

Generalized Methodology for the Quick Prediction of Variant SARS-CoV-2 Spike Protein Binding Affinities with Human Angiotensin-Converting Enzyme II

Alexander H. Williams and Chang-Guo Zhan*



Cite This: *J. Phys. Chem. B* 2022, 126, 2353–2360



Read Online

ACCESS |



Metrics & More

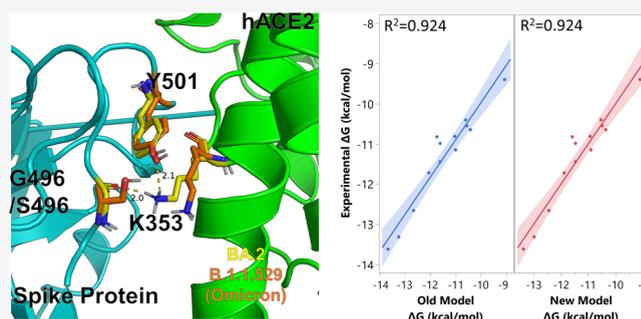


Article Recommendations



Supporting Information

ABSTRACT: Variants of the SARS-CoV-2 virus continue to remain a threat 2 years from the beginning of the pandemic. As more variants arise, and the B.1.1.529 (Omicron) variant threatens to create another wave of infections, a method is needed to predict the binding affinity of the spike protein quickly and accurately with human angiotensin-converting enzyme II (ACE2). We present an accurate and convenient energy minimization/molecular mechanics Poisson–Boltzmann surface area methodology previously used with engineered ACE2 therapeutics to predict the binding affinity of the Omicron variant. Without any additional data from the variants discovered after the publication of our first model, the methodology can accurately predict the binding of the spike/ACE2 variant complexes. From this methodology, we predicted that the Omicron variant spike has a K_d of ~ 22.69 nM (which is very close to the experimental K_d of 20.63 nM published during the review process of the current report) and that spike protein of the new “Stealth” Omicron variant (BA.2) will display a K_d of ~ 12.9 nM with the wild-type ACE2 protein. This methodology can be used with as-yet discovered variants, allowing for quick determinations regarding the variant’s infectivity versus either the wild-type virus or its variants.



INTRODUCTION

The SARS-CoV-2 (COVID-19) pandemic has become the greatest challenge facing the scientific community in the 21st century and is responsible for a total death toll of over 5 million globally and continues to be responsible for over 5000 deaths per day. SARS-CoV-2, like SARS-CoV before it, uses the membrane-bound human angiotensin-converting enzyme II (ACE2) as an entry point into human cells. To achieve this, the viral capsid of SARS-CoV-2 is lined with a spike protein, containing a receptor-binding domain (RBD) region with exceptional binding affinity to ACE2. This spike protein is used as the antigen within several vaccines to induce the antibody response and is also an important target for monoclonal antibody therapeutics (mAb). This link to multiple types of therapeutics and being the primary entry point to human cells leads to the spike protein being a large vulnerability within our treatments for SARS-CoV-2. Compared to other commonly mutating viruses (e.g., the seasonal influenza A virus), COVID-19 presents with both a reproduction rate nearly twice that of the flu and a mutation rate 1000-fold higher.¹ These mutations have led to variants with greater infectivity and the ability to break through the protection given by the COVID-19 vaccines (e.g., B.1.617.2 known as the Delta variant).² Overall, this high mutation rate has led to over 1500 known lineages of the

SARS-CoV-2 virus, with over 200 variants being discovered in 2021.³

With the geographic breadth that became available to the virus, it was inevitable that certain variants would start appearing and representing a greater threat than the wild-type virus. Indeed, one of the first mutations that quickly gained prominence with those studying the virus was the D614G mutation near the Furin cleavage site of the spike protein. This mutation arose so quickly that by the time that most lockdown procedures began in the United States (i.e., March 2020) the D614G mutation was already seen within 26% of all sequenced virus samples.⁴

As the pandemic continued into late 2020, the first of the Greek alphabet named variants was discovered within the United Kingdom. The B.1.1.7 variant, later named α variant, quickly gained a large foothold within the United Kingdom, and by mid-December accounted for over half of the sequenced viruses in the city of London.^{5,6} Unique to the

Received: December 20, 2021

Revised: February 24, 2022

Published: March 22, 2022



variant at that time was the relatively large number of mutations upon the spike protein, including the N501Y mutation within the (RBD). The impact that this mutation had on the binding affinity of the spike protein was initially unknown, as high-resolution Cryo-EM structures of the α variant spike protein were unavailable until early 2021. During this early period of the α variant's discovery, our lab was simultaneously developing a methodology to predict the binding affinity of several published ACE2 mimicking proteins that were early candidates for COVID-19 therapeutics.^{7,8} This methodology utilized a combination of energy minimization and molecular mechanics/Poisson–Boltzmann surface area (MM-PBSA) calculations starting from an available experimental structure of wild-type spike (RBD) binding with ACE2 to estimate the binding affinities of the spike mutants with WT ACE2, several engineered ACE2 proteins (e.g., ACE2.v2 and ACE2.v2.4) which had shown subnanomolar K_d with the Spike protein. Employing this methodology upon the B.1.1.7 spike/ACE2 complex, we obtained a prediction of 0.44 nM, a 55-fold increase in binding affinity over the wild-type spike protein in binding with ACE2, when experimental binding affinity was not available. This prediction was validated before our paper was published, as during the review process, the measurement for the B.1.1.7 spike/ACE2 complex's K_d was published as 0.8 nM, an error of only 0.36 nM for our computational prediction.

Since the publication of that initial methodology, the prevalence of COVID-19 variants has only intensified. The α variant quickly mutated to include the E484K mutation, spawning the B.1.351 (β) variant, and within the United States, the introduction of the B.1.617.2 (Delta) variant induced a new wave of COVID-19 cases that rivaled the mortality of the initial wave of the pandemic.⁹ As vaccination rates vary depending on the locale, and as COVID-19 becomes endemic, the threat of new variants being introduced will not end in the foreseeable future.^{10–13} Even now, the B.1.1.529 (Omicron) variant has caused concern due to the sheer number of mutations found upon the Spike protein (32 in total with 15 upon the RBD region).^{14–17} Additionally, the high mutation rate of SARS-CoV-2 has led to myriad lineages, making the testing of each lineage's binding affinity to the ACE2 protein a time-consuming and cost-prohibitive venture.³ As these variants continue to manifest, and nearly all lineages and sequenced samples of SARS-CoV-2 containing some mutation upon the spike protein (e.g., D614G),¹⁸ a methodology is required to quickly screen these variants to determine their spike protein's binding affinity to the ACE2 protein. This can allow for quick determinations on the potential for a new variant to have an increase in infectivity versus the wild-type spike protein or other variants. The availability of variant spike/ACE2 K_d values also presents an opportunity further generalize our model beyond the engineered ACE2 proteins that represented most of the measurements within our previous model.

