<u>39.9</u>	<u>4.1</u>
<u>Q493R</u>	<u>-5.2</u>	<u>-125.0</u>
<u>G496S</u>	<u>1.2</u>	<u>-16.5</u>
<u>Q498R</u>	<u>-16.8</u>	<u>-170.0</u>
<u>N501Y</u>	<u>-11.6</u>	<u>-13.5</u>
YS05H	4.0	6.8

<sup>a</sup>Underlined residues and values represent stronger interactions in the Omicron variant than those in the wild type.

residue in the structure generated using AlphaFold than that present in the original crystal structure.

The loop structure formed by residues between Leu455 and Phe490 has various conformations, but the other structures of the RBD are the same in the wild type and the Delta variant. The average RMSD for the RBD structures was 4.439 Å for the wild type, 2.865 Å for the Delta variant, and 2.794 Å for the Omicron variant. The Omicron variant had the least structural deviation among the three types of predicted RBD structures. The lowest RMSDs of the three types of RBDs were 2.292 Å for the wild type, 1.701 Å for the Delta variant, and 2.505 Å for the Omicron variant containing the Q493R mutation. The highest RMSDs of the three types of RBDs were 7.852 Å for the wild type, 7.488 Å for the Delta variant, and 3.253 Å for the Omicron variant containing the Q493R mutation. As structure prediction using AlphaFold included the structure relaxation process, the structure corresponding to the energy minimum could be predicted. The RBD structure of the Omicron variant was relatively rigid compared to the other RBD structures.

**3.2. Prediction of the Binding Affinity between RBD and hACE2.** Molecular dynamics is a popular in silico method for predicting the binding affinity and analyzing the interaction energy. However, the use of molecular dynamics to predict binding affinity in massive protein–protein complex structures



**Figure 2.** Distribution of pair interaction energy (a) between hACE2 and three types of spike proteins and (b) between each residue of hACE2 and the spike protein of the wild type and Omicron variant. Asterisks indicate statistically significant variation (Student's *t*-test, \*\*\**P*-value < 0.001).

**Table 3. Median Interaction Energy of Residues in the Receptor Binding Domain Interacting with hACE2 Generated Using Molecular Dynamics<sup>ab</sup>**

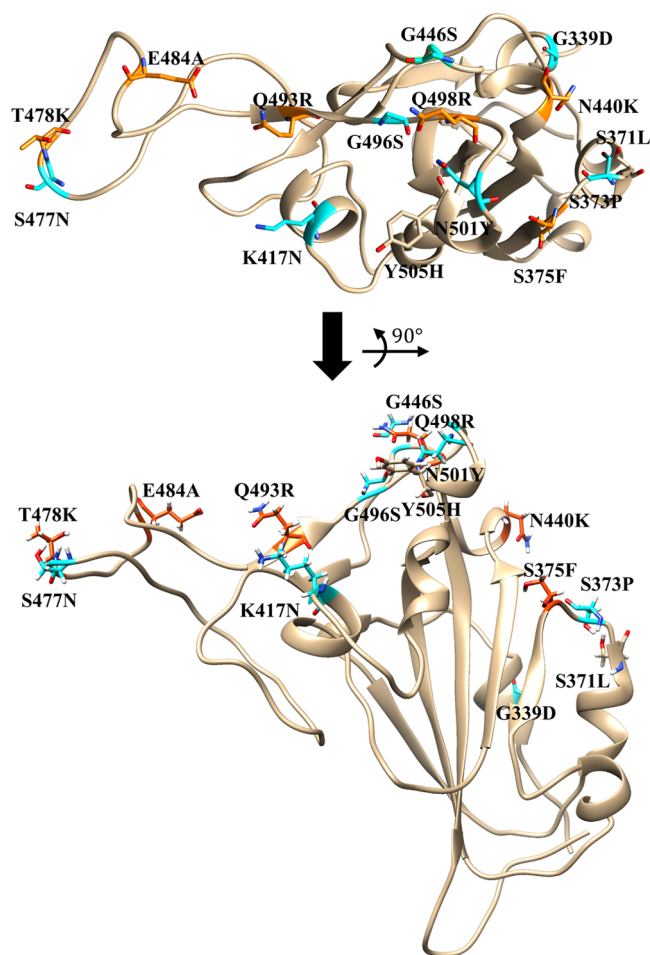
residue	median interaction energy (kcal/mol) with hACE2		P-value
	wild type	Omicron (Q493R)	
G339D	9.7	35.5	$3.3 \times 10^{-29}$
<u>S371L</u>	<u>-7.8</u>	<u>-14.6</u>	$1.6 \times 10^{-5}$
S373P	-9.6	1.5	$9.6 \times 10^{-21}$
<b>S375F</b>	<b>-4.3</b>	<b>-10.3</b>	$4.1 \times 10^{-5}$
K417N	-67.7	12.6	$9.6 \times 10^{-29}$
<b>N440K</b>	<b>13.2</b>	<b>-10.9</b>	$7.2 \times 10^{-22}$
G446S	-0.6	7.0	$2.8 \times 10^{-15}$
S477N	3.9	6.7	$4.9 \times 10^{-2}$
<b>T478K</b>	<b>-7.4</b>	<b>-15.0</b>	$5.7 \times 10^{-18}$
<b>E484R</b>	<b>10.5</b>	<b>6.3</b>	$6.1 \times 10^{-8}$
<b>Q493K</b>	<b>-30.0</b>	<b>-230.0</b>	$3.3 \times 10^{-29}$
G496S	-4.2	-2.4	$1.7 \times 10^{-1}$
<b>Q498R</b>	<b>-8.8</b>	<b>-117.4</b>	$4.7 \times 10^{-22}$
N501Y	-31.7	-23.9	$9.3 \times 10^{-4}$
Y505H	-4.6	-2.1	$3.1 \times 10^{-1}$

<sup>a</sup>Underlined residues and values indicate significantly stronger interactions in the Omicron variant than those in the wild type for which the *P*-value was less than  $5.0 \times 10^{-2}$ , and the median interaction energy of the Omicron variant was lower than that of the wild type. <sup>b</sup>Bold residues and values represent residues having significantly stronger interactions in protein–protein docking simulations for the Omicron variant than those of the wild type.

has a high cost, and the determination of the initial structure is difficult. The calculation cost involved in protein–protein docking simulations is relatively lower than that of molecular dynamics. We used protein–protein docking simulations as well as molecular dynamics for an *in silico* analysis of the S protein–receptor complex based on predictions of binding affinity between the hACE2 receptor and crystal structures for three mutations.

