JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

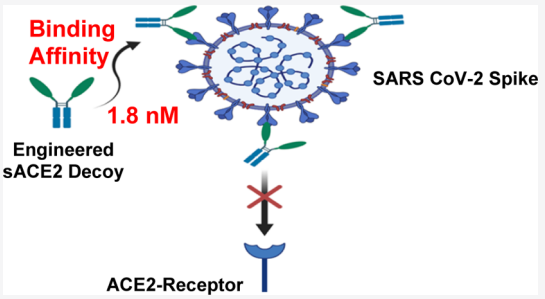
Computationally Designed ACE2 Decoy Receptor Binds SARS-CoV-2 Spike (S) Protein with Tight Nanomolar Affinity

Brandon Havranek, Kui K. Chan, Austin Wu, Erik Procko, and Shahidul M. Islam*

Cite This: https://doi.org/10.1021/acs.jcim.1c00783 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Even with the availability of vaccines, therapeutic options for COVID-19 still remain highly desirable, especially in hospitalized patients with moderate or severe disease. Soluble ACE2 (sACE2) is a promising therapeutic candidate that neutralizes SARS CoV-2 infection by acting as a decoy. Using computational mutagenesis, we designed a number of sACE2 derivatives carrying three to four mutations. The top-predicted sACE2 decoy based on the *in silico* mutagenesis scan was subjected to molecular dynamics and free-energy calculations for further validation. After illuminating the mechanism of increased binding for our designed sACE2 derivative, the design was verified experimentally by flow cytometry and BLI-binding experiments. The computationally designed sACE2 decoy (ACE2-**FFWF**) bound the receptor-binding domain of SARS-CoV-2 tightly with low nanomolar affinity and ninefold affinity enhancement over the wild type. Furthermore, cell surface expression was slightly greater than wild-type ACE2, suggesting that the design is well-folded and stable. Having an arsenal of high-affinity sACE2 derivatives will help to buffer against the emergence of SARS CoV-2 variants. Here, we show that computational methods have become sufficiently accurate for the design of therapeutics for current and future viral pandemics.



INTRODUCTION

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) responsible for coronavirus disease 2019 (COVID-19) is still wreaking havoc across the globe. At the time of writing this manuscript, there have been over 156 million cases and 3.3 million deaths worldwide.[1] While vaccines are rapidly advancing in clinical trials[2,3] and three vaccines have been approved by the FDA for emergency use in the United States,[4] there is still an unmet need to deliver therapeutic options to those infected, especially in hospitalized patients with moderate to severe disease in which monoclonal antibody drugs are less effective.

It has been well-characterized that the spike (S) glycoprotein of SARS-CoV-2 uses human angiotensin-converting enzyme 2 (ACE2) as a cell-entry receptor.[5,6] The S protein is a 1273 amino acid trimeric class I viral fusion protein that is proteolytically cleaved into S1 and S2 subunits that remain noncovalently bound in a prefusion conformation.[5,7] The S protein receptor-binding domain (RBD) is located in the S1 subunit and upon binding to ACE2, a conformational change occurs that causes S1 shedding and the proteolytic processing of S2 that is important for host cell−virus membrane fusion and cytosolic release of viral RNA.[8,9] The RBD of the S protein, containing residues 319−541, has been a target for many neutralizing antibodies (nAb's).[7,10,11] A few nAb's have shown great potency with neutralizing capability for multiple SARS-CoV-2 variants and more distantly related betacoronaviruses, but novel S mutations that have appeared in some cases allow for viral escape and loss of neutralization efficacy.[12] Most recently, a number of mutations have been identified in the S protein that increase virus transmission and likely allow partial immune escape.[13] Namely, N501Y increases binding with ACE2 20-fold[14] and H69-V70del deletes amino acids in an N-terminal domain epitope and has been found in viruses that have eluded the immune response in immunocompromised patients.[15] Another mutation, E484K, has provided SARS CoV-2 resistance to several monoclonal antibodies.[12,16] Emergency Use Authorization was withdrawn by the U.S. Food and Drug Administration for bamlanivimab as a montherapy due to the emergence of new virus variants.[17] Very recently, the BioNTech/Pfizer vaccine, BNT162b2, was tested against engineered SARS CoV-2 viruses containing mutations from the newly emerged United Kingdom (B.1.1.7 lineage) and South African (B.1.351 lineage) variants, N501Y, H69-V70del, D614G, and E484K.[18] Although these novel variants showed only small effects on vaccine neutralization capability,[18] clinically, it is still not clear how these mutations

Received: July 2, 2021

might impact vaccine efficacy, making broadly effective therapeutic options against SARS CoV-2 highly desirable.

Soluble ACE2 (sACE2) can be used as a therapeutic agent by acting as a decoy copy of ACE2, tricking the S protein of SARS-CoV-2 from binding to native ACE2 receptors and instead binding to the soluble decoy receptors (Figure S1); this may lead to a reduction in viral load. Most importantly, resistance to engineered ACE2 decoys will likely be rare because any mutation in the RBD that subsequently reduces binding affinity to our sACE2 decoy will also reduce binding affinity to the cell surface native ACE2 receptors, thus likely reducing or eliminating infectivity. In addition, these decoy receptors can be used against future outbreaks of SARS-associated betacoronaviruses.[14] Recombinant sACE2 has already been demonstrated to be safe in healthy subjects and patients with lung disease.[19,20] In addition to neutralizing SARS-CoV-2 infection, sACE2 through its enzymatic activity may directly treat respiratory distress pathophysiology that is often associated with COVID-19.[19] A case report by Zoufaly et al.[21] found that sACE2 did not impede the generation of nAb's and lead to a significant clinical improvement of a patient with severe COVID-19. Topline data from a phase 2 clinical trial of recombinant sACE2 in hospitalized COVID-19 patients demonstrated trends toward reduced death and a significant reduction in mechanical ventilation, associated with reduced viral load (Apeiron Biologics press release, March 12, 2021).

Due to the fact that human ACE2 has not evolved to recognize the SARS-CoV-2 S protein, it was hypothesized and shown that mutations within ACE2 may increase the affinity of RBD binding.[22] Multiple groups have since engineered sACE2 derivatives carrying a small number of mutations to increase affinity, achieving dissociation constants (K_D) in the sub-nanomolar range that rival affinity-matured monoclonal antibodies.[22−26] While most of these efforts have used in vitro selection by yeast or mammalian cell display, two groups have used computational design. In the first study, computational alanine scanning mutagenesis identified hot spots in the interface for focused saturation mutagenesis using the Rosetta macromolecular modeling suite,[27] yielding computationally designed sACE2 derivatives (carrying three or four mutations) with ∼10-fold improvement in affinity.[23] These were improved upon further by directed evolution. In the second study, sequence information from ACE2 orthologs was used to guide Rosetta-based modeling by restricting which mutations were preferred. Mutations shown experimentally to reduce ACE2/S affinity were also disallowed. The authors reported an even higher ∼100-fold affinity enhancement, albeit with a much higher mutational load of 8 substitutions.[24] The two computational design methods came to different solutions for affinity enhancement. We were interested in looking at this problem but without restricting the choice of substitutions based on hot spot predictions or evolutionary sequence information and furthermore providing an atomistic rationale for affinity enhancement through molecular dynamics (MD) simulations. As we show, we also arrive at another different set of mutations, emphasizing that there are multiple ways in which the affinity of ACE2 for the S protein can be improved for possible therapeutic purposes. Having a suite of unique, high-affinity ACE2 derivatives in hand may buffer against the emergence of diverse SARS-CoV-2 variants.

In silico saturation mutagenesis and de novo protein design using the Rosetta macromolecular modeling suite[27] have been applied extensively in protein engineering, antibody discovery, and therapeutic designs for SARS CoV-2 and other viruses.[23,28−34] Here, we use the Rosetta flex ddG method[35] to computationally design ACE2 for enhanced binding to the RBD of SARS CoV-2. The accuracy of the flex ddG method compares favorably to alternative methods and was shown to provide large improvements in predicting small-to-large side chain mutations, multiple simultaneous mutations, mutations in antibody−antigen interfaces, and stabilizing mutations.[35] To confirm the performance of flex ddG, we compared it with two other methods, FoldX[36] and SSIPe,[37] in which we found flex ddG to perform significantly better. Our computational results showed agreement with an experimental deep mutational scan (DMS)[22] as a benchmark data set. We combined, *in silico*, a number of the flex ddG predicted, enhancing mutations in three to four different mutation combinations. Using solely a computational approach, our designed ACE2 derivative with four mutations, **FFWF**, exhibited a ∼10-fold higher affinity than the wild type for the SARS CoV-2 RBD. Furthermore, using MD simulations, we probed the mechanism for the enhanced binding of the **FFWF** decoy which could be mainly attributed to increased van der Waals (VDW) and hydrophobic interactions.

RESULTS AND DISCUSSION

Computational Optimization of ACE2 for Tight SARS-CoV-2 RBD Binding. Recently, a DMS was conducted to identify mutations in ACE2 that lead to tighter binding to the SARS CoV-2 spike protein RBD.[22] Similarly, we set out to perform *in silico* saturation mutagenesis using Rosetta's flex ddG protocol to identify affinity enhancing mutations that improve ACE2 binding to the RBD and that could be combined to create a high affinity decoy receptor. For each mutation, 30,000 backrub steps were performed to account for protein conformational flexibility. In addition, backrub sampling was used to generate an ensemble of 35 models from the input structure of ACE2−RBD (PDB 6M0J).[38] Next, torsion minimization and side chain repacking were performed and the results across the ensemble were averaged to estimate the $\Delta\Delta G$ values. In particular, backrub has proved effective in sampling local side chain and backbone conformational changes to recapitulate natural protein plasticity, along with improving mutant side chain predictions.[39−41] Backrub captures protein dynamics that are often found on the millisecond/microseconds time scale,[39] suggesting that backrub is able to sample a sizable amount of natural conformational variability without the need for long-time scale MD. Due to the high computational cost of performing saturated mutagenesis using this protocol on a large complex such as ACE2−RBD, we used the DMS experiment to guide our search space.