To investigate the structural changes seen within the Omicron variant's complex with the ACE2 protein, we have applied our previous energy minimization/MM-PBSA methodology to this new variant. Additionally, we have employed this methodology to several other variants of COVID-19 with known K_d values for their spike/ACE2 complexes to further generalize our previously published model. These improvements to the model will allow for fast and accurate prediction

of known and yet-to-come COVID-19 variant spike/ACE2 binding affinity.

METHODS

We have previously shown that the energy minimization/MM-PBSA methodology starting from the available experimental structure of the wild-type protein complex can accurately rank small-molecule compounds in concordance with their *in vitro* binding data.^{19–24} Our previous study on the binding mode of the α variant revealed that this methodology is also applicable to protein/protein systems as well as protein/ligand systems. Prior to this study, our lab has used a similar methodology to investigate potential mutations for use in protein therapeutics. This methodology contrasts with other investigations that have focused on predicting the binding affinity of SARS-CoV-2 mutants; most studies of the Spike/ACE2 system employ some form of either all-atom^{25–27} or coarse-grained molecular dynamics simulation²⁸ to generate snapshots of the complex which are then used to estimate the binding affinity using either the MM-PBSA or generalized Born solvation area (MM-GBSA) method. Our energy minimization/MM-PBSA methodology differs from these as it only requires a single snapshot of the Spike/ACE2 complex in its post-minimization state. Using this approach allows one to avoid these costly and time-consuming molecular dynamics (MD) simulations to reach the equilibrium state of the mutated protein structure, while still accurately predicting binding free energies in line with *in vitro* experimental data.^{19–24} Additionally, the limitations placed upon MD simulations (e.g., force field cutoffs, periodic conditions, etc.) can introduce artifacts at longer time-scales,^{29–33} which are required to perform free-energy perturbation calculations.^{34,35}

Usage of this methodology assumes that the initial binding mode of either the ligand or protein with their target protein is close to their equilibrium state (or “unperturbed state”) and that any changes to either (i.e., changes to the atoms of a ligand or residues of a protein) represent a small perturbation to an otherwise reliable structure. The widespread availability of Spike/ACE2 Cryo-EM structures makes the finding the unperturbed state of this complex a non-issue, and through energy minimization, the perturbed state of the protein/protein complex can be brought to a local energy minimum close to the true unperturbed state of the mutated protein.

Since the publication of our methodology for the fast prediction of Spike/ACE2 binding affinities, multiple publications have reported the dissociation constants of prolific variants of the SARS-CoV-2 Spike protein with ACE2.^{36–46} With these data in hand, our model can be updated to better predict the binding affinity of emerging variants of concern such as the Omicron variant.¹⁵ Briefly, for each variant, the SARS-CoV-2 Spike protein (consisting of the Spike RBD region binding with ACE2) crystal structure (PDB: 7KMB)⁴⁷ was mutated using the PyMol mutagenesis tool, making sure to choose the lowest-energy rotamer of the mutated residue. Additionally, three engineered ACE2 proteins with known subnanomolar K_d were modeled to determine if the methodology could also include mutations to ACE2. The resulting mutated complex was then prepared using the pdb4amber tool of the AmberTools2020 package to remove any nonprotein residues and any residues that were not Amber-compliant. Important to note is the absence of any glycans or other similar post-translational modifications to the Spike or ACE2 proteins. While glycans have been implicated in facilitating the “up/

Table 1. Calculated Binding Free Energies of Spike/ACE2 Complexes with Multiple Variants

| variant of spike | experimental K_d (nM) ^a | experimental ΔG_{exp} (kcal/mol) ^b | ΔG_{PB} (kcal/mol) ^c | old model prediction (kcal/mol) ^d | old model prediction (nM) ^e | new model prediction ΔG (kcal/mol) ^f | new model prediction (nM) ^g |
|----------------------|--------------------------------------|--|--|--|--|---|--|
| α (B.1.1.7) | 2.90 | -11.72 | -81.93 | -12.05 | 1.66 | -11.88 | 2.20 |
| β (B.1.351) | 13.24 | -10.81 | -78.06 | -11.03 | 9.28 | -10.93 | 10.90 |
| Delta (B.1.617.2) | 17.51 | -10.65 | -75.74 | -10.41 | 26.07 | -10.36 | 28.53 |
| E484K | 13.01 | -10.82 | -80.73 | -11.73 | 2.84 | -11.59 | 3.62 |
| γ (P.1) | 7.54 | -11.15 | -77.94 | -10.99 | 9.78 | -10.90 | 11.45 |
| K417N | 141.95 | -9.40 | -70.63 | -9.06 | 252.02 | -9.10 | 235.54 |
| L452R/ E484Q | 4.60 | -11.44 | -80.25 | -11.61 | 3.51 | -11.47 | 4.41 |
| N440K | 9.91 | -10.99 | -80.24 | -11.60 | 3.52 | -11.47 | 4.43 |
| Omicron (B.1.1.529) | 20.63 ⁱ | -10.55 | -76.29 | -10.56 | 20.39 | -10.49 | 22.69 |
| BA.2 | N/A | N/A | -77.64 | -10.91 | 11.19 | -10.83 | 12.98 |
| WT | 26.37 | -10.40 | -76.45 | -10.60 | 18.97 | -10.53 | 21.22 |
| ACE2.v2.4 | 0.60 | -12.66 | -84.20 | -12.65 | 0.61 | -12.44 | 0.86 |
| ACE2.v2 | 0.20 | -13.31 | -86.40 | -13.24 | 0.23 | -12.99 | 0.35 |
| ACE2v2.4/ b.1.1.7 | 0.12 | -13.62 | -88.00 | -13.66 | 0.11 | -13.38 | 0.18 |
| | | | RMSD ^h | 0.35 | | 0.32 | |

^aExperimental K_d determined for each variant of the Spike/ACE2 complex using the average ΔG_{exp} values within Table S1.^{39–46} ^bExperimental binding affinity converted to Gibbs binding free energy: $\Delta G_{\text{exp}} = -RT \ln(K_d)$. ^cCalculated binding free energy of complex using the MM-PBSA approach. ^dCorrected binding free energy using the original linear regression model.⁷ ^ePredicted K_d of the system using the equation: $K_d = e^{\Delta G_{\text{corr}}/-RT}$ for the original model (eq 2) prediction. ^fCorrected binding free energy using eq 1. ^gPredicted K_d of the system using the equation: $K_d = e^{\Delta G_{\text{corr}}/-RT}$ for eq 1 prediction. ^hRMSD of the predicted ΔG_{PB} vs ΔG_{exp} . ⁱPublished during submission of the original manuscript.