First, we performed protein–protein docking simulations to predict the binding affinity of different RBDs with hACE2. Our results showed 96 binding poses in the wild type, 98 in the Delta variant, 90 in the Omicron variant containing the Q493K mutation, and 85 in the Omicron variant containing the Q493R mutation. Among the binding poses that can be determined by docking simulations, a binding pose with a high probability of being located around the binding site of hACE2 was selected, and the distributions of PIE between the three types of RBDs and hACE2 were accordingly analyzed using FMO-based calculations (Figure 1a).

The median values of pair interaction energies were 340.0 kcal/mol for the wild type, 25.2 kcal/mol for the Delta variant, and -382.1 and -196.4 kcal/mol for the Omicron variants containing the Q493K and Q493R mutations, respectively. The *P*-values obtained by the Kruskal–Wallis rank-sum test were  $2.940 \times 10^{-15}$  and  $1.140 \times 10^{-11}$  for the Omicron variants containing the Q493K and Q493R mutations, respectively, indicating that the median interaction energies for each RBD type were significantly different. The median PIE of the Omicron variant was stronger than that of the other variants. In a previous study, the binding affinity between hACE2 and the RBD of the Delta variant ( $K_d = 4.6$  nM) was shown to be stronger than that observed in the wild type ( $K_d = 21.3$  nM).<sup>33</sup> Another study shows the binding affinity between



**Figure 3.** Mutation points of the spike protein of the Omicron variant depicted in the spike protein of the wild type. Cyan-colored residues represent residues in the Omicron variant with weaker interactions than those in the wild type. Orange-colored residues represent residues in the Omicron variant with stronger interactions than those in the wild type.

hACE2 and the RBD of the Omicron variant containing Q493K to be stronger than that of the wild type.<sup>34</sup> In our study, the predicted binding affinity of the RBD of the wild type and the Delta variant for hACE2 could be considered reliable due to the strong correlation between the predicted values and experimental data.

The interaction energy types between 13 mutation residues and hACE2 using PIE DA are described in Tables S2–S6. Among five interaction energy types, electrostatic energy was the major type of energy that the Omicron variant had a stronger pair interaction than the wild type. Among the seven mutation residues in the Omicron variant, six mutation residues in the Omicron variant except for S375F had a stronger electrostatic energy than those in the wild type. Only S375F in the Omicron variant had a stronger solvation energy than that in the wild type. Dispersion energy, known as the major interaction type in protein–protein interaction, was calculated to have a value between -1.5 and 0.0 kcal/mol in the wild type and -1.2 and 0.0 in the Omicron variant (Table S5).

The PIE in the crystal structures was -894.4 kcal/mol for the wild type, -980.8 kcal/mol for the Delta variant, and -1444.5 kcal/mol for the Omicron variant. The Pearson

**Table 4. Median Pair Interaction Energy between the Receptor Binding Domain and Eight Different Monoclonal Antibodies<sup>a</sup>**

monoclonal antibody	median total interaction energy (kcal/mol)			
	wild type	Delta	Omicron (Q493K)	Omicron (Q493R)
regdanvimab	909.6	1078.2 (1.2 × 10 <sup>-4</sup> ) <sup>b</sup>	<u>1347.0</u> (8.1 × 10 <sup>-11</sup> ) <sup>c</sup> (7.8 × 10 <sup>-4</sup> ) <sup>d</sup>	<u>1401.9</u> (7.1 × 10 <sup>-10</sup> ) <sup>c</sup> (1.6 × 10 <sup>-3</sup> ) <sup>d</sup>
bamlavimab	1202.1	1371.2 (1.3 × 10 <sup>-2</sup> ) <sup>b</sup>	<u>1661.9</u> (1.4 × 10 <sup>-8</sup> ) <sup>c</sup> (5.5 × 10 <sup>-4</sup> ) <sup>d</sup>	<u>1858.2</u> (8.0 × 10 <sup>-14</sup> ) <sup>c</sup> (2.3 × 10 <sup>-8</sup> ) <sup>d</sup>
etesivimab	895.4	1052.2 (4.2 × 10 <sup>-3</sup> ) <sup>b</sup>	925.2 (4.9 × 10 <sup>-1</sup> ) <sup>c</sup> (4.2 × 10 <sup>-2</sup> ) <sup>d</sup>	907.8 (9.5 × 10 <sup>-1</sup> ) <sup>c</sup> (1.1 × 10 <sup>-2</sup> ) <sup>d</sup>
casirivimab	1002.6	990.0 (8.0 × 10 <sup>-1</sup> ) <sup>b</sup>	976.5 (8.2 × 10 <sup>-1</sup> ) <sup>c</sup> (8.7 × 10 <sup>-1</sup> ) <sup>d</sup>	1149.2 (1.9 × 10 <sup>-1</sup> ) <sup>c</sup> (1.0 × 10 <sup>-1</sup> ) <sup>d</sup>
imdevimab	967.4	910.3 (9.7 × 10 <sup>-1</sup> ) <sup>b</sup>	<u>1147.3</u> (4.3 × 10 <sup>-4</sup> ) <sup>c</sup> (5.6 × 10 <sup>-4</sup> ) <sup>d</sup>	<u>1133.1</u> (2.4 × 10 <sup>-3</sup> ) <sup>c</sup> (3.6 × 10 <sup>-3</sup> ) <sup>d</sup>
cilgavimab	744.1	734.8 (8.2 × 10 <sup>-1</sup> ) <sup>b</sup>	711.8 (9.0 × 10 <sup>-1</sup> ) <sup>c</sup> (9.3 × 10 <sup>-1</sup> ) <sup>d</sup>	669.3 (3.1 × 10 <sup>-1</sup> ) <sup>c</sup> (4.137 × 10 <sup>-1</sup> ) <sup>d</sup>
tixagevimab	990.4	958.9 (9.1 × 10 <sup>-1</sup> ) <sup>b</sup>	1118.2 (3.7 × 10 <sup>-2</sup> ) <sup>c</sup> (5.3 × 10 <sup>-2</sup> ) <sup>d</sup>	<u>1089.5</u> (2.3 × 10 <sup>-2</sup> ) <sup>c</sup> (3.3 × 10 <sup>-2</sup> ) <sup>d</sup>
sotrovimab	1006.8	873.7 (4.1 × 10 <sup>-2</sup> ) <sup>b</sup>	1004.5 (8.2 × 10 <sup>-2</sup> ) <sup>c</sup> (3.4 × 10 <sup>-2</sup> ) <sup>d</sup>	949.7 (1.1 × 10 <sup>-1</sup> ) <sup>c</sup> (6.3 × 10 <sup>-1</sup> ) <sup>d</sup>