First, flex ddG is optimized to predict the effect of mutations on protein−protein binding using Rosetta's all-atom energy function,[42−44] where the change in free energy upon mutation is reweighted using a nonlinear reweight scheme that is similar to the generalized additive model (GAM),[45] where Monte Carlo sampling is used to fit a sigmoidal function to the individual distributions of energy function terms. On a large test set, this was found to reduce the absolute errors between the *in silico* prediction and experimental values while also reducing the frequency of outliers.[35] With this in mind, the energy function is predominately used to predict the effect of mutations close to the protein's binding interface. Using a 4 Å distance cutoff, we focused on 17 ACE2 residues in close

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

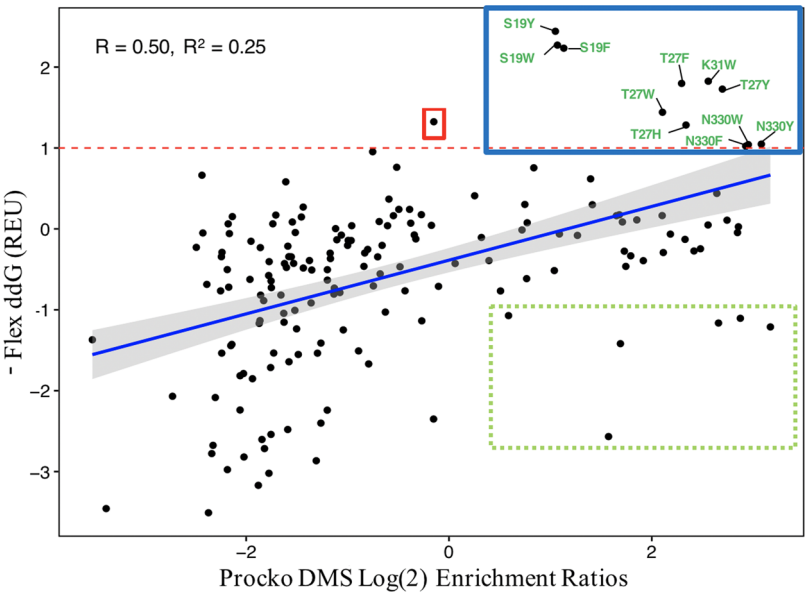


Figure 1. Flex ddG REU values vs log(2) enrichment ratios from the Procko DMS experiment.[22] Flex ddG values are reversed to match log(2) enrichment values. Positive values indicate tighter binding, while negative values are deleterious. The blue box contains mutations that enhance calculated binding by over 1 REU and are also enriched in the DMS. The red box contains mutation S19H, which is considered a false positive (mutation that is over 1 REU cutoff by flex ddG but deleterious by DMS). The light green dashed box contains outlier mutations Q24T, H34P, H34S, H34V, H34A, and E35A, predicted to be very deleterious by flex ddG but enriched for RBD binding by DMS. The dashed red line marks an arbitrary 1 REU cutoff, and a regression line is shown in blue with 95% confidence interval in light gray.

proximity to the RBD: S19, Q24, T27, F28, K31, H34, E35, E37, D38, Y41, Q42, L45, Y83, N330, K353, G354, and R357. To narrow down to the most important residues, we filtered the set based on data from the DMS. From the 17 identified residues, if at least one amino acid substitution for a particular residue was found to enrich binding according to the DMS experiment, it was included in the *in silico* mutagenesis scan. This filtering procedure lead to nine residues, S19, Q24, T27, K31, H34, E35, Y41, Q42, and N330, to target for ACE2 affinity optimization.

***In Silico* Prediction Agrees Well with DMS Data.** Aiming to show that *in silico* mutagenesis can be used effectively and efficiently for affinity optimization, we compared the results from our computational approach to the DMS experiment.[22] In total, the results from 171 mutations (9 residues × 19 amino acids) were compared to the log(2) enrichment ratios of the DMS data set (Figure 1).

Overall, the results from computational mutagenesis agree well with the experimental DMS data in which we find a moderate Pearson's correlation coefficient (R) of 0.50 and a moderate coefficient of determination (R^2 of 0.25). Using the same 171 mutations from DMS for comparison, flex ddG[35] was compared to both the fixed backbone approach from FoldX[36,46] and the custom energy function with evolutionary sequence profiles from SSIPe.[37] Flex ddG substantially outperformed both methods, with FoldX and SSIPe having Pearson's correlation coefficients of 0.35 and 0.26, respectively (Figure S2). Clearly, flex ddG, which takes into account conformational sampling and backbone flexibility that mimics natural protein plasticity, improves the correlation of mutant predictions with the DMS experiment.

Most importantly, the mutations in the blue box in Figure 1 show that the most enriching mutations found from computational mutagenesis also agree with the experimental data from DMS. In fact, two of the mutations with the greatest enhancement of calculated $\Delta\Delta G$, T27Y and N330Y, were previously used in combination to create a high-affinity decoy receptor.[22] The mutations with the greatest enhancement of calculated $\Delta\Delta G$ were exclusively substitutions for aromatic side chains, which can contribute disproportionately to VDW packing. We further confirmed the performance of the computational method by examining predicted false positives. Here, we define false positives as mutations in ACE2 predicted by Rosetta flex ddG to increase binding to the RBD but found to be deleterious by DMS. Applying a 1 Rosetta energy unit (REU) cutoff, there was only a single false-positive mutation S19H, shown in the red box in Figure 1. Thus, based on benchmarking to a DMS data set of single amino acid mutations, there is confidence that the *in silico* mutagenesis method can reasonably identify affinity-enhancing mutations for subsequent validation by wet lab experiments.

While the number of predicted false positives is very low, there remains room for the *in silico* methods to be improved. Notably, flex ddG failed to identify a number of enriching mutations (i.e., false negatives). In particular, mutations Q24T, H34P, H34S, H34V, H34A, and E35A, shown in the dotted light green box in Figure 1, were predicted by flex ddG to be deleterious (>1 REU loss of binding energy) but were enriched in the DMS experiment. It seems that this had less to do with the type of mutation but rather difficulty in predicting the effects for one specific residue, in particular, H34, which is located at the central bend in ACE2 helix 1. The protocol may have difficulty in accurately predicting conformation around helix deformations.

Combining Mutations to Explore Additive Effects. Even though single mutations can increase ACE2−RBD binding, mutations must be combined in order to reach affinities that rival potent monoclonal antibodies.[22,23] We identified S19Y, S19F, S19W, T27W, T27F, T27Y, T27H, K31W, N330W, N330Y, and N330F as the most beneficial

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

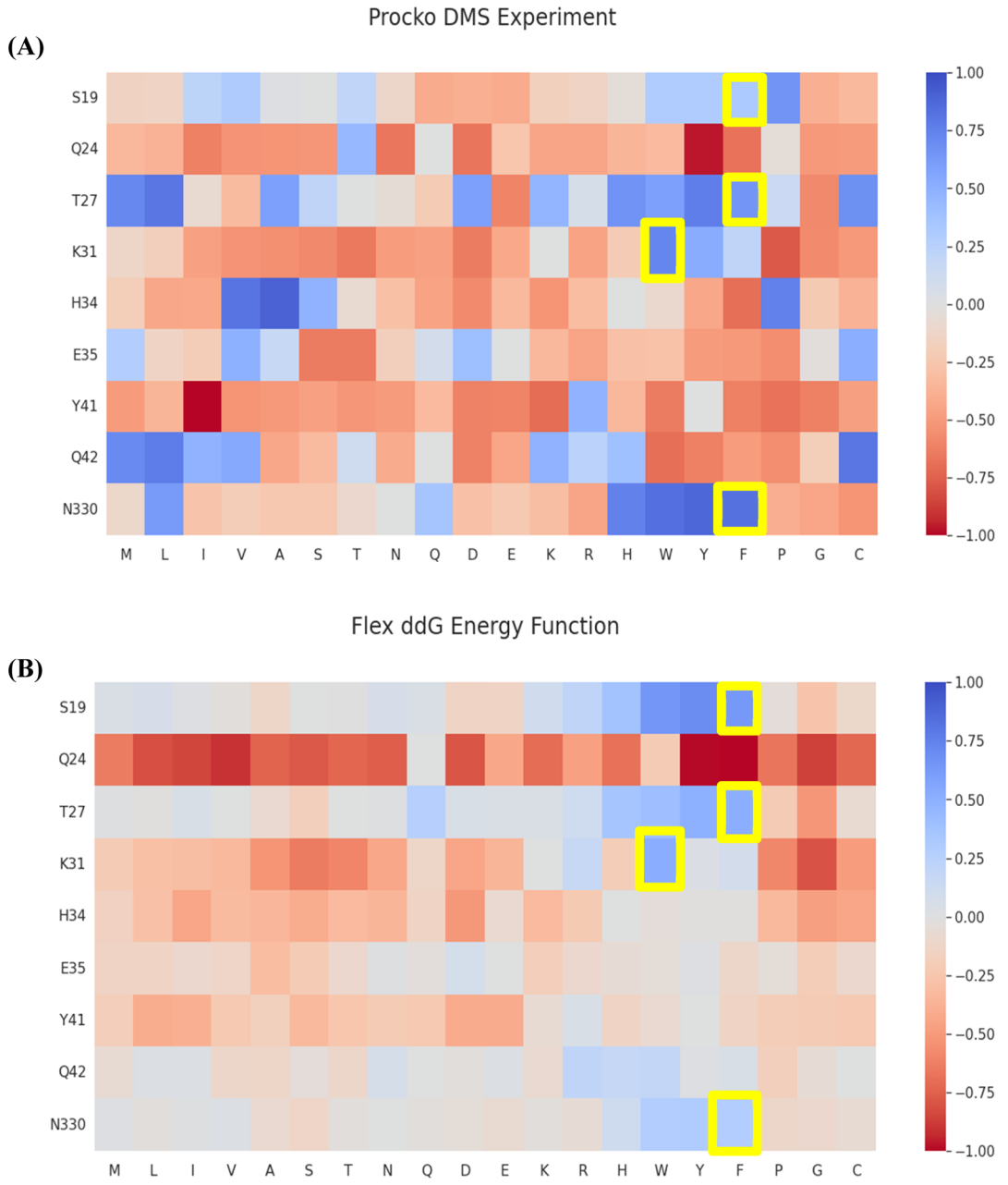


Figure 2. Heat maps showcasing binding scores for all 171 mutations from the saturated mutagenesis of 9 ACE2 residues in the Procko DMS experiment[22] (A) and Flex ddG method (B). All binding values are rescaled from −1 (decreased binding shown in red) to +1 (enriched binding shown in blue) for visualization purposes. Yellow outlined boxes identify mutations S19F, T27F, K31W, and N330F that were used to create the **FFWF** combinatorial mutant.

mutations above the 1 REU cutoff, with all of these mutations also agreeing with the DMS experiment (Figure 2). In particular, residues S19, T27, and N330 seem to be hot spot sites that emphasize how ACE2 affinity for SARS CoV-2 S can readily be enhanced, due to the fact that we found multiple mutations at these positions that can increase binding (Figure 2). Because flex ddG takes into account conformational effects around the site of the mutation, it may be more adept at accounting for epistasis and modeling the effects of multiple mutations introduced together. We combined predicted affinity enhancing mutations in 10 different sets of three or four mutations and computed their cumulative effects also using flex ddG (Table S1). The **FFWF** set containing mutations S19F, T27F, K31W, and N330F, shown in the outlined yellow boxes in Figure 2, was found to lead to the largest calculated increase in binding, −5.2 REU (Table S1). The predicted binding enhancement was close to double that of any single mutation alone.