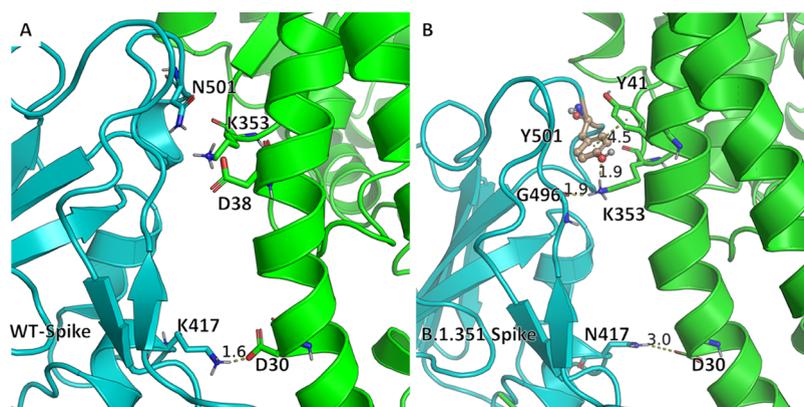


Figure 1. (A) Binding mode of wild-type spike protein, highlighting the $+/-$ charge interaction of K417 with nearby D30. In the B.1.351 variant, this residue is replaced with N417, which is only able to weakly hydrogen-bond with the D30 residue. However, the N501Y mutation is able to introduce a strong hydrogen bond with nearby D38 and a weaker hydrogen bond with K353 partially offsetting this loss in $+/-$ charge interaction. Additionally, the aromatic ring of the mutated Y501 can form a $\pi-\pi$ stacking interaction with nearby Y41 within ACE2.

down” conformation of the entire Spike protein⁴⁸ and shielding the virus from immune recognition,⁴⁹ the RBD region being investigated in this study is notably less glycanated.⁴⁹ This is confirmed by numerous Spike protein crystal structures only containing one glycation site (PDB: 6M0J, 6LZG, 6XE1, and 7KMB) far removed from the Spike/ACE2 binding interface. Additionally, these glycanated amino acids within the crystal structures rarely contain more than the base GlcNAc saccharide, truncating the glycan significantly. Conversely, *in vitro* studies of the ACE2 protein have shown that the removal of the glycans has limited the impact on recognition or binding affinity with the SARS-CoV-2 spike protein.⁵⁰ For these reasons, we chose to ignore the glycanation of the Spike and ACE2 proteins. Glycans were ignored in both the wild-type and mutant structures, ensuring

that the truncated GlcNAc saccharide would not interfere with the simulations.

Each resulting Spike/ACE2 complex’s parameters and coordinates were prepared using the LEaP module of the AmberTools2020 package using the ff99SB force field, which has been previously shown to have the best predictive ability to rank entities within the MM-PBSA methodologies.^{20,21,51–53} Each complex was then energy-minimized using the SANDER module of the Amber2020 package using a two-step energy minimization procedure. Each step consisted of 2000 steps of steepest descent energy minimization, followed by 3000 steps of conjugate gradient descent, the first step placed a 10 kcal/Å³ restraint on all nonhydrogen atoms, while the second step placed a 2 kcal/Å³ restraint on the backbone C_α atoms. Finally, the complex’s ΔG_{PB} was estimated using the MMPBSA.py module of the AmberTools2020 package using the MM-PBSA

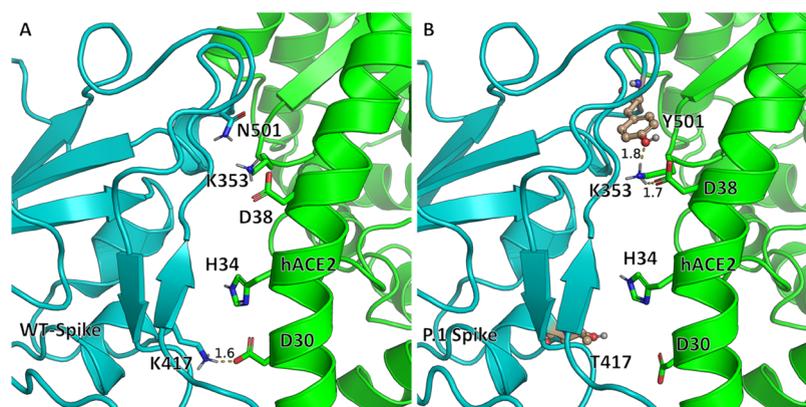


Figure 2. (A) Binding mode of the wild-type spike protein with ACE2, highlighting the K417 interaction with nearby D30 of ACE2. In the P.1 variant, this strong $+/-$ charge interaction is replaced with T417, which is unable to even hydrogen-bond with D30 like the K417N mutation of the β variant, weakening the overall binding affinity between the two proteins. However, the N501Y mutation is able to introduce a hydrogen bond with nearby K353.

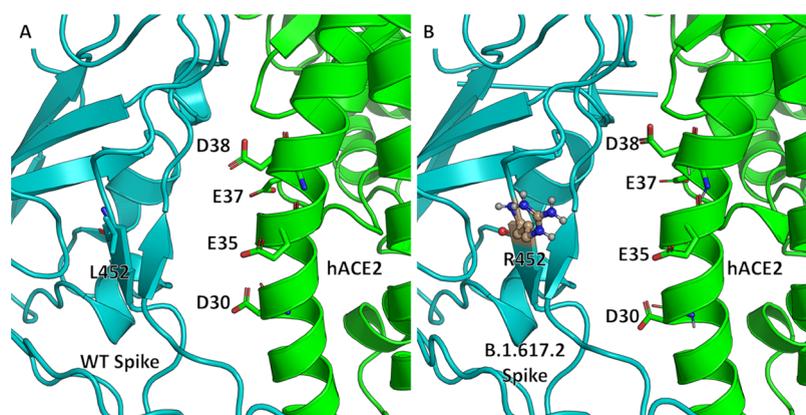


Figure 3. (A) Binding mode of WT spike protein with ACE2, with the D30 to D38 negatively charged cluster of residues in stick representation. Mutation of L452 to R452 in the Delta variant (B) places a positively charged residue in proximity to this cluster of negatively charged residue on ACE2, increasing the binding affinity between the two proteins.

methodology; the entropy for these binding modes was estimated using the quasi-harmonic methodology built into the MMPBSA.py program.⁵⁴ As expected, the MM-PBSA values obtained for each complex overestimate the absolute binding free energy of the complex. To account for this overestimation, the experimental dissociation constants were used to create a linear regression equation to convert between our obtained MM-PBSA values and the experimental values (Tables 1 and S1).