<sup>a</sup>Underlined monoclonal antibodies are monoclonal antibodies for which the binding affinity with the Omicron variant was significantly weaker than that with the wild type and the Delta variant. <sup>b</sup>Values represent the *P*-value for the median total interaction energy between the wild type and the Delta variant. <sup>c</sup>Values represent the *P*-value for the median total interaction energy between the wild type and the Omicron variant. <sup>d</sup>Values represent the *P*-value for the median total interaction energy between the Delta and Omicron variants.

**Table 5. Residues in the Spike Protein Having Lower Pair Interaction Energy Compared to the Wild Type and the Omicron Variant<sup>a</sup>**

monoclonal antibody	residues
regdanvimab	S375F, N440K, T478K, E484A, Q493K, Q498R, Y505H
bamlavimab	S375F, N440K, T478K, E484A, Q493K, Q498R, Y505H
etesivimab	S375F, N440K, T478K, Q493K, Q498R
casirivimab	G339D, S375F, S477N, Q498R
imdevimab	G339D, S375F, N501Y, Y505H
cilgavimab	G339D, S371L, S373P, K417N, S477N, N501Y
tixagevimab	G339D, S375F, Y505H
sotrovimab	K417N, N440K

<sup>a</sup>Underlined residues are the residues in which the pair interaction energy for the Omicron variant was weaker than that of the wild type for the highest number of monoclonal antibodies.

correlation value of the interaction energy for various docking poses and crystal structures was 0.95, which indicates a high correlation and the consequent high predictability of our method. The PIE between each residue of the RBDs and all residues of hACE2 was calculated to determine differences in the interaction energy of RBD residues and hACE2 between the wild type and the Omicron variant (Figure 1b). PIE values for S375F, N440K, T478K, E484A, Q493K/R, and Q498R in the RBD of the two Omicron variants indicated that they had a stronger interaction than the corresponding residues in the wild type (Table 1). The PIE value of the S371L fragment,

which was only present in the RBD of the Omicron variant containing the Q493R mutation, indicated the fragment to have a significantly stronger interaction than the corresponding residues of the wild type. The G339D, S373P, K417N, G446S, S477N, G496S, and N501Y fragments of the two Omicron variants were found to have weaker interactions than the corresponding residues of the wild type (Table 1). In the crystal structure, S375F, N440K, S477N, T478K, E484A, Q493R, G496S, Q498R, and N501Y interacted more strongly with the omicron variant than with the wild type (Table 2).

The six residues apart from S371L that were predicted to have strong interactions with hACE2 in protein–protein docking simulations also showed strong interactions with hACE2 in the Omicron variant than those in the wild type when analyzed experimentally. The 10 of 13 mutated residues in the Omicron variant with significantly different PIEs showed that the trends of predicted differences in PIE between the wild type and Omicron variants were consistent when used for the predicted binding poses and when using the crystal structure. Therefore, the predictions for the binding affinity between hACE2 and the RBD of the predicted structures in our study can be considered to indicate the reliability of our method.

The energies of five interaction types using PIEDA for the crystal structures are described in Table S7. Electrostatic energy is a major interaction type that makes pair interaction energy in Omicron stronger than in the wild type. Pearson's correlation values of five interaction energy types in the crystal structure and the predicted structure were 0.99 in electrostatic energy, 0.04 in exchange energy, 0.83 in charge transfer energy, 0.77 in dispersion energy, and 0.98 in solvation energy. In

N501Y, the difference of dispersion energy between the crystal structure of the wild type and the Omicron variant was calculated to be 4.3 kcal/mol, which is the biggest among the other residues due to the T-shaped  $\pi$ - $\pi$  stacking interaction between Tyr41 in hACE2 and Tyr501 in the S protein of the Omicron variant (Figure S2). The  $\pi$ - $\pi$  stacking interaction energy changes sensitively with the location and angle between two benzene substructures. Since the positions and angles between the two benzene substructures were T-shaped,  $\pi$ - $\pi$  stacking interactions could be formed in the predicted structures, which differ from one predicted structure to another; the residues where strong dispersive interactions can be formed could not be predicted. Since the energies of five interaction types were high, there was a correlation between the crystal structure and the predicted structure, except for the exchange energy, and the energy trends of the five interaction types of each residue in the S protein could be predicted with our method with high accuracy.