MD Simulations Illuminate the Mechanism for Increased Binding of ACE2-FFWF. Aiming to quantify the specific interactions of **FFWF** that correspond to the increase in binding predicted by flex ddG, we performed 200 ns classical MD simulations for both the sACE2 wild-type system and the **FFWF** system (with S19F, T27F, K31W, and N330F mutations). From the 200 ns simulations of RBD-bound ACE2 proteins, we calculated the binding enthalpy using the molecular mechanics generalized Born surface area (MM/GBSA) method, and a pairwise energy decomposition was

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

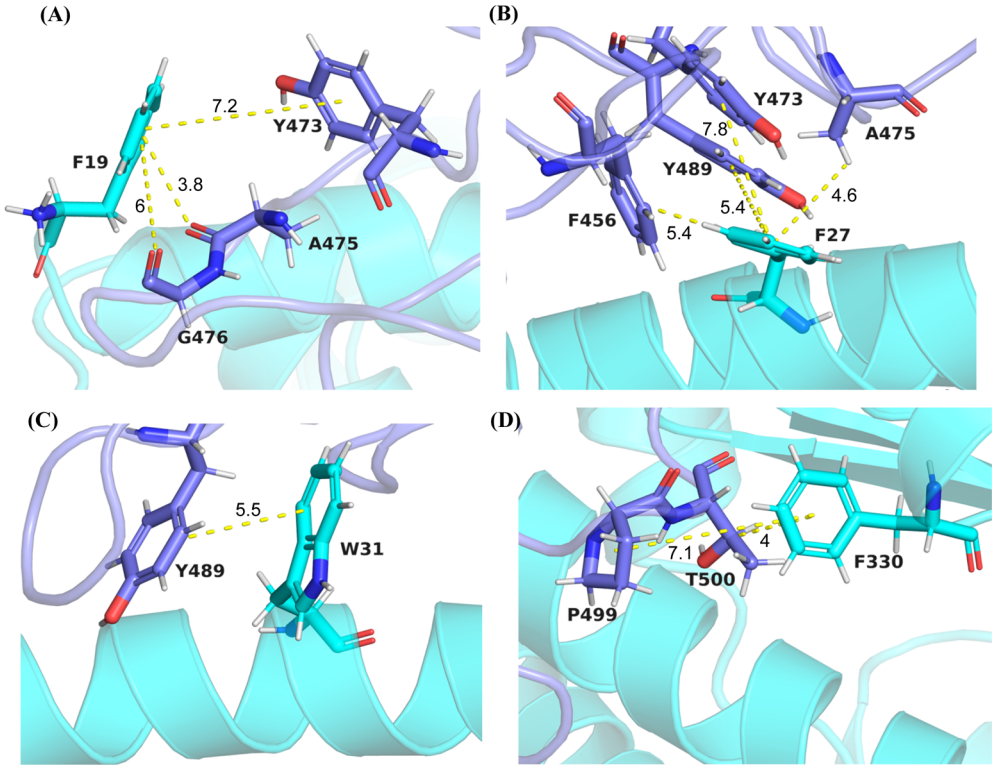


Figure 3. Interactions contributing to favorable interaction energy from mutated ACE2 residues (A) F19, (B) F27, (C) W31, and (D) F330 are shown in blue color and SARS CoV-2 RBD residues are shown in purple. Yellow dashed lines represent distances, in angstroms, from aromatic rings in ACE2 residues to the closest side chains in the RBD.

conducted to calculate the interaction energy between pairs of residues in the system. The binding enthalpy of RBD binding to wild-type ACE2 was calculated to be −21.8 ± 0.2 kcal/mol, with VDW interactions accounting for −65.6 kcal/mol, electrostatic energy −417.7 kcal/mol, polar solvation energy 471.1 kcal/mol, and nonpolar energy contribution −9.7 kcal/mol. Experimentally, multiple groups have measured the K_D of wild-type ACE2 and SARS-CoV-2 RBD as ∼20 nM[22,47,58] corresponding to −11 kcal/mol, indicating that the MM/GBSA calculations have overestimated the energies. However, this is to be expected since we did not compute the entropic cost of binding. We therefore consider these calculated energies as relative indicators for comparison purposes, not as absolute values and instead of free energy of binding, we refer to the binding values as "binding enthalpy" throughout the manuscript. In comparison, the **FFWF** system had binding enthalpy that was roughly 10 kcal/mol more favorable, with a calculated free energy of −31.4 ± 0.1 kcal/mol, which also agrees with the increased binding estimates from flex ddG. Not surprisingly, the increase in free energy was almost entirely contributed to the increase in VDW interactions, −79.2 kcal/mol. This is due to the fact that all four substitutions in the **FFWF** system were to larger aromatic side chains (phenylalanine or tryptophan) that will contribute to increased VDW interactions with the RBD. The electrostatic energy was calculated at −412.4 kcal/mol, polar solvation energy 471.0 kcal/mol, and a slight increase in nonpolar energy, −10.8 kcal/mol (Table S2).

Furthermore, the pairwise interaction energy decomposition allowed us to compare the contribution of binding for the S19, T27, K31, and N330 residues in ACE2 for the wild-type versus **FFWF** mutant. This provided insights into the differences between the wild-type and **FFWF** interactions with the RBD and quantitative information into how the mutations contributed to the overall increase in binding. We analyzed the interactions between ACE2 and the RBD that contributed to at least −0.50 kcal/mol of energy or higher. In the wild type, there were no favorable interactions for Ser19 contributing over −0.50 kcal/mol. After mutation to Phe19, we observed a number of favorable hydrophobic interactions including Phe19−Ala475 (−1.41 kcal/mol; the ACE2 residue is listed first, the RBD residue second), Phe19−Tyr473 (−0.65 kcal/mol), and Phe19−Gly476 (−0.62 kcal/mol) (Figure 3A) (Table S3).

Next, we analyzed Thr27 in the wild type, where we found a number of favorable interactions: Thr27−Phe456 (−1.58 kcal/mol), Thr27−Tyr489 (−1.55 kcal/mol), and Thr27−Tyr473 (−0.98 kcal/mol). After mutation to Phe27, the interaction energies with aromatic RBD residues Phe456 and Tyr473 were increased due to enhanced hydrophobic and pi-stacking interactions, and an additional contact pair, Phe27−Ala475 (−1.0 kcal/mol), was observed (Figure 3B). Six favorable interactions were found for Lys31 (Table S3). After mutation to Trp31, there were five favorable interactions. However, the calculated interaction energy between Trp31−Tyr489, at −3.87 kcal/mol, was approximately double of any interaction in the wild type (Figure 3C). Finally, ACE2-Asn330 in the wild type had one contact with RBD-Thr500 at −1.16 kcal/mol. After mutation to Phe330, we calculated a substantial increase in the interaction with Thr500 to −2.46 kcal/mol and an additional interaction with Pro499, −0.65 kcal/mol (Figure 3D). Our MM/GBSA with pairwise decomposition analysis suggests that the increase in binding for the **FFWF** mutant is largely due to VDW contributions from the larger side chains

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

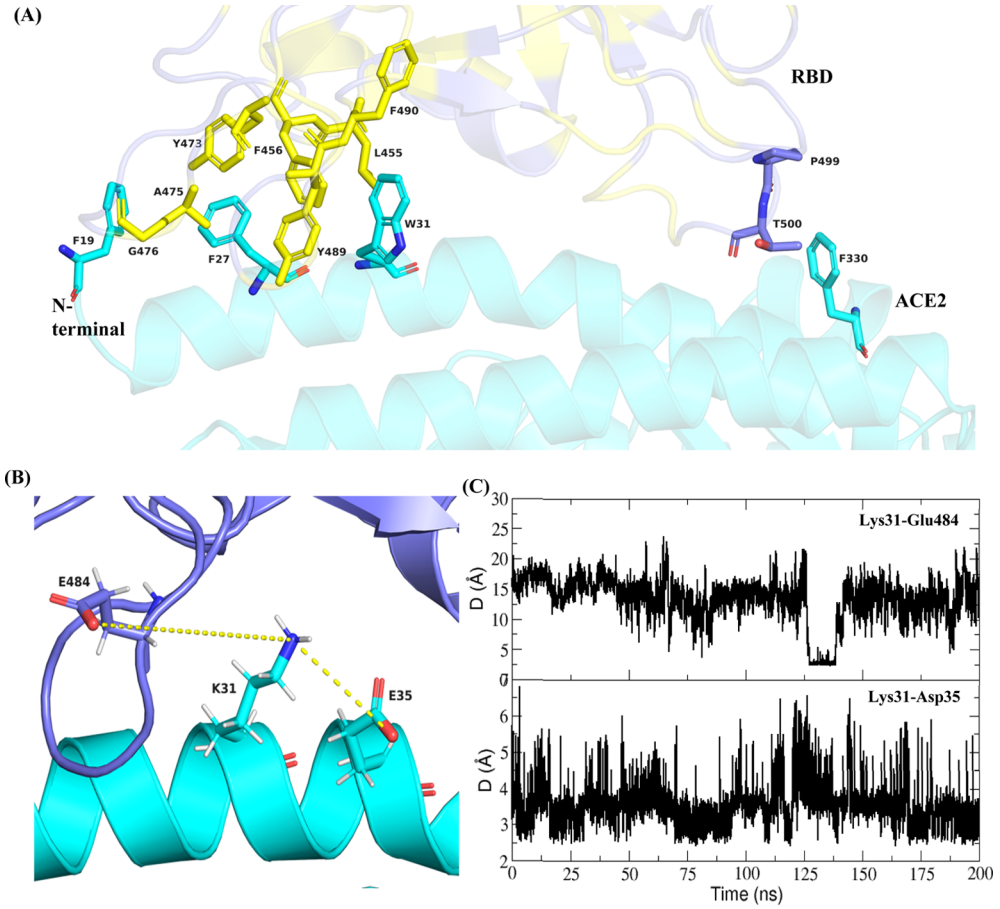


Figure 4. (A) Depiction of key interface interactions between ACE2 and RBD. ACE2 residues F19, F27, W31, and F330 in the **FFWF** mutant are shown in blue. Hydrophobic RBD residues are shown in yellow with all other nonhydrophobic residues shown in purple. (B) Salt bridges ACE2-Lys31 to ACE2-Asp35 and ACE2-Lys31 to RBD-Glu484 in wild-type ACE2−RBD, defined as the distance between oxygen atoms of acidic residues and nitrogen atoms of basic residues and within 3.2 Å in at least one frame of the simulation. ACE2 is shown in blue and RBD in purple. (C) Time-dependent distance *D* between Lys31−Glu484 and Lys31−Asp35 salt bridges in wild-type ACE2−RBD.