Results and Discussion. The obtained binding modes for the variants used within our linear regression model can be found in Figures 1–5. The linear regression analysis data can be seen in Figure 6 and Table 1. The MM-PBSA correction equation obtained from the linear regression analysis of the raw MM-PBSA calculations *vs* the *in vitro* K_d data collected from the spike/ACE2 complexes can be found within eq 1 (see below).

Structural Analysis of the SARS-CoV-2 Variant Spike Proteins. Structural Analysis of the B.1.351 β Variant. The β variant shares the N501Y mutation with the α variant but introduces two key mutations that worsen the binding mode with ACE2. As in the α variant, the N501Y mutation induces an additional hydrogen bond with nearby K353, along with a π - π stacking interaction with nearby Y41 of ACE2, which has been reported within other computational studies of the

N501Y mutation.²⁵ Additionally, the movement of K353 to hydrogen-bond with the mutated Y501 residue induces another hydrogen bond with the backbone of G496.⁷ However, the K417N mutation removes a strong $+/-$ charge interaction with D30 of ACE2, replacing it with a weak hydrogen bond with D30, resulting in a lower binding affinity of 13.2 nM *vs* the α variant's 2.9 nM. The deleterious effects of the K417N mutation have been previously reported, and single mutation K417N variants of the Spike protein massively decrease the K_d to nearly 200 nM.^{43,46} Additionally, the E484K mutation introduces another positively charged residue into the spike protein, increasing the $+/-$ charge interactions with the overall negative charge of ACE2.

Structural Analysis of the P.1 γ Variant. The γ variant suffers from the same residue incompatibility of the β variant, as the K417T mutation similarly removes a strong $+/-$ charge interaction between K417 and D30. The N501Y mutation plays the same role as seen within the α and β variants, providing an additional hydrogen bond with D38 and K353 of ACE2. This weakening of the $+/-$ charge interaction leads to a K_d of 4.8 nM *vs* the α variant's 0.8 nM.

Structural Analysis of the B.1.617.2 Delta Variant. The Delta variant contains two residue mutations upon the RBD region: L452R and T478K. Both mutations induce a change from an uncharged to a positively charged residue,

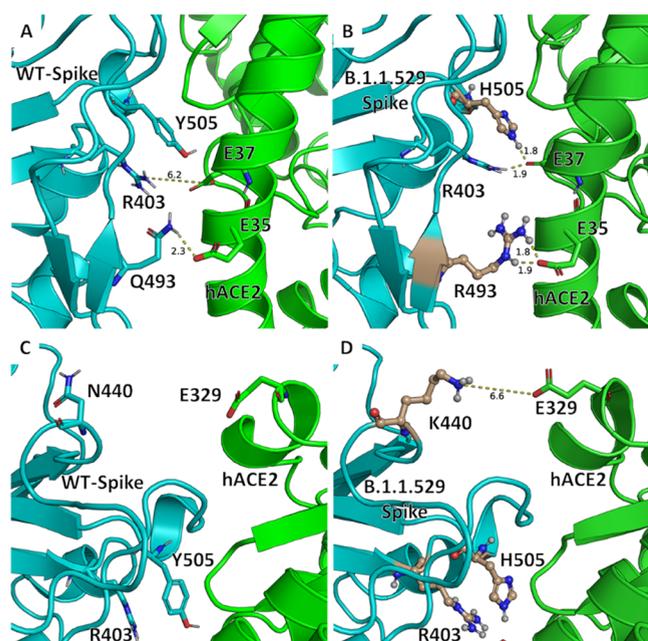


Figure 4. (A) Binding mode of wild-type Spike protein with human ACE2. The E35 and E37 residues do not closely bind with any residues on the wild-type spike protein. (B) Binding mode of B.1.1.529 spike protein, with mutated residues in brown ball and sticks. R493 and H505 contribute $+/-$ charge interactions with E35 and E37 respectively. (C) N440 on the wild-type protein does not bind with ACE2; however, upon mutation to K440 produces a $+/-$ charge interaction with E329 of ACE2 (D).

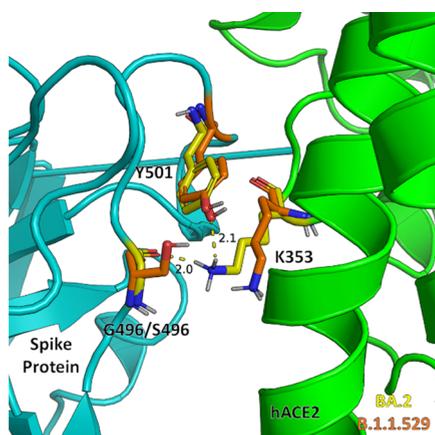


Figure 5. Magnified binding mode of the binding interface of the Omicron/BA.2 Spike variants with ACE2. The omicron variant (Orange Sticks) introduces the G496S mutation, which places steric strain on K353, moving it away from the interface and breaking two hydrogen bonds seen within the BA.2 variant (yellow sticks). With this mutation removed, K353 can hydrogen-bond with G496 and Y501, an interaction shared with both the α and β variants.

complimenting the overall negative charge of the ACE2 protein. L452R is in proximity to a cluster of negatively charged residues including D30, E35, E37, and D38, while T478K is in proximity to a nearby E87. The complimentary charge/charge interactions increase the binding affinity between the Delta variant's spike protein and ACE2.

Structural Analysis of the B.1.428 Epsilon Variant. The Epsilon variant only contains the L452R mutation on the RBD region of the spike protein. Like the Delta variant, this

mutation acts to increase the $+/-$ electrostatic interactions between the spike protein and ACE2. This additional interaction between the two proteins leads to a greater binding affinity at 1.2 nM *vs* the wild type's 22 nM. However, the lack of additional mutations to increase infectivity or to avoid the immune response (e.g., E484K) has led to an extremely low prevalence among all COVID-19 cases, with the last reported case arising in September 2021.