The binding affinities of the predicted RBDs to hACE2 were also predicted using molecular dynamics. The contact regions of hACE2 and the S protein are located on their surfaces and are thus in a hydrophilic environment. The RBD of the Omicron variant has a relatively high frequency of mutations owing to its sequence being comparatively longer than that of the other variants. Hence, it is highly likely that there might be major changes in the RBD structure of the Omicron variant compared to that of the wild type. The altered structure of the RBD of the Omicron variant might also have led to changes in the structure of the binding site of hACE2. Such flexibility of hACE2 not being considered in molecular docking simulations necessitated molecular dynamics simulations to be performed to generate binding poses that take this flexibility into consideration.

Figure S1d shows a predicted RBD structure (purple color) overlapping with the hACE2 structure. As all RBD structures of the Omicron variant containing the Q493R mutation obtained by superimposition were seen to overlap with hACE2, we used the RBD structures of the wild type and the Omicron variant containing the Q493K mutation for molecular dynamics but excluded those of the Omicron variant containing the Q493R mutation.

The three initial types of S protein structures are shown in Figure S1. Changes in RMSD of the complexes formed between hACE2 and the three types of S proteins are shown in Figures S3–S5. The RMSD of the S protein structures of the complex formed between hACE2 and the RBD of the Delta variant was greater than 20 Å. Therefore, this complex was considered to have unstable binding poses. Each of the top 10 complex structures was selected based on their potential energy minima from 28 structures of complexes formed by hACE2 with the three types of S proteins. FMO calculations were performed using 280 complex structures to determine the interaction energy distribution of each variant. We obtained 258 calculable complex structures, of which 81 were those of the wild type, 90 were of the Delta variant, and 87 were of the Omicron variants containing the Q493K mutation. In addition, considering the changes in the hACE2 and S protein structures occurring in the molecular dynamic simulations, FMO calculations were performed on the structures of 516 monomers of the selected complex structures to calculate the total interaction energy of the monomers. The median PIE was  $-759.9$  kcal/mol for the wild type,  $-1164.1$  kcal/mol for the

Delta variant, and  $-1553.4$  kcal/mol for the Omicron variant containing the Q493K mutation (Figure 2a).

The results of protein–protein docking simulations showed the S protein of the Omicron variant to have the highest binding affinity to hACE2 compared to that of the wild type and the Delta variant. In addition, the Pearson correlation value between the median PIEs obtained using protein–protein docking simulations and those obtained using molecular dynamics was 0.995. However, compared to the PIE calculated by means of molecular dynamics, the PIE calculated by means of protein–protein docking simulations differs to a large extent from the experimentally obtained PIE of the crystal structure, and a high correlation is observed between the median PIE of both simulation-based approaches.

We analyzed pair interaction energy to binding poses obtained from molecular dynamics simulations to determine the accuracy of differences in PIE between RBD residues and hACE2 obtained using protein–protein docking simulations. The residues S371L, S375F, N440K, T478K, E484A, Q493K, and Q498R showed significantly stronger interactions in the Omicron variant than the corresponding residues in the wild type (Table 3). The residues G339D, S373P, K417N, G446S, S477N, and N501Y showed significantly weaker interactions in the Omicron variant than the corresponding residues in the wild type (Table 3). The strongly interacting residues obtained by the protein–protein docking simulations formed a subset of those obtained by molecular dynamics, and S371L, which did not show significant differences in PIE obtained from protein–protein docking simulations between the wild type and the Omicron variant, was newly identified as a strongly interacting residue. The weakly interacting residues obtained by molecular dynamics were identical to those obtained from the protein–protein docking simulations. As observed with the protein–protein docking simulations, 10 out of 13 residues, except for two residues that did not have significant differences in PIE, showed that the trends of predicted differences in PIE between the wild type and Omicron variant were consistent with the trends of differences in experimentally observed PIE values.

The S protein of the Omicron variant has a stronger interaction with hACE2 in the region of contact, and the chain that mainly interacts with hACE2 is different from that present in the wild type. The mutated residues of the Omicron variants are shown in Figure 3. The left image of Figure 3 shows the main interacting residue in the S protein to be shifted from the left to the right chain.

Comparing the results of protein–protein docking simulations and molecular dynamics for binding affinity prediction showed S371L or G496S to be residues with significant differences in PIE. The Y505H residue did not show a significant difference in PIE between the two prediction methods. Therefore, the protein–protein docking simulation method could be considered better for binding affinity prediction than molecular dynamics considering these metrics and the calculation cost of each prediction method. In addition, the interaction energy between the residues of the RBD and the receptor or between the entire RBD and the receptor was described well by the median PIE.

**3.3. Prediction of Individual Binding Affinity between the RBD and Eight Different Monoclonal Antibodies.** Protein–protein docking simulations can be performed in the antigen-binding domain of monoclonal antibodies since it is a relatively rigid structure. In addition, based on results we obtained for the prediction of binding