and a substantial increase in hydrophobic and pi-stacking interactions with RBD residues. It is especially intuitive that mutating polar ACE2 residues S19 and T27 to hydrophobic amino acids F19 and F27 will substantially enhance binding since these amino acids fall within a region of the binding interface that is dominated by hydrophobic residues in the RBD (Figure 4A).

Residue 27 of ACE2 sits in a pocket of hydrophobic RBD residues including Y473, F456, Y489, and A475 (Figure 4A), which explains why T27Y, T27F, and T27W were all predicted to be favorable mutations for increased binding (Figure 2), since hydrophobic and aromatic pi-stacking interactions with three of the aromatic RBD residues likely dominate. ACE2 residue 19 benefits from a larger nonpolar side chain that can take advantage of hydrophobic interactions with RBD residues G476, A475, and Y473 at the region of the binding interface near the ACE2 N-terminus (Figure 4A). Wang et al.[48] noted that the large role of hydrophobic interactions at the interface of ACE2−RBD is largely thought to contribute to SARS-CoV-2-enhanced binding to ACE2 compared to SARS-CoV. It is less intuitive why mutating K31 to Trp leads to increased binding because the K31W mutation disrupts two salt bridges between Lys31−Asp35 in ACE2 and Lys31−Glu484 in ACE2−RBD (Figure 4B), albeit these are solvent-exposed salt bridges where the charged groups will have competition with hydrogen bonds to water. However, ACE2-W31 is able to interact with RBD-Y489 within 6 Å (Figure 3C), and the calculated increase in binding is mainly due to pi-stacking interactions in the W31−Y489 pair that overshadows the loss of the salt bridge, making the W31−Y489 interaction more favorable by 1.9 kcal/mol over the interaction in the wild type. In fact, the Lys31-Glu484 salt bridge in the wild type appears to be weak, as it persists for only ∼10 ns in 125−135 ns simulation, with the two residues being separated by ∼15 Å throughout most of the simulation (Figure 4C). Unlike the other three mutations that reside near the ACE2 N-terminus, the N330F mutation resides at the opposite end of the interface and sits behind the major alpha helix interacting portion of ACE2 (Figure 4A). Additionally, ACE2-N330F contacts a relatively nonhydrophobic surface of the RBD in what is known as the binding loop (Figure 4A). The N330F mutation likely improves packing against the aliphatic portion of RBD-Thr500 (Figure 3D). Furthermore, the N330F mutation acquires an additional interaction not seen in the wild type with RBD-Pro499 (Figure 3D and Table S3). Proline and aromatic residues, such as F330, can interact favorably due to both the hydrophobic effect and interaction between the pi aromatic ring and polarized C−H bonds, which is known as the CH/π interaction.[49] In addition, since the proline side chain is conformationally restricted, interactions with aromatic residues take place with minimal entropic or steric penalty.[49] Many studies have reported on the role CH/π interactions play

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(A)

Wildtype ACE2				FFWF			
Donor/Acceptor Pair	f(%)	r (Å)	θ (deg)	Donor/Acceptor Pair	f(%)	r (Å)	θ (deg)
$Lys353^{O}/Gly502^{H}$	90%	2.92	162.6	$Lys353^{O}/Gly502^{H}$	90%	2.91	162.6
$Asp355^{O}/Thr500^{H}$	54%	2.72	162.6	$Asp355^{O}/Thr500^{H}$	71%	2.72	162.7
$Tyr83^{OH}/Asn487^{H}$	47%	2.77	157.9	$Tyr83^{OH}/Asn487^{H}$	64%	2.77	157.6
$His34^{N}/Tyr453^{H}$	38%	2.90	160.2	$His34^{N}/Tyr453^{H}$	28 %	2.94	157.3
$Glu35^{O}/Gln493^{H}$	26%	2.95	160.6	$Gln24^{N}/Ala475^{O}$	16%	2.98	157.7
$Glu35^{O}/Gln493^{H}$	21%	2.96	159.1	$Tyr41^{OH}/Thr500^{H}$	14%	2.96	160.1
$Tyr41^{OH}/Thr500^{H}$	16%	2.90	158.4	$Glu35^{O}/Tyr505^{H}$	14%	2.81	157.5
$Gln24^{N}/Ala475^{O}$	11%	3.00	157.4	$Glu35^{O}/Tyr505^{H}$	13%	2.80	162.4
$Gln24^{N}/Asn487^{O}$	10%	3.04	158.5	$His34^{N}/Ser494^{H}$	10%	3.08	157.7

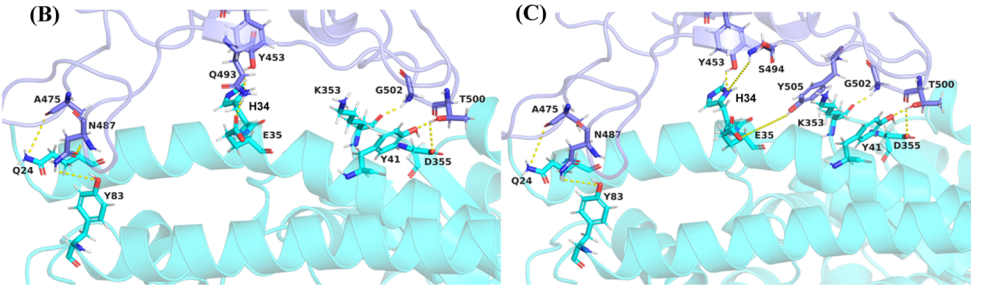


Figure 5. (A) Average lengths (r), angles (θ), and occupancy (f) of intermolecular hydrogen bonds between wild-type ACE2 and RBD, as well as between the **FFWF** mutant and RBD from 200 ns MD simulations. Only occupancies (f) over 10% are reported. Hydrogen-bond distances are given with respect to the heavy atoms. (B) Visual representation of hydrogen bonds found from 200 ns MD simulation in wild-type ACE2−RBD. (C) Visual representation of hydrogen bonds found from 200 ns MD simulation in the **FFWF** mutant. ACE2 is shown in blue and RBD is shown in purple. Hydrogen bonds are shown with yellow dotted lines.

in stabilizing not only the protein structure but also protein−protein interactions.[50−54] Therefore, the ACE2−F330−RBD−P499 interaction may be very important for increased binding.

Although we found that enhanced binding for the **FFWF** ACE2 mutant largely derives from increased VDW and hydrophobic interactions, mutating the S19, T27, K31, and N330 residues could have unintended consequences on hydrogen bonding between ACE2 and the RBD. Therefore, we compared the hydrogen bonding between ACE2 and the RBD in both the wild-type and **FFWF** systems. H-bonding was classified as bonds between H and N, O, and F within 3.5 Å and bond angle greater than 120°. No remarkable differences were observed between the wild-type and **FFWF** systems. Both systems had nine total hydrogen bonds with occupancies of 10% or greater during the 200 ns MD simulations. In both systems, the hydrogen bonding was very similar in the contact pairs, total number of H bonds, and occupancies (Figure 5). No hydrogen bonding was gained nor lost when the four ACE2 residues were mutated, again consistent with the gain in binding being mainly due to hydrophobic interactions. Moreover, since none of these residues were participating in H-bonding with the RBD to begin with, this may explain why we were able to mutate them to aromatic residues with no deleterious effect on calculated binding.

Moreover, a free-energy landscape (FEL) was plotted along the first two eigenvector components, PC1 and PC2, for residues in the binding interface, 19−87 and 325−330 in ACE2 plus 438−506 in the RBD (Figure S3D), in order to extract the common states throughout the simulation (Figure S3A,B). It is worth noting that 200 ns is not nearly enough time to explore all protein conformational states. Nonetheless, the FEL provides some insights into the dynamics of the wild-type vs **FFWF** systems. In the wild-type system, the ACE2−RBD binding interface existed in mainly four populated states, which suggests that the wild-type binding interface exists in various conformations during the simulation (Figure S3B). In comparison, the **FFWF** mutant's binding interface was clustered in mainly one to two prominent states, suggesting that the mutations might "lock" the interface in a more dominant state with tighter binding (Figure S3A). This phenomenon is also seen in reference to the probability density of the distances between the binding interface residues in the ACE2−RBD interface (Figure S3C). The wild-type system displayed a less-prominent peak and relative wider distribution in distances between interface residues, suggesting that the conformation at the binding interface was more flexible. In comparison, the **FFWF** system had a much more narrow peak with probability density shifted toward lower distances, which again shows that the four mutations helped to stabilize the interface with the RBD and ACE2 more tightly bound with less flexibility and improved interface rigidity.

In Vitro Validation of the FFWF ACE2 Mutant. It is plausible that our designed **FFWF** system is likely to have increased binding with the RBD based on the results from flex ddG and MD simulations with both MM/GBSA free-energy calculations and the pairwise decomposition. Therefore, we set out to verify whether the results generated from our computational approach also agree with in vitro binding experiments (Figure 6).