Structural Analysis of the Omicron Variant. The Omicron variant contains novel mutations that are not commonly seen within other prevalent variants of SARS-CoV-2. Of note is the mutation Q493R, which is in close proximity to E35 on ACE2. This mutation creates a very strong $+/-$ charge interaction between the two proteins, thus increasing the binding affinity. Additionally, another unique mutation in the receptor-binding motif (RBM) is the Y505H mutation; this mutation is similar to the Q493R mutation as places the charged nitrogen within 2 Å of a nearby E19 residue in ACE2, thus inducing another $+/-$ charge interaction. Finally, the N440K mutation of the Omicron variant induces another, albeit weaker, $+/-$ charge interaction with E329 of the ACE2 protein. However, the Omicron variant also contains the K417N mutation, which similarly destroys the strong $+/-$ interaction with D30 as seen within the β variant.

Structural Analysis of the BA.2 "Stealth Omicron" Variant. While the mutations of the original B.1.1.529 Omicron and BA.2 Stealth Omicron variants are very similar, one residue change stands out. BA.2 lacks the G496S mutation seen within the Omicron variant. The residue change of G496S is responsible for applying steric pressure to the K353 residue in the Omicron variant, blocking it from making the same hydrogen-bond interactions seen within the α or β variants, affording them a much lower K_d value. Removal of this residue change allows the K353 residue to bind with Y501 and G496. These additional hydrogen bonds afford the BA.2 variant a lower K_d compared to the original Omicron variant (Table 1).

Spike/ACE2 Linear Regression Model. As noted previously within our report detailing the energy minimization/MM-PBSA methodology, the estimated binding energy far exceeds that of the actual binding energy (Table 1). However, with a simple linear regression, these predictions can accurately be brought into line with the experimental results (eq 1).

$$\Delta G_{\text{corr}} = 0.247\Delta G_{\text{PB}} + 8.32 \text{ kcal/mol} \quad (1)$$

$$\Delta G_{\text{corr}} = 0.265\Delta G_{\text{PB}} + 9.66 \text{ kcal/mol} \quad (2)$$

The predictions from our model correlate well with the published ΔG_{exp} values for each of the SARS-CoV-2 variants (Table 1). In comparison with our previous model, the root-mean-square error (RMSE) of the ΔG_{corr} in comparison to the ΔG_{exp} has decreased. Notably, both models have very low RMSE values when predicting the ΔG , further validating our original model's (eq 2) predictions. Interestingly, the correlation coefficient remains close to our previously reported coefficients (0.265 for the original spike/ACE2 model and 0.3057 for our CB₁/CB₂ model). When comparing our results to previous attempts to estimate the binding affinity of the SARS-CoV-2 mutants, our model was more successful than similar methodologies that employed all-atom molecular dynamics (MD) simulations along with MM-GBSA estimation of the binding affinity, where the N501Y mutation on the RBD was predicted as deleterious to the binding of the Spike/ACE2 complex.²⁵ While other attempts to estimate the binding

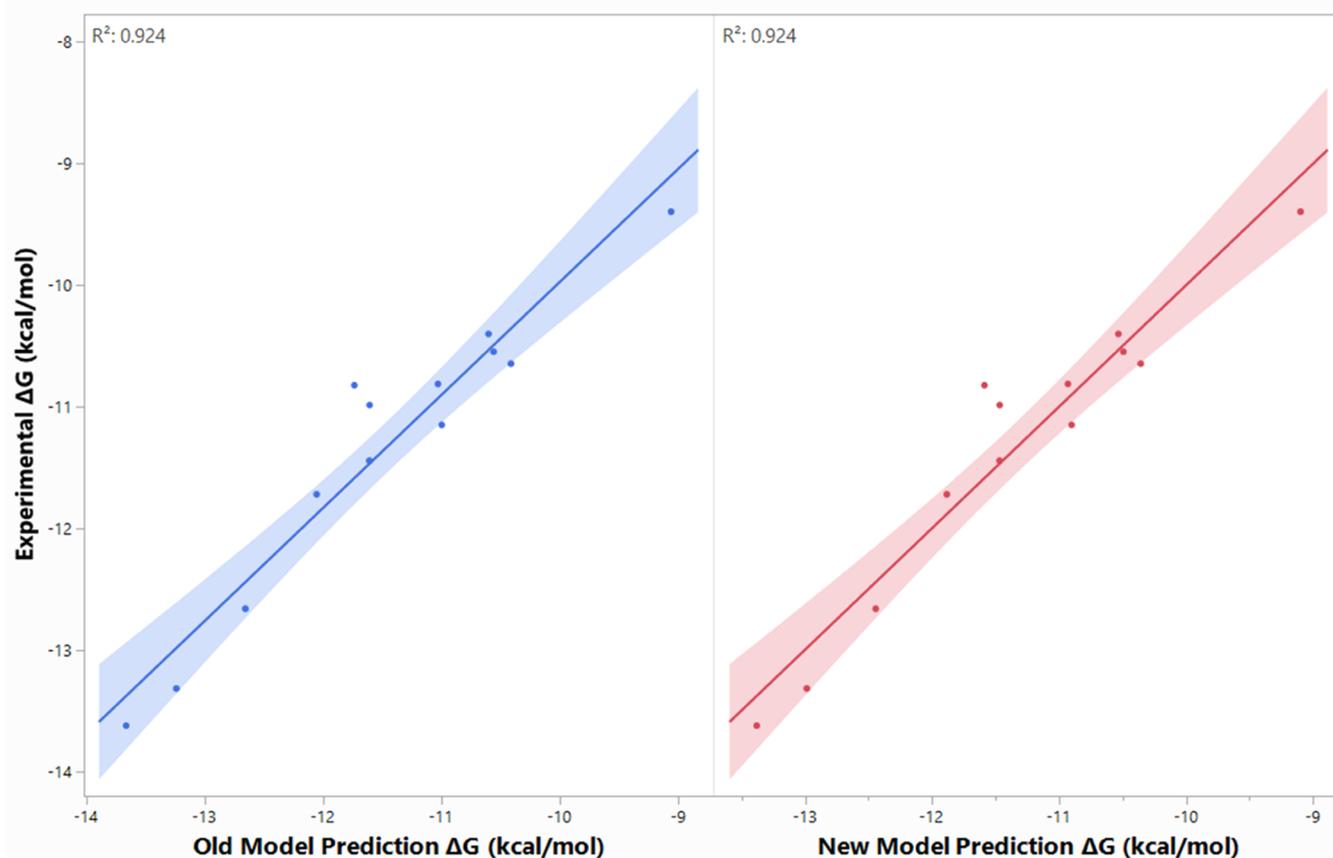


Figure 6. Comparison of the previously published antibody model and our updated model. Each graph shows the correlation of the ΔG_{corr} values produced by the respective equations against the ΔG_{exp} values for the included Spike/ ACE2 variants. While both models have the same R^2 value, the new model's predictions have an RMSD of 0.32 *vs* 0.35 for the older model.