affinity between RBD and hACE2, protein–protein docking simulations could be considered for predicting binding affinity than molecular dynamics. We therefore performed a protein–protein docking simulation to predict the individual binding affinity between three types of RBD and eight different monoclonal antibodies that were used to predict the neutralizing antibody titer compared to that of the wild type. The median PIE of the eight monoclonal antibodies with three types of RBD is shown in Figures S6–S13. The median interaction energy of the RBD of the Omicron variant with regdanvimab, bamlavimab, estesivimab, and imdevimab was significantly lower than that of the wild type (Table 4). The binding affinity of the Omicron variant for estesivimab was stronger than that of the Delta variant. Tixagevimab showed significantly lower binding affinity to the Omicron variant containing the Q493R mutation but not to the one containing the Q493K mutation, compared to the wild type and the Delta variant. In a previous study, analysis of binding with regdanvimab showed that the Delta variant ( $K_d = 0.540$  nM) has a lower binding affinity with hACE2 than that of the wild type ( $K_d = 0.056$  nM).<sup>35</sup> Another study showed that, for all eight monoclonal antibodies, the binding affinity of the Omicron variant was lower than that of the Delta variant, and six monoclonal antibodies, excluding cilgavimab and sotrovimab, did not bind to all S proteins in the Omicron variant.<sup>36</sup> Excluding three out of the total eight monoclonal antibodies analyzed, the binding affinity did not show a significant difference for four among five monoclonal antibodies and was lower for the Omicron variant containing the Q493R mutation than for the Delta variant. Therefore, considering that the tendencies of binding affinity between monoclonal antibodies and each of the three types of RBD were predicted with high accuracy, the binding affinity between RBD structures and novel monoclonal antibodies can be efficiently predicted using protein–protein docking simulations and our proposed energy analysis method.

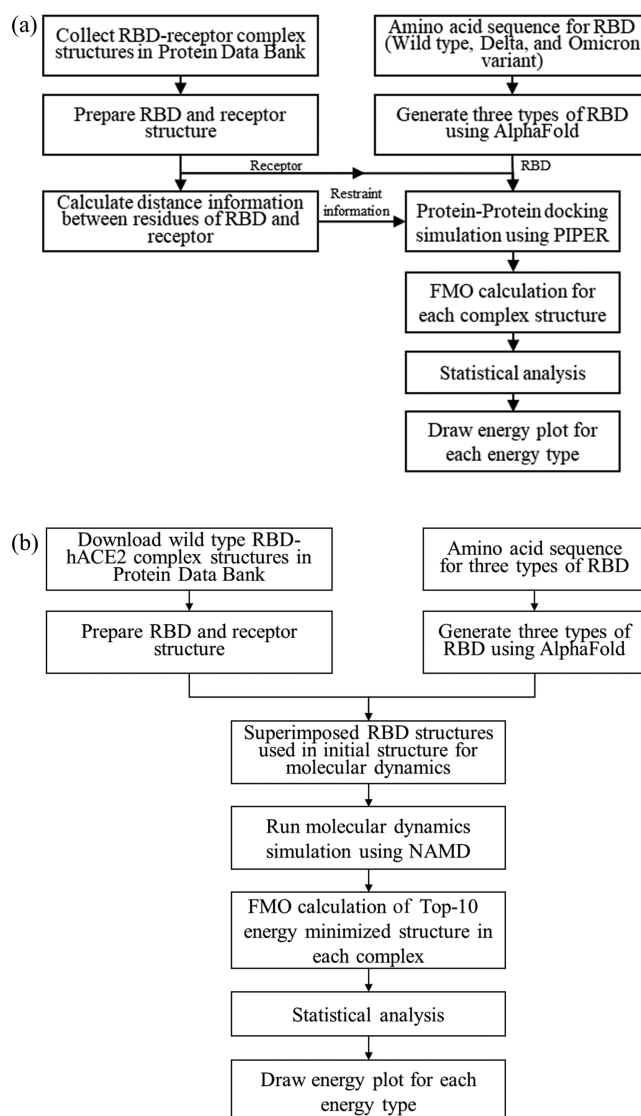
To identify the key mutation site that significantly lowered PIE in the binding of the RBD to the four monoclonal antibodies, we performed PIEDA using all of the binding poses of the eight monoclonal antibodies. The median PIEs of 15 mutation sites are described in Tables S8–S15. Among the 15 mutation sites in the S protein of the Omicron variant, S375F was a common mutation site showing the interaction to be weakened with four monoclonal antibodies, and not with sotrovimab, among the antibodies that had significantly differing PIE from that of the wild type (Table 5). For median PIEs between the residues of the RBD and hACE2, S375F showed a stronger interaction with hACE2 in the Omicron variant than that in the wild type. Therefore, the S375F mutation can be associated with increased infectivity and immune escape ability.

In our study, the analysis of PIE distribution focused on the RBD domain of the spike protein. When developing a new antibody that increases the binding affinity of the spike protein with the RBD domain by modifying the existing antibody, the analysis of PIE distribution was focused on the existing antibodies. The antibody residues with lower binding affinity in the Omicron mutant RBD compared to the wild-type RBD should be identified. The interaction energy between the residues of RBD and the antibody could be obtained using FMO calculation, and interaction types were determined by PIEDA. The amino acid sequences of new antibodies are made by changing the residues in the antibody with a low binding

affinity by reflecting the chemical properties of the residues in the mutated RBD domain that lower the binding affinity. The amino acid sequence of the candidate antibody having a higher binding affinity than the existing antibody can be identified by predicting the binding affinity using this method.

#### 4. CONCLUSIONS

In this study, we proposed a new binding affinity prediction method, the accuracy of which was confirmed by using it to



**Figure 4.** Flowchart showing (a) protein–protein docking simulation using the receptor binding domain of the spike protein, hACE2, and eight monoclonal antibodies and (b) molecular dynamic simulation using the predicted receptor binding domain structure superimposed on the crystal structure of the wild-type receptor binding domain.

predict the binding affinity among three types of RBDs and hACE2. The tendencies of differences in PIE between the three types of S proteins and hACE2 could be predicted using the proposed binding affinity prediction method. The predicted PIE differences between hACE2 and the three types of S proteins obtained using protein–protein docking simulations or molecular dynamics simulations were highly correlated with the PIE difference using the crystal structure.