The binding of soluble RBD-sfGFP to full-length myc-ACE2 expressed on the surface of human cells was qualitatively analyzed using dual-color flow cytometry, and we observed

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

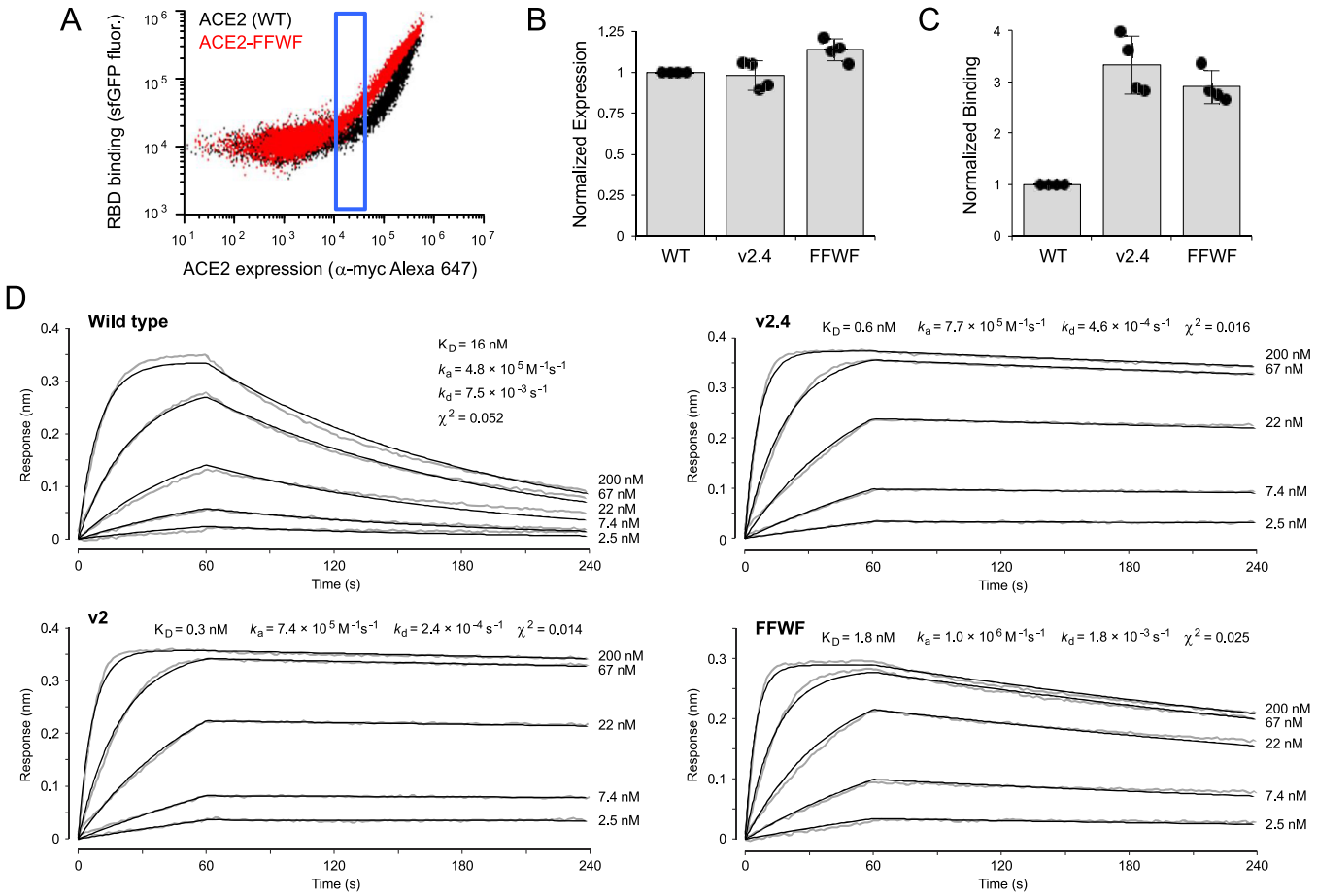


Figure 6. ACE2 mutant **FFWF** binds SARS-CoV-2 RBD with tighter affinity. (A) Expi293F cells were transfected with full-length myc-ACE2 and analyzed by flow cytometry (representative data from $n = 4$). Anti-myc Alexa 647 fluorescence for the full population was used to assess surface ACE2 expression levels. The fluorescence signal of bound RBD-sfGFP was measured within the blue gate to control for differences in expression levels. (B) Normalized surface expression of myc-ACE2 mutants by flow cytometry ($n = 4$, mean ± SD). (C) Normalized binding of RBD-sfGFP measured by flow cytometry ($n = 4$, mean ± SD). (D) Purified $sACE2_2$-IgG1 proteins were immobilized and association (time 0−60 s) and dissociation (time 60−240 s) of soluble, monomeric RBD-8h was measured by BLI. RBD-8h concentrations are indicated to the right of each response trace. Responses are gray, and fitted curves are black.

increased binding for the **FFWF** mutant (Figure 6A). The computationally designed **FFWF** mutant was also compared to the wild-type and two previously described ACE2 variants (v2 and v2.4, with four and three mutations, respectively) that were derived from solely an experimental DMS approach (Figure 6B−D).[22] The normalized binding of RBD-sfGFP measured via flow cytometry was observed to be slightly less than ACE2.v2.4 but still much greater than the wild type (Figure 6C). We also found that the surface expression for the myc-ACE2 **FFWF** mutant was slightly greater than both the wild-type and v2.4 proteins, suggesting that the mutational load of **FFWF** does not negatively impact protein folding (Figure 6B). Finally, purified dimeric $sACE2_2$-IgG1 proteins were immobilized to an anti-IgG1 biosensor and incubated with monomeric RBD-8h to quantitatively measure monovalent affinity by biolayer interferometry (BLI; Figure 6D). The dissociation constant K_D of RBD for wild-type ACE2 was measured at 16 nM, which is in agreement with another report.[58] The computationally designed **FFWF** mutant has ninefold higher affinity than the wild-type for RBD, with the increased binding mainly attributed to the slower off-rate. The K_D of 1.8 nM for the **FFWF** mutant is in line with SARS CoV-2 monoclonal antibodies[55−58] and other therapeutics such as nanobodies[59−62] that potently neutralize SARS CoV-2. Although **FFWF** did not bind as tightly as v2 or v2.4 designed mutants, these designs were constructed after multiple rounds of experimental mutagenesis. Nonetheless, we demonstrate that using a computational approach can be feasible for affinity optimization and combined with MD can reduce the number of wet lab experiments, while also achieving a low nanomolar affinity binder.

CONCLUSIONS

In this work, we used a systematic approach that involved computational mutagenesis to design improved sACE2 binders. The top-predicted decoy based on the *in silico* mutagenesis scan, **FFWF**, was subjected to MD and free-energy calculations to verify the design. After investigating the molecular mechanism for enhanced binding, the **FFWF** mutant was verified experimentally by flow cytometry and BLI experiments. The computationally designed **FFWF** decoy had a binding affinity of 1.8 nM that was ninefold tighter than the wild-type and is on par with other anti-SARS-CoV-2 monoclonal antibodies,[57] miniproteins,[30,63] and nanobodies.[60] There have been various strategies for dealing with ACE2 peptidase activity for designed decoys, with some groups

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

opting to mutate ACE2 residues to abolish peptidase activity.[23,64] In our design, catalytic residues are left intact, which likely means that the **FFWF** decoy retains at least some catalytic activity. It is envisioned that ACE2 peptidase activity may be beneficial for treating lung injury from SARS-CoV-2,[65] supported by promising results in a phase 2 clinical trial where severe COVID-19 patients were administered catalytically active, wild-type, recombinant sACE2 (Apeiron Biologics press release, March 12, 2021). ACE2 decoys also present numerous other advantages such as the low likelihood of being impacted by escape mutations that develop in the spike protein and cross-neutralization of other betacoronaviruses that utilize ACE2 as an entry receptor.[14] It is possible that **FFWF** or a similar sACE2 derivative, delivered either intranasally or intravenously, may be suitable as a prophylactic or therapeutic treatment. Computational approaches present a promising approach for the design and optimization of biologics and can be employed to screen a large amount of designs and are especially beneficial in cases where combinatorial diversity in experimental libraries is too large for in vitro selection methods. The approach we have described here further emphasizes how computational methods have become sufficiently accurate for the design of therapeutics for future viral pandemics and against current protein targets.

MATERIALS AND METHODS

***In Silico* Mutagenesis.** Flex ddG[35] was performed using the recommended parameters in the flex ddG tutorial in the GitHub repository (https://github.com/Kortemme-Lab/flex_ddG_tutorial). Briefly, the RBD-ACE2 crystal structure (PDB ID 6M0J)[38] was minimized. All minimizations were performed with harmonic restraints on pairwise atom distances to their values in the input structure. Restraints were added to all alpha carbon pairs within 9 Å using a harmonic score potential with a standard deviation parameter set to 0.5 Å and were added to the Rosetta score function with a term weight of 1.0. The backrub method[30] in Rosetta was used to create an ensemble of 35 different models. Backrub performs local sampling of backbone and side chain degrees of freedom of pivot residues, defined as those with neighbor atoms (C-β) within 8 Å of the mutated position. In total, 35,000 backrub steps were performed for each mutation. For all 35 generated ensemble models, the Rosetta "packer" was used to optimize side chain conformations using discrete rotameric sampling[42] and simulated annealing[66] for the wild-type sequence. In conjunction, the "packer" optimized the side chain conformations for the introduced mutations in all 35 ensembles. Again, all 35 models with both wild-type and mutant sequences were minimized to optimize side chain and backbone torsions and the same restraints from above were applied. The wild-type and mutant complexes, along with the unbound partners, were scored individually for each ensemble, and the $\Delta\Delta G$ score is calculated from the average of the 35 ensembles. Flex ddG uses the Rosetta Talaris all-atom energy function,[42−44] and the $\Delta\Delta G$ values are reweighted using a nonlinear reweight scheme that is similar to the GAM to improve correlation with experimental values.[45]

The FoldX[36,67] (http://foldxsuite.crg.eu/) version 5.0[46] energy function was used to estimate the interaction energy between ACE2−RBD mutations. We used a Python wrapper, MutateX[68] (https://github.com/ELELAB/mutatex), to support the systematic substitution of wild-type residues to all possible 20 standard amino acids using multithreading calculations. Briefly, the *RepairPDB* module from FoldX was applied to optimize the conformation of the crystal structure by repairing residues with unfavorable torsion angles, VDW clashes, and total energy. Asn, Gln, and His residues are flipped 180° to prevent incorrect rotamer assignment. Residues with bad energies were identified and mutated to themselves along with their neighbors to explore different rotamer combinations in order to find new energy minima. The *BuildModel* from FoldX independently mutates each residue at every position. The *AnalyseComplex* in FoldX is used to estimate the $\Delta\Delta G$ of interaction between all possible ACE2 mutations and the RBD relative to the original wild-type residue. FoldX calculations were carried out from PDB 6M0J,[38] which is the structure of SARS-CoV-2 RBD complexed with its receptor, ACE2. Five independent runs (numberOfRuns set to 5) were carried out and averaged. All calculations were performed using a temperature of 298 K, an ionic strength of 0.05 M, and a pH of 7.