affinity with a combination of MD/MM-GBSA methodologies do show the ability to broadly predict binding affinities that correlate with experimental data,²⁶ meta-analyses have shown that using the MD/MM-GBSA method gives worse correlating predictions than with the energy minimization/MM-PBSA method in protein/peptide systems.²¹ While the MM-PBSA methodology is more computationally expensive than the MM-GBSA calculation, the use of energy minimization *vs* an all-atom MD simulation drastically cuts down on the computing time required for each Spike/ACE2 complex simulation.

To determine whether our methodology was generalizable to other Spike/ACE2 complexes, we repeated the method upon a crystal structure of the B.1.1.7 α Variant Spike with ACE2 (PDB: 7MJN) with the same set of Spike variants, including the WT. While the correlation factor ($R^2 = 0.70$ for eq 3) was not as strong as with the usage of the WT crystal structure (Table S2, see the Supporting Information), our methodology was still able to determine which mutations were advantageous to the binding affinity of the Spike protein (*e.g.*, N501Y), and which were deleterious (*e.g.*, K417N). The use of a new crystal structure requires the use of a linear regression equation, as the baseline α variant crystal structure ΔG_{PB} is different from that of the WT crystal ΔG_{PB} (eq 3.) Additionally, the prediction for the $\alpha \rightarrow$ WT binding affinity exhibited the smallest overall error, showing that the small perturbations of the crystal structure (*i.e.*, one mutation Y501N) lead to the best results in this methodology, whereas

larger mutations (*i.e.*, α to β , α to γ , *etc.*) can produce larger errors.

$$\Delta G_{\text{corr}} = 0.0844\Delta G_{\text{PB}} - 6.48 \text{ kcal/mol} \quad (3)$$

Conclusions. Using our previously published methodology for the prediction of Spike/ACE2 binding affinities, we have successfully created a model able to predict the binding affinity quickly and accurately for multiple variants of SARS-CoV-2. With this model, we are also able to extrapolate the binding affinity for the Omicron variant spike's K_d for ACE2 at ~ 22.69 nM *vs* the recently published experimental *in vitro* value of 20.63 nM (which became available during the review process of the current report.)^{42,43} Additionally, the newly classified variant of concern BA.2 (“Stealth Omicron”) appears to have an even stronger connection with ACE2 with a K_d of 12.93 nM. Compared to other methods that have previously attempted to estimate the binding affinity of these variants with ACE2, our energy minimization/MM-PBSA methodology has a stronger correlation with the experimental *in vitro* values and was able to successfully predict the binding affinity of the Omicron variant without its inclusion in the established linear regression model. With the computational power saved with this methodology, one could systematically predict the binding affinity of hundreds of potential variants of the Spike protein and determine their effect on the binding affinity against ACE2.

Additionally, we have investigated the energy-minimized structures of several of these mutants and developed a

structural basis for the appearance of the mutations seen within these variants. Of note is the N501Y mutation originally seen within the α variant, which alone was able to increase the binding affinity of the Spike protein nearly 10-fold. Additionally, it appears that not all mutations of the Spike protein are advantageous, as the K417N mutation by itself degrades the binding affinity of the Spike protein nearly 10-fold. This combination of mutations partially explains why the Omicron variant, which contains both N501Y and K417N, does not have a substantially different binding affinity compared to the WT spike protein, as the charge–charge interactions gained from the N501Y mutation are nullified from the loss in a strong charge/charge interaction from the K417N mutation.

This methodology can be applied to any future variant of SARS-CoV-2 and will allow for the quick prediction of its binding affinity with ACE2 without the need for time-consuming experiments such as bilayer interferometry. Additionally, future efforts will be made to adapt this methodology for use on SARS-CoV-2 antibodies to determine the structural effects of mutations on their binding with the Spike protein.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.1c10718>.

Available experimental data and references (Table S1) and additional computational data associated with eq 3 (Table S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Chang-Guo Zhan – Molecular Modeling and Biopharmaceutical Center and Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lexington, Kentucky 40536, United States;
✉ orcid.org/0000-0002-4128-7269; Phone: 859-323-3943; Email: zhan@uky.edu; Fax: 859-257-7585

Author

Alexander H. Williams – Molecular Modeling and Biopharmaceutical Center and Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lexington, Kentucky 40536, United States

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jpcb.1c10718>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by the funding of the Molecular Modeling and Biopharmaceutical Center at the University of Kentucky College of Pharmacy, the National Science Foundation (NSF grant CHE-1111761), and the National Institutes of Health (P20 GM130456). The authors also acknowledge the Computer Center at University of Kentucky for supercomputing time on their Lipscomb Compute Cluster, and their NVIDIA V100 nodes.

■ REFERENCES

(1) Liu, L.; Zeng, F.; Rao, J.; Yuan, S.; Ji, M.; Lei, X.; Xiao, X.; Li, Z.; Li, X.; Du, W.; et al. Comparison of clinical features and outcomes of

medically attended COVID-19 and influenza patients in a defined population in the 2020 respiratory virus season. *Front. Public Health* **2021**, *9*, No. 587425.

(2) Christensen, P. A.; Olsen, R. J.; Long, S. W.; Subedi, S.; Davis, J. J.; Hodjat, P.; Walley, D. R.; Kinsley, J. C.; Saavedra, M. O.; Pruitt, L.; et al. Delta variants of SARS-CoV-2 cause significantly increased vaccine breakthrough COVID-19 cases in Houston, Texas. *Am. J. Pathol.* **2021**, *192*, 320.

(3) Rambaut, A.; Holmes, E. C.; O’Toole, Á.; Hill, V.; McCrone, J. T.; Ruis, C.; du Plessis, L.; Pybus, O. G. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **2020**, *5*, 1403–1407.

(4) Zhang, L.; Jackson, C. B.; Mou, H.; Ojha, A.; Rangarajan, E. S.; IZard, T.; Farzan, M.; Choe, H. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. *BioRxiv* **2020**, 2020.06.12.148726 DOI: [10.1101/2020.06.12.148726](https://doi.org/10.1101/2020.06.12.148726).