The six residues of Omicron variants that increased the binding affinity of the virus to hACE2 were predicted using the proposed method. The proposed method was applied to the monoclonal antibody as well as to receptors. Among the eight commercial monoclonal antibodies, four monoclonal antibodies showed a significantly decreased binding affinity with the RBD of the Omicron variant. S375F was found to be a residue that commonly increases the binding affinity of hACE2 with the RBD of the Omicron variant, and it can be considered as a residue that majorly affects the immune escape capability of SARS-CoV-2.

The advantage of the proposed binding affinity prediction method was that the significance of the prediction result could be statistically confirmed. This enabled a more accurate analysis since the significance was presented by a statistical method rather than an absolute numerical value of energy determined by the researcher. The high accuracy of the newly proposed binding affinity prediction method can thus enable efficient prediction of binding affinity using protein–protein docking simulations for any new variant of SARS-CoV-2 that might emerge. Furthermore, if the three-dimensional structures of the receptor or ligand and the binding site of the receptor protein are known, using the approach employed for predicting the binding affinity of RBD-hACE2 or RBD-monoclonal antibody complexes, the binding affinity for protein–protein complexes in general can be predicted efficiently (Figure 4).

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.2c00100>.

Predicted Delta and Omicron variant RBD structures superimposed on crystal wild-type RBD structures; RMSD variation using molecular dynamics; distribution plot of the pair interaction energy by wild type, Delta variant, and Omicron variant; list of eight monoclonal antibodies; median interaction energy of residues in the receptor binding domain interacting with eight monoclonal antibodies; and flowchart for protein–protein docking simulation and molecular dynamic simulation (docx) (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Daeui Park** – Department of Predictive Toxicology, Korea Institute of Toxicology, Daejeon 34114, Republic of Korea; Center for Convergent Research of Emerging Virus Infection, Korea Research Institute of Chemical Technology, Daejeon 34114, Republic of Korea; [orcid.org/0000-0002-9452-5849](https://orcid.org/0000-0002-9452-5849); Phone: +82-42-860-8844; Email: [daeui.park@kitox.re.kr](mailto:daeui.park@kitox.re.kr); Fax: +82-42-610-8157

### Authors

**Sungbo Hwang** – Department of Predictive Toxicology, Korea Institute of Toxicology, Daejeon 34114, Republic of Korea; [orcid.org/0000-0002-1610-5259](https://orcid.org/0000-0002-1610-5259)

**Seung-Hwa Baek** – Department of Predictive Toxicology, Korea Institute of Toxicology, Daejeon 34114, Republic of Korea; Center for Convergent Research of Emerging Virus Infection, Korea Research Institute of Chemical Technology, Daejeon 34114, Republic of Korea

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jcim.2c00100>

### Author Contributions

S.H.: Conceptualization, data curation, formal analysis, methodology, resources, software, validation, visualization, writing—original draft, writing—review & editing. D.P.: Conceptualization, funding acquisition, project administration, supervision, writing—original draft, writing—review & editing. S.-H.B.: Writing—original draft, writing—review & editing. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Notes

The authors declare no competing financial interest.

All data calculated in our study can be downloaded in Mendeley Data.<sup>37</sup> The spike RBD-receptor complex structure was downloaded from Protein Data Bank<sup>38,39</sup> (<https://www.rcsb.org/>). The S protein structures of four types such as wild type, Delta variant, and Omicron variants (Q493K and Q493R) were predicted by AlphaFold<sup>12</sup> (version 2.0.1, <https://github.com/deepmind/alphafold>, accessed by November 17, 2021). We selected predicted 10 structures for the S protein of four types. The PDB files of the structure are available in zip file format in the Supporting Information (Predicted\_S\_protein\_of\_variants.zip). For docking simulation, the protein structures were prepared by the Dock Prep module of Chimera<sup>21</sup> version 1.14. Then, the protein–protein docking simulation between the S protein and receptors including hACE2 and eight mAbs was performed using the restraint mode<sup>40</sup> of the PIPER<sup>22</sup> program (<https://cluspro.org/downloads.php>) in ClusPro<sup>41</sup> (<https://cluspro.org/>). The binding poses between spike-RBD and receptors are available in the zip file format in the Supporting Information, named Docking\_pose files. FMO calculation was performed based on the binding poses using the version June 30, 2020, R1 GAMESS.<sup>32</sup> The energy plot was generated using ggplot2 library version 3.3.3 of R version 4.1.0. For molecular dynamics, the coordinate and structure files were generated by CHARMM-GUI<sup>25,26</sup> (<https://charmm-gui.org/>). Water molecules around the RBD-receptor complex were generated using the Solvate plugin of visual molecular dynamics<sup>27</sup> version 1.9.3. The structure and coordinate files for three types (wild type, Delta variant, and Omicron Q493K variant) of the solvated RBD-hACE2 complex are available in zip file format in the Supporting Information (Molecular\_dynamics\_input\_file.zip). Molecular dynamics was performed using NAMD<sup>24</sup> version 2.12. The workflows for protein–protein docking simulation and molecular dynamics simulation are described in Figure 4.

### ■ ACKNOWLEDGMENTS

This work was supported by the National Research Council of Science and Technology (NST) funded by the Ministry of Science and ICT, Republic of Korea (Grant No. CRC-16-01-KRICT) and the Korea Institute of Toxicology, Republic of Korea (Grant No. 1711133848).