SSIPe[37] binding affinity change estimator was accessed from https://zhanglab.dcmb.med.umich.edu/SSIPe/. Briefly, SSIPe is a method to calculate binding affinity changes in protein−protein interactions upon mutation at the protein−protein interface. SSIPe combines structural and sequence profiles with a physical energy function, EvoFF,[69] in order to predict the impact of mutations on protein−protein interactions in the interface.

MD Simulations. MD simulations for the wild-type and **FFWF** systems were performed using the AMBER 18 package.[70] All MD simulations were performed using PDB 6M0J,[38] which contains the SARS-CoV-2 RBD complexed with its receptor human ACE2 solved with X-ray diffraction at 2.45 Å resolution. The systems were prepared using the solution builder from CHARMM-gui[71] utilizing the CHARMM36m additive force field for proteins.[72,73] The systems were fitted using a rectangular water box that had a radius of 10 Å from the complex's surface and solvated using a series of TIP3P water molecules. In order to mimic physiological conditions, 0.15 M KCl ions were included using the Monte-Carlo ion placing method. A steepest decent energy minimization was carried out using CPU for 2500 cycles and then the conjugate gradient algorithm was used for 5000 cycles. All systems were subjected to an equilibration period of 2 ns under (canonical ensemble) *NVT* conditions, where the volume (V), amount of substance (N), and temperature (T) are conserved. In order to restrain each of the complexes during equilibration, a positional restraint of 1 kcal/mol was implemented. The temperature was set at 303.15 K and was maintained using Langevin dynamics.[74] After minimization and equilibration of the systems were conducted, long MD simulations of 200 ns were performed for each system. Long MD simulations were performed under *NPT* conditions where the temperature was kept at 303.15 K and pressure at 1 atm to mimic experimental conditions. A friction coefficient, γ, of 1.0 ps^{-1} was used for the Langevin thermostat, and the pressure was held constant with the Monte Carlo barostat. Integration was performed using a leap-frog algorithm with a 2 fs time step. All bonds involving hydrogen atoms were constrained to their equilibrium values using SHAKE.[75] Periodic boundary conditions were applied to all simulations with a nonbonded cutoff of 10 Å, and the particle-mesh-Ewald method[76] was used to treat all long-range interactions. All equilibration and production MD simulations were performed

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

using CUDA acceleration (*pmemed.cuda*)[77] inside of AMBER 18 using NVIDIA P100 GPUs.

Free-Energy Calculations. Binding free energies between ACE2−RBD for the wild-type and **FFWF** systems were performed using the MM/GBSA method using the MMGBSA python script[78] inside of Ambertools20. The Ambertools20[79] ante-MMGBSA python script[78] was used to create the complex, receptor, and ligand parameter files from the solvated parameter files used in MD simulations. For MM/GBSA, the binding free energies were calculated from 1000 independent frames using the last 120 ns from the 200 ns explicit-solvent MD simulations. The first 80 ns were discarded for equilibration. The Generalized born method, developed by Onufriev and company,[80] was set to igb = 2 to estimate the solvation energy. The radii were set to mbondi2 and the salt concentration was set to 0.15 M. Additionally, the dielectric constant of solvent and dielectric constant of solute were set to 78.5 and 1.0, respectively, which are Amber default and recommended values. Wang et al.[81] investigated 21 protein−protein complexes to assess the performance of MM/GBSA under different conditions and also suggested using a solute dielectric constant of 1.0. Finally, the solvent-accessible surface area (SASA) was calculated using γ = 0.0072 kcal/mol/Å^2 and β = 0.0 kcal/mol, respectively.

The binding free energy (ΔG_{bind}) between the RBD and ACE2 is calculated as

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM,gas} + \Delta G_{sol} - T\Delta S \quad (1)$$

where $\Delta E_{MM,gas}$, ΔG_{sol}, and $-T\Delta S$ are the changes of the gas phase molecular mechanics energy, the solvation free energy, and the conformational entropy upon binding, respectively. The $\Delta E_{MM,gas}$ is the sum of the internal energy ($\Delta E_{internal}$) arising from bond, angle, and dihedral interactions, nonbonded electrostatic energy (ΔE_{ELE}), and VDW energy (ΔE_{VDW}).

$$\Delta E_{MM,gas} = \Delta E_{int} + \Delta E_{ELE} + \Delta E_{VDW} \quad (2)$$

In MM/GBSA calculations, the ΔE_{int} is cancelled since the complex, receptor, and ligand parameter files are created from the same trajectory. The solvation free energy (ΔG_{sol}) is calculated from the sum of the polar solvation energy (ΔG_{GB}) using the generalized Born model, and the nonpolar energy is calculated based on the solvent-accessible surface area (ΔG_{surf}) according to the LCPO algorithm.[82]

$$\Delta G_{sol} = \Delta G_{GB(PB)} + \Delta G_{surf} \quad (3)$$

The conformational entropy change, $-T\Delta S$, is usually computed by normal-mode analysis on a set of conformational snapshots taken from MD simulations. In this case, contribution from entropy is neglected because of its large computational cost and low prediction accuracy. Moreover, entropy is likely to be similar for all systems as they differ by only several mutations.

Pairwise Decomposition Energy Analysis. A pairwise energy decomposition was performed to calculate the interaction energy between pairs of residues in the wild-type and **FFWF** systems. The decomposition scheme was decomposed on a pairwise per-residue basis using idecomp = 4. Decomposition energy schemes were developed by Gohlke and co-workers[83] and energy terms are decomposed using the following equation

$$\Delta E^{Decomp} = \sum_{\substack{j\in COMPLEX,\\ j\in RECEPTOR}} (E^{COMPLEX}(i,j) - E^{RECEPTOR}(i,j)) + \sum_{\substack{j\in COMPLEX,\\ j\in LIGAND}} (E^{COMPLEX}(i,j) - E^{LIGAND}(i,j)) \quad (4)$$

Where the first and second terms are the average contribution over the snapshots, i, from MD simulations in residues j on the receptor (ACE2) and ligand (RBD), respectively. The $E_{GBTOTAL}(i,j)$ is a contribution of gas phase and solvation energies

$$E_{GBTOTAL}(i,j) = E_{GAS}(i,j) + E_{GBSOLV}(i,j) = E_{VDW}(i,j) + E_{ELE}(i,j) + E_{GB}(i,j) + E_{SURF}(i,j) \quad (5)$$

The per-residue decomposition energy includes nonbonded electrostatic energy (ΔE_{ELE}), VDW energy (ΔE_{VDW}), polar solvation energy (ΔG_{GB}) from the generalized Born model, and the nonpolar energy. Entropy is not included in the decomposition method.

Plasmids. For testing full-length ACE2 protein by flow cytometry, **FFWF** (S19F, T27F, K31W, and N330F) and v2.4 (T27Y, L79T, and N330Y) mutations were introduced by overlap extension PCR into pCEP4-myc-ACE2 (Addgene #141185) to generate pCEP4-myc-ACE2-**FFWF** and pCEP4-myc-ACE2.v2.4, respectively. The extracellular region was then PCR-amplified and inserted into the NheI-BamHI sites of plasmid pcDNA3-sACE2-WT(732)-IgG1 (Addgene #154104) to encode from N- to C-terminus: an HA signal peptide, myc epitope tag, ACE2-**FFWF** residues 19−732, gly−ser linker, and IgG1 Fc. Plasmids for expression of RBD-8h (Addgene #145145), RBD-sfGFP (Addgene #141184), wild-type sACE2$_2$.-IgG1 (Addgene #154104), sACE2$_2$.v2-IgG1 (Addgene #154105), and sACE2$_2$.v2.4-IgG1 (Addgene #154106) are previously described.[22]

Flow Cytometry. Expi293F cells (Thermo Fisher Scientific) were grown at 37 °C, 125 rpm, 8% CO_2 in Expi293 Expression Medium (Thermo Fisher Scientific). Cells at 2 × 10^6/mL were transfected using ExpiFectamine 293 (Thermo Fisher Scientific) with 300 ng pCEP4-myc-ACE2 plasmid per mL of culture. Cells were collected 27 h later by centrifugation (600*g*, 60 s), washed with Dulbecco's phosphate buffered saline (PBS) supplemented with 0.2% bovine serum albumin (BSA), and incubated for 30 min on ice in PBS−BSA with 1/50 dilution of expression medium containing RBD-sfGFP (prepared as previously described[22]) plus 1/250 dilution anti-myc Alexa 647 (clone 9B11, Cell Signaling Technology). Cells were washed twice, resuspended in PBS−BSA, and analyzed on a BD Accuri C6 using instrument software. The main cell population was gated by forward and side scatter properties, followed by measurement of mean Alexa 647 fluorescence to determine changes in myc-ACE2 surface expression. To assess RBD-sfGFP binding, cells were further gated based on Alexa 647 fluorescence to control for any differences in myc-ACE2 expression (see blue gate in Figure 6A) and the mean sfGFP fluorescence was then recorded. In all cases, background fluorescence of cells transfected with empty pCEP4 vector was subtracted (i.e., Δ mean fluorescence was determined) prior to normalization against the wild-type ACE2 sample.