(5) Mahase, E. Covid-19: What have we learnt about the new variant in the UK? *BMJ* **2020**, *371*, No. m4944.

(6) Wise, J. Covid-19: New coronavirus variant is identified in UK. *BMJ* **2020**, *371*, No. m4857.

(7) Williams, A. H.; Zhan, C.-G. Fast Prediction of Binding Affinities of the SARS-CoV-2 Spike Protein Mutant N501Y (UK Variant) with ACE2 and Miniprotein Drug Candidates. *J. Phys. Chem. B* **2021**, *125*, 4330–4336.

(8) Cao, L.; Goresnik, I.; Coventry, B.; Case, J. B.; Miller, L.; Kozodoy, L.; Chen, R. E.; Carter, L.; Walls, A. C.; Park, Y.-J.; et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **2020**, *370*, 426–431.

(9) Kupferschmidt, K.; Wadman, M. Delta variant triggers new phase in the pandemic. *Science* **2021**, *372*, 1375–1376.

(10) McAteer, J.; Yildirim, I.; Chahroudi, A. The VACCINES Act: deciphering vaccine hesitancy in the time of COVID-19. *Clin. Infect. Dis.* **2020**, *71*, 703–705.

(11) Dror, A. A.; Eisenbach, N.; Taiber, S.; Morozov, N. G.; Mizrahi, M.; Zigran, A.; Srouji, S.; Sela, E. Vaccine hesitancy: the next challenge in the fight against COVID-19. *Eur. J. Epidemiol.* **2020**, *35*, 775–779.

(12) Sallam, M. COVID-19 vaccine hesitancy worldwide: a concise systematic review of vaccine acceptance rates. *Vaccines* **2021**, *9*, 160.

(13) Coustasse, A.; Kimble, C.; Maxik, K. COVID-19 and vaccine hesitancy: a challenge the United States must overcome. *J. Ambulatory Care Manage.* **2021**, *44*, 71–75.

(14) Newsroom, C., CDC Statement on B. 1.1. 529 (Omicron variant).

(15) Torjesen, I. Covid-19: Omicron may be more transmissible than other variants and partly resistant to existing vaccines, scientists fear. *BMJ* **2021**, *375*, No. n2943.

(16) Rao, S.; Singh, M. The Newly Detected B. 1.1. 529 (Omicron) Variant of SARS-CoV-2 With Multiple Mutations: Implications for transmission, diagnostics, therapeutics, and immune evasion. *DHR Proc.* **2021**, *1*, 7–10.

(17) Sahoo, J. P.; Samal, K. C. World on Alert: WHO Designated South African New COVID Strain (Omicron/B. 1.1. 529) as a Variant of Concern. *Biot. Res. Today* **2021**, *3*, 1086–1088.

(18) Chen, C.; Nadeau, S. A.; Yared, M.; Voinov, P.; Stadler, T. CoV-Spectrum: Analysis of globally shared SARS-CoV-2 data to Identify and Characterize New Variants, arXiv:2106.08106. arXiv.org e-Print archive. <https://arxiv.org/abs/2106.08106> (submitted Jun 15, 2021).

(19) Yau, M. Q.; Emtage, A. L.; Chan, N. J.; Doughty, S. W.; Loo, J. S. Evaluating the performance of MM/PBSA for binding affinity prediction using class A GPCR crystal structures. *J. Comput.-Aided Mol. Des.* **2019**, *33*, 487–496.

(20) Chen, F.; Sun, H.; Wang, J.; Zhu, F.; Liu, H.; Wang, Z.; Lei, T.; Li, Y.; Hou, T. Assessing the performance of MM/PBSA and MM/GBSA methods. 8. Predicting binding free energies and poses of protein–RNA complexes. *RNA* **2018**, *24*, 1183–1194.

(21) Weng, G.; Wang, E.; Chen, F.; Sun, H.; Wang, Z.; Hou, T. Assessing the performance of MM/PBSA and MM/GBSA methods.