### ■ ABBREVIATIONS

FMO, fragment molecular orbital; hACE2, human angiotensin-converting enzyme 2; MM, molecular mechanical; PIEDA, pair interaction energy decomposition analysis; QM, quantum mechanical; RBD, receptor binding domain; RMSD, root-

mean-squared distance; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

## REFERENCES

- (1) Koopmans, M. The Novel Coronavirus Outbreak: What We Know and What We Don't. *Cell* **2020**, *180*, 1034–1036.
- (2) Li, X.; Geng, M.; Peng, Y.; Meng, L.; Lu, S. Molecular Immune Pathogenesis and Diagnosis of COVID-19. *J. Pharm. Anal.* **2020**, *10*, 102–108.
- (3) Won, J. H.; Lee, H. The Current Status of Drug Repositioning and Vaccine Developments for the COVID-19 Pandemic. *Int. J. Mol. Sci.* **2020**, *21*, No. 9775.
- (4) Huang, Y.; Yang, C.; Xu, X. F.; Xu, W.; Liu, S. W. Structural and Functional Properties of SARS-CoV-2 Spike Protein: Potential Antiviral Drug Development for COVID-19. *Acta Pharmacol. Sin.* **2020**, *41*, 1141–1149.
- (5) Bangaru, S.; Ozorowski, G.; Turner, H. L.; Antanasijevic, A.; Huang, D.; Wang, X.; Torres, J. L.; Diedrich, J. K.; Tian, J.-H.; Portnoff, A. D.; Patel, N.; Massare, M. J.; Yates, J. R.; Nemazee, D.; Paulson, J. C.; Glenn, G.; Smith, G.; Ward, A. B. Structural Analysis of Full-Length SARS-CoV-2 Spike Protein from an Advanced Vaccine Candidate. *Science* **2020**, *370*, 1089–1094.
- (6) Geneva: World Health Organization. WHO COVID-19 Dashboard, <https://covid19.who.int> (accessed December 20, 2021).
- (7) Rambaut, A.; Loman, N.; Pybus, O.; Barclay, W.; Barrett, J.; Carabelli, A.; Connor, T.; Peacock, T.; Robertson, D. L.; Volz, E. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations, <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563> (accessed January 22, 2022).
- (8) Liu, Y.; Liu, J.; Plante, K. S.; Plante, J. A.; Xie, X.; Zhang, X.; Ku, Z.; An, Z.; Scharton, D.; Schindewolf, C.; Widen, S. G.; Menachery, V. D.; Shi, P.-Y.; Weaver, S. C. The N501Y Spike Substitution Enhances SARS-CoV-2 Infection and Transmission. *Nature* **2021**, *602*, 294–299.
- (9) Kumar, V.; Singh, J.; Hasnain, S. E.; Sundar, D. Possible Link between Higher Transmissibility of Alpha, Kappa and Delta Variants of SARS-CoV-2 and Increased Structural Stability of Its Spike Protein and HACE2 Affinity. *Int. J. Mol. Sci.* **2021**, *22*, No. 9131.
- (10) Kumar, S.; Thambiraja, T. S.; Karuppanan, K.; Subramaniam, G. Omicron and Delta Variant of SARS-CoV-2: A Comparative Computational Study of Spike Protein. *J. Med. Virol.* **2021**, *94*, 1641–1649.
- (11) Genovese, L.; Zaccaria, M.; Farzan, M.; Johnson, W.; Momeni, B. Investigating the Mutational Landscape of the SARS-CoV-2 Omicron Variant via Ab Initio Quantum Mechanical Modeling. *bioRxiv* **2021**1521–16 DOI: 10.1101/2021.12.01.470748.
- (12) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589.
- (13) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules. *Chem. Phys. Lett.* **1999**, *313*, 701–706.
- (14) Lim, H.; Baek, A.; Kim, J.; Kim, M. S.; Liu, J.; Nam, K.-Y.; Yoon, J.; No, K. T. Hot Spot Profiles of SARS-CoV-2 and Human ACE2 Receptor Protein Interaction Obtained by Density Functional Tight Binding Fragment Molecular Orbital Method. *Sci. Rep.* **2020**, *10*, No. 16862.
- (15) Kim, J.; Lim, H.; Moon, S.; Cho, S. Y.; Kim, M.; Park, J. H.; Park, H. W.; No, K. T. Hot Spot Analysis of YAP-TEAD Protein-Protein Interaction Using the Fragment Molecular Orbital Method and Its Application for Inhibitor Discovery. *Cancers* **2021**, *13*, No. 4246.
- (16) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nature* **2020**, *581*, 215–220.
- (17) RCSBPDB - 7V8B: Local refinement of SARS-CoV-2 S-Delta variant (B.1.617.2) RBD and Angiotensin-converting enzyme 2 (ACE2) ectodomain, <https://www.rcsb.org/structure/7V8B> (accessed February 25, 2022).
- (18) RCSBPDB - 7T9L: Cryo-EM structure of SARS-CoV-2 Omicron spike protein in complex with human ACE2 (focused refinement of RBD and ACE2), <https://www.rcsb.org/structure/7T9L> (accessed February 25, 2022).
- (19) Elbe, S.; Buckland-Merrett, G. Data, Disease and Diplomacy: GISAID's Innovative Contribution to Global Health. *Global Challenges* **2017**, *1*, 33–46.
- (20) Shu, Y.; McCauley, J. GISAID: Global Initiative on Sharing All Influenza Data – From Vision to Reality. *Eurosurveillance* **2017**, *22*, No. 30494.
- (21) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
- (22) Kozakov, D.; Brenke, R.; Comeau, S. R.; Vajda, S. PIPER: An FFT-Based Protein Docking Program with Pairwise Potentials. *Proteins* **2006**, *65*, 392–406.
- (23) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hémin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, No. 044130.
- (24) Huang, J.; MacKerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135–2145.
- (25) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865.
- (26) Jo, S.; Cheng, X.; Islam, S. M.; Huang, L.; Rui, H.; Zhu, A.; Lee, H. S.; Qi, Y.; Han, W.; Vanommeslaeghe, K.; MacKerell, A. D.; Roux, B.; Im, W. CHARMM-GUI PDB Manipulator for Advanced Modeling and Simulations of Proteins Containing Nonstandard Residues. In *Biomolecular Modelling and Simulations*; Karabencheva-Christova, T., Ed.; Academic Press, 2014; Vol. 96, pp 235–265.
- (27) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (28) Fedorov, D. G.; Kitaura, K. Extending the Power of Quantum Chemistry to Large Systems with the Fragment Molecular Orbital Method. *J. Phys. Chem. A* **2007**, *111*, 6904–6914.
- (29) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.
- (30) Nakano, T.; Kaminuma, T.; Sato, T.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. Fragment Molecular Orbital Method: Application to Polypeptides. *Chem. Phys. Lett.* **2000**, *318*, 614–618.
- (31) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3094.
- (32) Barca, G. M. J.; Bertoni, C.; Carrington, L.; Datta, D.; De Silva, N.; Deustua, J. E.; Fedorov, D. G.; Gour, J. R.; Gunina, A. O.; Guidez, E.; Harville, T.; Irle, S.; Ivanic, J.; Kowalski, K.; Leang, S. S.; Li, H.; Li, W.; Lutz, J. J.; Magoulas, I.; Mato, J.; Mironov, V.; Nakata, H.; Pham, B. Q.; Piecuch, P.; Poole, D.; Pruiett, S. R.; Rendell, A. P.; Roskop, L. B.; Ruedenberg, K.; Sattasathuchana, T.; Schmidt, M. W.; Shen, J.; Slipchenko, L.; Sosonkina, M.; Sundriyal, V.; Tiwari, A.; Galvez Vallejo, J. L.; Westheimer, B.; Wloch, M.; Xu, P.; Zahariev, F.; Gordon, M. S. Recent Developments in the General Atomic and