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Protein Purification. ExpiCHO-S cells (Thermo Fisher Scientific) were grown in ExpiCHO Expression Medium (Thermo Fisher Scientific) at 37 °C, 125 rpm, 8% CO_2 and were transfected with plasmid DNA (500 ng/mL of culture) using ExpiFectamine CHO (Thermo Fisher Scientific). After 18−22 h, temperature was dropped to 33 °C and ExpiFectamine CHO Enhancer (6 μL per mL of culture) plus ExpiCHO Feed (240 μL per mL of culture) (Thermo Fisher Scientific) were added. An additional 240 μL of ExpiCHO Feed per mL of culture was added on day 5. CO_2 was decreased to 7% on day 9 and to 6% on day 11. On day 14, the expression medium was cleared of cells (800*g*, 4 °C, 10 min) and the pH was adjusted to ∼7.5 with 1 M Tris base. Particulates were removed by a higher speed centrifugation step (15,000*g*, 4 °C, 20 min). A total of 2 mL of PBS-equilibrated KanCapA resin (Kaneka Corporation, equilibrated in PBS) was added for each 100 mL of medium supernatant. IgG1-fused protein was captured in the batch mode for 2 h, 4 °C, under rotation. The sample was flowed through a chromatography column to capture the resin, washed with 10 column volumes (CV) PBS, and protein-eluted with 4 CV 60 mM sodium acetate pH 3.7. The acidic eluate was rapidly neutralized using a collection tube prefilled with 2 CV 1 M Tris pH 8.0, followed by further adjustment of the pH to 7 with 1 to 2 CV 1 M Tris base. The protein was concentrated with a centrifugal filtration device (Millipore, 50 kD MW cutoff) and separated on a HiLoad Superdex 200 pg 16/600 column (GE Healthcare) equilibrated in PBS. Protein fractions were pooled and concentrated, and aliquots were snap frozen in liquid N_2 and stored at −80 °C. SARS-CoV-2 RBD (Wuhan isolate) with an 8his affinity tag was expressed in Expi293F cells from plasmid pcDNA3-SARS-CoV-2-S-RBD-8his (Addgene #145145) and purified as previously described using Ni-NTA affinity resin followed by size exclusion chromatography.[22]

BioLayer Interferometry. Anti-human IgG biosensors (Molecular Devices) were equilibrated in assay buffer (10 mM HEPES pH 7.6, 150 mM NaCl, 3 mM EDTA, 0.05% polysorbate 20, and 0.5% nonfat dry milk), and $sACE2_2$-IgG1 proteins (100 nM in assay buffer) were immobilized on the biosensor surface for 600 s. Loaded sensors were dipped in assay buffer containing the indicated concentrations of RBD-8h (sequence from SARS-CoV-2 Wuhan isolate) and transferred back to buffer for dissociation. Data were collected on an Octet RED96a and analyzed using instrument software (Molecular Devices) with a 1:1 binding model (global fit).

DATA AND SOFTWARE AVAILABILITY

All data are available upon reasonable request. MM/GBSA and pairwise energy decomposition was performed using AmberTools 20 with academic license. MD simulations were conducted using Amber 18 under license at Extreme Science and Engineering Discovery Environment (XSEDE) COMET GPU at SDSC. Mutagenesis using Flex ddG was performed with open source software and python scripts provided in GitHub tutorial (https://github.com/Kortemme-Lab/flex_ddG_tutorial) and required the use of Rosetta (v 2020.08) under academic license. Mutagenesis using FoldX was performed with a mutatex open source software and python scripts provided in GitHub tutorial (https://github.com/ELELAB/mutatex) and required the use of FoldX (Suite 5.0) under academic license. Mutagenesis using SSIPe was performed from Zhang Lab hosted on their web server at the University of Michigan (https://zhanglab.dcmb.med.umich.edu/SSIPe/) available to all users. R software (v 3.6.3) was used to make correlation plots. Xmgrace (v 5.1.4) was used to generate data plots. PyMOL (v 2.5.0) open-source version was used to generate molecular graphics and is available at https://github.com/schrodinger/pymol-open-source.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00783.

Mechanistic illustration of sACE2 binding with SARS CoV-2 and preventing viral infection; scatter plots with a regression line for mutagenesis results from FoldX and SSIPe compared with the experimental DMS experiment; Flex ddG REU values for 10 sACE2 mutant systems; binding free energies for wild-type and **FFWF** mutant systems and full pairwise decomposition energies from MM/GBSA calculations; and FEL and distance distributions for ACE2−RBD interface residues (PDF)

AUTHOR INFORMATION

Corresponding Author

Shahidul M. Islam − *Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois 60607, United States;* orcid.org/0000-0001-5769-6844; Phone: 312-355-3767; Email: mshahidi@uic.edu

Authors

Brandon Havranek − *Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois 60607, United States*

Kui K. Chan − *Orthogonal Biologics Inc., Urbana, Illinois 61801, United States*

Austin Wu − *Department of Computer Science, Northwestern University, Evanston, Illinois 60208, United States*

Erik Procko − *Department of Biochemistry and Cancer Center at Illinois, University of Illinois, Urbana, Illinois 61801, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c00783

Notes

The authors declare the following competing financial interest(s): KKC and EP are cofounders of Orthogonal Biologics, Inc., which licenses IP regarding engineered ACE2 derivatives from the University of Illinois. EP is an inventor on patent filings by the University of Illinois.

ACKNOWLEDGMENTS

This work is supported by the COVID-19 High Performance Computing Consortium project and used the Extreme Science and Engineering Discovery Environment (XSEDE) COMET GPU at SDSC through allocation TG-CHE200091, which is supported by the National Science Foundation grant number ACI-1548562.

REFERENCES

(1) Dong, E.; Du, H.; Gardner, L. An Interactive Web-Based Dashboard to Track COVID-19 in Real Time. *Lancet Infect. Dis.* **2020**, *20*, 533−534.

(2) Fourth large-scale COVID-19 vaccine trial begins in the United States. https://www.nih.gov/news-events/news-releases/fourth-large-scale-covid-19-vaccine-trial-begins-united-states Accessed May 1, 2021.