9. Prediction reliability of binding affinities and binding poses for protein–peptide complexes. *Phys. Chem. Chem. Phys.* **2019**, *21*, 10135–10145.
- (22) Rastelli, G.; Rio, A. D.; Degliesposti, G.; Sgobba, M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *J. Comput. Chem.* **2010**, *31*, 797–810.
- (23) Petukh, M.; Li, M.; Alexov, E. Predicting binding free energy change caused by point mutations with knowledge-modified MM/PBSA method. *PLoS Comput. Biol.* **2015**, *11*, No. e1004276.
- (24) Gromiha, M. M.; Yugandhar, Ka.; Jemimah, S. Protein–protein interactions: scoring schemes and binding affinity. *Curr. Opin. Struct. Biol.* **2017**, *44*, 31–38.
- (25) Spinello, A.; Saltalamacchia, A.; Borisek, J.; Magistrato, A. Allosteric cross-talk among spike's receptor-binding domain mutations of the SARS-CoV-2 South African variant triggers an effective hijacking of human cell receptor. *J. Phys. Chem. Lett.* **2021**, *12*, 5987–5993.
- (26) Antony, P.; Vijayan, R. Molecular Dynamics Simulation Study of the Interaction between Human Angiotensin Converting Enzyme 2 and Spike Protein Receptor Binding Domain of the SARS-CoV-2 B.1.617 Variant. *Biomolecules* **2021**, *11*, 1244.
- (27) Spinello, A.; Saltalamacchia, A.; Magistrato, A. Is the rigidity of SARS-CoV-2 spike receptor-binding motif the hallmark for its enhanced infectivity? Insights from all-atom simulations. *J. Phys. Chem. Lett.* **2020**, *11*, 4785–4790.
- (28) Nguyen, H. L.; Lan, P. D.; Thai, N. Q.; Nissley, D. A.; O'Brien, E. P.; Li, M. S. Does SARS-CoV-2 bind to human ACE2 more strongly than does SARS-CoV? *J. Phys. Chem. B* **2020**, *124*, 7336–7347.
- (29) Patra, M.; Karttunen, M.; Hyvönen, M. T.; Falck, E.; Lindqvist, P.; Vattulainen, I. Molecular dynamics simulations of lipid bilayers: major artifacts due to truncating electrostatic interactions. *Biophys. J.* **2003**, *84*, 3636–3645.
- (30) Condic-Jurkic, K.; Subramanian, N.; Mark, A. E.; O'Mara, M. L. The reliability of molecular dynamics simulations of the multidrug transporter P-glycoprotein in a membrane environment. *PLoS One* **2018**, *13*, No. e0191882.
- (31) Igaev, M.; Kutzner, C.; Bock, L. V.; Vaiana, A. C.; Grubmüller, H. Automated cryo-EM structure refinement using correlation-driven molecular dynamics. *eLife* **2019**, *8*, No. e43542.
- (32) Tieleman, D. P.; Marrink, S.-J.; Berendsen, H. J. A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. *Biochim. Biophys. Acta, Rev. Biomembr.* **1997**, *1331*, 235–270.
- (33) Yesylevskyy, S. O.; Rivel, T.; Ramseyer, C. The influence of curvature on the properties of the plasma membrane. Insights from atomistic molecular dynamics simulations. *Sci. Rep.* **2017**, *7*, No. 16078.
- (34) Li, Z.; Li, X.; Huang, Y.-Y.; Wu, Y.; Liu, R.; Zhou, L.; Lin, Y.; Wu, D.; Zhang, L.; Liu, H.; Xu, X.; Yu, K.; Zhang, Y.; Cui, J.; Zhan, C.-G.; Wang, X.; Luo, H.-B. Identify potent SARS-CoV-2 main protease inhibitors via accelerated free energy perturbation-based virtual screening of existing drugs. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 27381–27387.
- (35) Zou, J.; Yin, J.; Fang, L.; Yang, M.; Wang, T.; Wu, W.; Bellucci, M. A.; Zhang, P. Computational Prediction of Mutational Effects on SARS-CoV-2 Binding by Relative Free Energy Calculations. *J. Chem. Inf. Model.* **2020**, *60*, 5794–5802.
- (36) Williams, M. A.; Hall, D. R.; Hulswit, R. J.; Bowden, T. A.; Fry, E. E.; et al. Antibody evasion by the P. 1 strain of SARS-CoV-2. *Cell* **2021**, *184*, 2939.
- (37) Motozono, C.; Toyoda, M.; Zahradnik, J.; Ikeda, T.; Saito, A.; Tan, T. S.; Ngare, I.; Nasser, H.; Kimura, I.; Uriu, K.; et al. An emerging SARS-CoV-2 mutant evading cellular immunity and increasing viral infectivity. *BioRxiv* **2021**, *15*, No. 2021.04.02.438288.
- (38) Leach, A.; Ilca, F. T.; Akbar, Z.; Ferrari, M.; Bentley, E. M.; Mattiuzzo, G.; Onuoha, S.; Miller, A.; Ali, H.; Rabbitts, T. H. A tetrameric ACE2 protein broadly neutralizes SARS-CoV-2 spike variants of concern with elevated potency. *Antiviral Res.* **2021**, *194*, No. 105147.
- (39) 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* **2021**, *77*, 111.
- (40) Tian, F.; Tong, B.; Sun, L.; Shi, S.; Zheng, B.; Wang, Z.; Dong, X.; Zheng, P. N501Y mutation of spike protein in SARS-CoV-2 strengthens its binding to receptor ACE2. *eLife* **2021**, *10*, No. e69091.
- (41) Han, P.; Li, L.; Liu, S.; Wang, Q.; Zhang, D.; Xu, Z.; Han, P.; Li, X.; Peng, Q.; Su, C.; et al. Receptor binding and complex structures of human ACE2 to spike RBD from Omicron and Delta SARS-CoV-2. *Cell* **2022**, *185*, 630.
- (42) Zhang, X.; Wu, S.; Wu, B.; Yang, Q.; Chen, A.; Li, Y.; Zhang, Y.; Pan, T.; Zhang, H.; He, X. SARS-CoV-2 Omicron strain exhibits potent capabilities for immune evasion and viral entrance. *Signal Transduction Targeted Ther.* **2021**, *6*, No. 430.
- (43) Cameroni, E.; Bowen, J. E.; Rosen, L. E.; Saliba, C.; Zepeda, S. K.; Culap, K.; Pinto, D.; VanBlargan, L. A.; De Marco, A.; di Iulio, J.; et al. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* **2021**, *602*, 664–670.
- (44) Chan, K. K.; Tan, T. J.; Narayanan, K. K.; Procko, E. An engineered decoy receptor for SARS-CoV-2 broadly binds protein S sequence variants. *Sci. Adv.* **2021**, *7*, No. eabf1738.
- (45) Han, P.; Su, C.; Zhang, Y.; Bai, C.; Zheng, A.; Qiao, C.; Wang, Q.; Niu, S.; Chen, Q.; Zhang, Y.; et al. Molecular insights into receptor binding of recent emerging SARS-CoV-2 variants. *Nat. Commun.* **2021**, *12*, No. 6103.
- (46) Laffèber, C.; de Koning, K.; Kanaar, R.; Lebbink, J. H. Experimental evidence for enhanced receptor binding by rapidly spreading SARS-CoV-2 variants. *J. Mol. Biol.* **2021**, *433*, No. 167058.
- (47) Zhou, T.; Tsybovsky, Y.; Gorman, J.; Rapp, M.; Cerutti, G.; Chuang, G.-Y.; Katsamba, P. S.; Sampson, J. M.; Schön, A.; Bimela, J.; et al. Cryo-EM Structures of SARS-CoV-2 Spike without and with ACE2 Reveal a pH-Dependent Switch to Mediate Endosomal Positioning of Receptor-Binding Domains. *Cell Host Microbe* **2020**, *28*, 867–879.e5.
- (48) Sztain, T.; Ahn, S.-H.; Bogetti, A. T.; Casalino, L.; Goldsmith, J. A.; Seitz, E.; McCool, R. S.; Kearns, F. L.; Acosta-Reyes, F.; Maji, S.; et al. A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat. Chem.* **2021**, *13*, 963–968.
- (49) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci. Rep.* **2020**, *10*, No. 14991.
- (50) Allen, J. D.; Watanabe, Y.; Chawla, H.; Newby, M. L.; Crispin, M. Subtle influence of ACE2 glycan processing on SARS-CoV-2 recognition. *J. Mol. Biol.* **2021**, *433*, No. 166762.
- (51) Chen, F.; Liu, H.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T. Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking. *Phys. Chem. Chem. Phys.* **2016**, *18*, 22129–22139.
- (52) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discovery* **2015**, *10*, 449–461.
- (53) Sun, H.; Duan, L.; Chen, F.; Liu, H.; Wang, Z.; Pan, P.; Zhu, F.; Zhang, J. Z.; Hou, T. Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Phys. Chem. Chem. Phys.* **2018**, *20*, 14450–14460.
- (54) Miller, B. R., III; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.