Molecular Electronic Structure System. *J. Chem. Phys.* **2020**, *152*, No. 154102.

(33) Augusto, G.; Mohsen, M. O.; Zinkhan, S.; Liu, X.; Vogel, M.; Bachmann, M. F. In Vitro Data Suggest That Indian Delta Variant B.1.617 of SARS-CoV-2 Escapes Neutralization by Both Receptor Affinity and Immune Evasion. *Allergy* **2022**, *77*, 111–117.

(34) Lan, J.; He, X.; Ren, Y.; Wang, Z.; Zhou, H.; Fan, S.; Zhu, C.; Liu, D.; Shao, B.; Liu, T.-Y.; Wang, Q.; Zhang, L.; Ge, J.; Wang, T.; Wang, X. Structural and Computational Insights into the SARS-CoV-2 Omicron RBD-ACE2 Interaction. *bioRxiv* **2022**, DOI: 10.1101/2022.01.03.474855.

(35) Ryu, D.-K.; Kang, B.; Noh, H.; Woo, S.-J.; Lee, M.-H.; Nuijten, P. M.; Kim, J.-I.; Seo, J.-M.; Kim, C.; Kim, M.; Yang, E.; Lim, G.; Kim, S.-G.; Eo, S.-K.; Choi, J.; Song, M.; Oh, S.-S.; Chung, H.-Y.; Tijjsma, A. S. L.; van Baalen, C. A.; Kwon, K.-S.; Lee, S.-Y. The in Vitro and in Vivo Efficacy of CT-P59 against Gamma, Delta and Its Associated Variants of SARS-CoV-2. *Biochem. Biophys. Res. Commun.* **2021**, *578*, 91–96.

(36) Planas, D.; Saunders, N.; Maes, P.; Guivel-Benhassine, F.; Planchais, C.; Buchrieser, J.; Bolland, W.-H.; Porrot, F.; Staropoli, I.; Lemoine, F.; Péré, H.; Veyer, D.; Puech, J.; Rodary, J.; Baela, G.; Dellicour, S.; Raymenants, J.; Gorissen, S.; Geenen, C.; Vanmechelen, B.; Wawina-Bokalanga, T.; Martí-Carreras, J.; Cuypers, L.; Sève, A.; Hocqueloux, L.; Prazuck, T.; Rey, F.; Simon-Lorrière, E.; Bruel, T.; Mouquet, H.; André, E.; Schwartz, O. Considerable Escape of SARS-CoV-2 Variant Omicron to Antibody Neutralization. *Nature*. **2022**, *602*, 671–675, DOI: 10.1101/2021.12.14.472630.

(37) Structural dataset using SARS-CoV-2 Spike binding affinity prediction - Mendeley Data, <https://data.mendeley.com/datasets/g6bbwmwkm/1> (accessed March 3, 2022).

(38) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(39) Burley, S. K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chen, L.; Crichlow, G. V.; Christie, C. H.; Dalenberg, K.; Di Costanzo, L.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ganesan, S.; Goodsell, D. S.; Ghosh, S.; Green, R. K.; Guranović, V.; Guzenko, D.; Hudson, B. P.; Lawson, C. L.; Liang, Y.; Lowe, R.; Namkoong, H.; Peisach, E.; Persikova, I.; Randle, C.; Rose, A.; Rose, Y.; Sali, A.; Segura, J.; Sekharan, M.; Shao, C.; Tao, Y.-P.; Voigt, M.; Westbrook, J. D.; Young, J. Y.; Zardecki, C.; Zhuravleva, M. RCSB Protein Data Bank: Powerful New Tools for Exploring 3D Structures of Biological Macromolecules for Basic and Applied Research and Education in Fundamental Biology, Biomedicine, Biotechnology, Bioengineering and Energy Sciences. *Nucleic Acids Res.* **2021**, *49*, D437–D451.

(40) Xia, B.; Vajda, S.; Kozakov, D. Accounting for Pairwise Distance Restraints in FFT-Based Protein–Protein Docking. *Bioinformatics* **2016**, *32*, 3342–3344.

(41) Desta, I. T.; Porter, K. A.; Xia, B.; Kozakov, D.; Vajda, S. Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure* **2020**, *28*, 1071–1081.e3.