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(3) Phase 3 Clinical Testing in the US of AstraZeneca COVID-19 Vaccine Candidate Begins. https://www.nih.gov/news-events/news-releases/phase-3-clinical-testing-us-astrazeneca-covid-19-vaccine-candidate-begins Accessed May 1, 2021.
(4) COVID-19 Vaccines. https://www.fda.gov/emergency-preparedness-and-response/coronavirus-disease-2019-covid-19/covid-19-vaccines Accessed May 1, 2021.
(5) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281−292.
(6) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271−280.
(7) Huang, Y.; Yang, C.; Xu, X.-f.; Xu, W.; Liu, S.-w. Structural and Functional Properties of SARS-CoV-2 Spike Protein: Potential Antivirus Drug Development for COVID-19. *Acta Pharmacol. Sin.* **2020**, *41*, 1141−1149.
(8) Walls, A. C.; Tortorici, M. A.; Snijder, J.; Xiong, X.; Bosch, B.-J.; Rey, F. A.; Veesler, D. Tectonic Conformational Changes of a Coronavirus Spike Glycoprotein Promote Membrane Fusion. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 11157−11162.
(9) Tortorici, M. A.; Veesler, D. Structural Insights into Coronavirus Entry. *Adv. Virus Res.* **2019**, *105*, 93−116.
(10) Brouwer, P. J. M.; Caniels, T. G.; van der Straten, K.; Snitselaar, J. L.; Aldon, Y.; Bangaru, S.; Torres, J. L.; Okba, N. M. A.; Claireaux, M.; Kerster, G.; Bentlage, A. E. H.; van Haaren, M. M.; Guerra, D.; Burger, J. A.; Schermer, E. E.; Verheul, K. D.; van der Velde, N.; van der Kooi, A.; van Schooten, J.; van Breemen, M. J.; Bijl, T. P. L.; Sliepen, K.; Aartse, A.; Derking, R.; Bontjer, I.; Kootstra, N. A.; Wiersinga, W. J.; Vidarsson, G.; Haagmans, B. L.; Ward, A. B.; de Bree, G. J.; Sanders, R. W.; van Gils, M. J. Potent Neutralizing Antibodies from COVID-19 Patients Define Multiple Targets of Vulnerability. *Science* **2020**, *369*, 643−650.
(11) Ju, B.; Zhang, Q.; Ge, J.; Wang, R.; Sun, J.; Ge, X.; Yu, J.; Shan, S.; Zhou, B.; Song, S.; Tang, X.; Yu, J.; Lan, J.; Yuan, J.; Wang, H.; Zhao, J.; Zhang, S.; Wang, Y.; Shi, X.; Liu, L.; Zhao, J.; Wang, X.; Zhang, Z.; Zhang, L. Human Neutralizing Antibodies Elicited by SARS-CoV-2 Infection. *Nature* **2020**, *584*, 115−119.
(12) Baum, A.; Fulton, B. O.; Wloga, E.; Copin, R.; Pascal, K. E.; Russo, V.; Giordano, S.; Lanza, K.; Negron, N.; Ni, M.; Wei, Y.; Atwal, G. S.; Murphy, A. J.; Stahl, N.; Yancopoulos, G. D.; Kyratsous, C. A. Antibody Cocktail to SARS-CoV-2 Spike Protein Prevents Rapid Mutational Escape Seen with Individual Antibodies. *Science* **2020**, *369*, 1014−1018.
(13) Kupferschmidt, K. Mutant Coronavirus in the United Kingdom Sets off Alarms, but Its Importance Remains Unclear. *Science* **2020**, DOI: 10.1126/science.abg2626.
(14) Chan, K. K.; Tan, T. J. C.; 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.
(15) Kemp, S. A.; Collier, D. A.; Datir, R.; Gayed, S.; Jahun, A.; Hosmillo, M.; Ferreira, I. A. T. M.; Rees-Spear, C.; Mlcochova, P.; Lumb, I. U.; Roberts, D.; Chandra, A.; Temperton, N.; Sharrocks, K.; Blane, E.; Briggs, J. A.; Smith, K. G.; Bradley, J. R.; Smith, C.; Goldstein, R.; Goodfellow, I. G.; Smielewska, A.; Skittrall, J. P.; Gouliouris, T.; Gkrania-Klotsas, E.; Illingworth, C. J. R.; McCoy, L. E.; Gupta, R. K. Neutralising Antibodies Drive Spike Mediated SARS-CoV-2 Evasion. **2020**, 2020.12.05.20241927. medRxiv.
(16) Ku, Z.; Xie, X.; Davidson, E.; Ye, X.; Su, H.; Menachery, V. D.; Li, Y.; Yuan, Z.; Zhang, X.; Muruato, A. E.; i Escuer, A. G.; Tyrell, B.; Doolan, K.; Doranz, B. J.; Wrapp, D.; Bates, P. F.; McLellan, J. S.; Weiss, S. R.; Zhang, N.; Shi, P.-Y.; An, Z. Molecular Determinants and Mechanism for Antibody Cocktail Preventing SARS-CoV-2 Escape. *Nat. Commun.* **2021**, *12*, 469.
(17) Coronavirus (COVID-19) Update: FDA Revokes Emergency Use Authorization for Monoclonal Antibody Bamlanivimab. https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-revokes-emergency-use-authorization-monoclonal-antibody-bamlanivimab Accessed May 1, 2021.
(18) Xie, X.; Liu, Y.; Liu, J.; Zhang, X.; Zou, J.; Fontes-Garfias, C. R.; Xia, H.; Swanson, K. A.; Cutler, M.; Cooper, D.; Menachery, V. D.; Weaver, S. C.; Dormitzer, P. R.; Shi, P.-Y. Neutralization of SARS-CoV-2 Spike 69/70 Deletion, E484K and N501Y Variants by BNT162b2 Vaccine-Elicited Sera. *Nat. Med.* **2021**, *27*, 620.
(19) Haschke, M.; Schuster, M.; Poglitsch, M.; Loibner, H.; Salzberg, M.; Bruggisser, M.; Penninger, J.; Krähenbühl, S. Pharmacokinetics and Pharmacodynamics of Recombinant Human Angiotensin-Converting Enzyme 2 in Healthy Human Subjects. *Clin. Pharmacokinet.* **2013**, *52*, 783−792.
(20) Khan, A.; Benthin, C.; Zeno, B.; Albertson, T. E.; Boyd, J.; Christie, J. D.; Hall, R.; Poirier, G.; Ronco, J. J.; Tidswell, M.; Hardes, K.; Powley, W. M.; Wright, T. J.; Siederer, S. K.; Fairman, D. A.; Lipson, D. A.; Bayliffe, A. I.; Lazaar, A. L. A Pilot Clinical Trial of Recombinant Human Angiotensin-Converting Enzyme 2 in Acute Respiratory Distress Syndrome. *Crit. Care* **2017**, *21*, 234.
(21) Zoufaly, A.; Poglitsch, M.; Aberle, J. H.; Hoepler, W.; Seitz, T.; Traugott, M.; Grieb, A.; Pawelka, E.; Laferl, H.; Wenisch, C.; Neuhold, S.; Haider, D.; Stiasny, K.; Bergthaler, A.; Puchhammer-Stoeckl, E.; Mirazimi, A.; Montserrat, N.; Zhang, H.; Slutsky, A. S.; Penninger, J. M. Human Recombinant Soluble ACE2 in Severe COVID-19. *Lancet Respir. Med.* **2020**, *8*, 1154−1158.
(22) Chan, K. K.; Dorosky, D.; Sharma, P.; Abbasi, S. A.; Dye, J. M.; Kranz, D. M.; Herbert, A. S.; Procko, E.; Herbert, E. P. Engineering Human ACE2 to Optimize Binding to the Spike Protein of SARS Coronavirus 2. *Science* **2020**, *369*, 1261.
(23) Glasgow, A.; Glasgow, J.; Limonta, D.; Solomon, P.; Lui, I.; Zhang, Y.; Nix, M. A.; Rettko, N. J.; Zha, S.; Yamin, R.; Kao, K.; Rosenberg, O. S.; Ravetch, J. V.; Wiita, A. P.; Leung, K. K.; Lim, S. A.; Zhou, X. X.; Hobman, T. C.; Kortemme, T.; Wells, J. A. Engineered ACE2 Receptor Traps Potently Neutralize SARS-CoV-2. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 28046−28055.
(24) Cohen-Dvashi, H.; Weinstein, J.; Katz, M.; Eilon, M.; Mor, Y.; Shimon, A.; Strobelt, R.; Shemesh, M.; Fleishman, S. J.; Diskin, R. Coronacept—a Potent Immunoadhesin against SARS-CoV-2. **2020**, 2020.08.12.247940. bioRxiv.
(25) Higuchi, Y.; Suzuki, T.; Arimori, T., Ikemura, N.; Mihara, E.; Kirita, Y.; Ohgitani, E.; Mazda, O.; Motooka, D.; Nakamura, S.; Sakai, Y.; Itoh, Y.; Sugihara, F.; Matsuura, Y.; Matoba, S.; Okamoto, T.; Takagi, J.; Hoshino, A. High Affinity Modified ACE2 Receptors Protect from SARS-CoV-2 Infection in Hamsters. **2020**, 2020.09.16.299891. bioRxiv.
(26) Sims, J. J.; Greig, J. A.; Michalson, K. T.; Lian, S.; Martino, R. A.; Meggersee, R.; Turner, K. B.; Nambiar, K.; Dyer, C.; Hinderer, C.; Horiuchi, M.; Yan, H.; Huang, X.; Chen, S.-J.; Wilson, J. M. Intranasal Gene Therapy to Prevent Infection by SARS-CoV-2 Variants. **2021**, 2021.04.09.439149. bioRxiv.
(27) Leman, J. K.; Weitzner, B. D.; Lewis, S. M.; Adolf-Bryfogle, J.; Alam, N.; Alford, R. F.; Aprahamian, M.; Baker, D.; Barlow, K. A.; Barth, P.; Basanta, B.; Bender, B. J.; Blacklock, K.; Bonet, J.; Boyken, S. E.; Bradley, P.; Bystroff, C.; Conway, P.; Cooper, S.; Correia, B. E.; Coventry, B.; Das, R.; De Jong, R. M.; DiMaio, F.; Dsilva, L.; Dunbrack, R.; Ford, A. S.; Frenz, B.; Fu, D. Y.; Geniesse, C.; Goldschmidt, L.; Gowthaman, R.; Gray, J. J.; Gront, D.; Guffy, S.; Horowitz, S.; Huang, P.-S.; Huber, T.; Jacobs, T. M.; Jeliazkov, J. R.; Johnson, D. K.; Kappel, K.; Karanicolas, J.; Khakzad, H.; Khar, K. R.; Khare, S. D.; Khatib, F.; Khramushin, A.; King, I. C.; Kleffner, R.; Koepnick, B.; Kortemme, T.; Kuenze, G.; Kuhlman, B.; Kuroda, D.; Labonte, J. W.; Lai, J. K.; Lapidoth, G.; Leaver-Fay, A.; Lindert, S.; Linsky, T.; London, N.; Lubin, J. H.; Lyskov, S.; Maguire, J.; Malmström, L.; Marcos, E.; Marcu, O.; Marze, N. A.; Meiler, J.; Moretti, R.; Mulligan, V. K.; Nerli, S.; Norn, C.; Ó'Conchúir, S.; Ollikainen, N.; Ovchinnikov, S.; Pacella, M. S.; Pan, X.; Park, H.; Pavlovicz, R. E.; Pethe, M.; Pierce, B. G.; Pilla, K. B.; Raveh, B.; Renfrew, P. D.; Burman, S. S. R.; Rubenstein, A.; Sauer, M. F.; Scheck, A.; Schief, W.; Schueler-Furman, O.; Sedan, Y.; Sevy, A. M.; Sgourakis, N. G.; Shi, L.; Siegel, J. B.; Silva, D.-A.; Smith, S.; Song,

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Y.; Stein, A.; Szegedy, M.; Teets, F. D.; Thyme, S. B.; Wang, R. Y.-R.; Watkins, A.; Zimmerman, L.; Bonneau, R. Macromolecular Modeling and Design in Rosetta: Recent Methods and Frameworks. *Nat. Methods* **2020**, *17*, 665−680.
(28) Chatterjee, P.; Ponnapati, M.; Kramme, C.; Plesa, A. M.; Church, G. M.; Jacobson, J. M. Targeted Intracellular Degradation of SARS-CoV-2 via Computationally Optimized Peptide Fusions. *Commun. Biol.* **2020**, *3*, 715.
(29) Xue, T.; Wu, W.; Guo, N.; Wu, C.; Huang, J.; Lai, L.; Liu, H.; Li, Y.; Wang, T.; Wang, Y. Single Point Mutations Can Potentially Enhance Infectivity of SARS-CoV-2 Revealed by in Silico Affinity Maturation and SPR Assay. **2020**, 2020.12.24.424245. bioRxiv.
(30) Cao, L.; Goreshnik, I.; Coventry, B.; Case, J. B.; Miller, L.; Kozodoy, L.; Chen, R. E.; Carter, L.; Walls, A. C.; Park, Y.-J.; Strauch, E.-M.; Stewart, L.; Diamond, M. S.; Veesler, D.; Baker, D. De Novo Design of Picomolar SARS-CoV-2 Miniprotein Inhibitors. *Science* **2020**, *370*, 426.
(31) Chevalier, A.; Silva, D.-A.; Rocklin, G. J.; Hicks, D. R.; Vergara, R.; Murapa, P.; Bernard, S. M.; Zhang, L.; Lam, K.-H.; Yao, G.; Bahl, C. D.; Miyashita, S.-I.; Goreshnik, I.; Fuller, J. T.; Koday, M. T.; Jenkins, C. M.; Colvin, T.; Carter, L.; Bohn, A.; Bryan, C. M.; Fernández-Velasco, D. A.; Stewart, L.; Dong, M.; Huang, X.; Jin, R.; Wilson, I. A.; Fuller, D. H.; Baker, D. Massively Parallel de Novo Protein Design for Targeted Therapeutics. *Nature* **2017**, *550*, 74−79.
(32) Procko, E.; Berguig, G. Y.; Shen, B. W.; Song, Y.; Frayo, S.; Convertine, A. J.; Margineantu, D.; Booth, G.; Correia, B. E.; Cheng, Y.; Schief, W. R.; Hockenbery, D. M.; Press, O. W.; Stoddard, B. L.; Stayton, P. S.; Baker, D. A Computationally Designed Inhibitor of an Epstein-Barr Viral Bcl-2 Protein Induces Apoptosis in Infected Cells. *Cell* **2014**, *157*, 1644−1656.
(33) Berger, S.; Procko, E.; Margineantu, D.; Lee, E. F.; Shen, B. W.; Zelter, A.; Silva, D.-A.; Chawla, K.; Herold, M. J.; Garnier, J.-M.; Johnson, R.; Maccoss, M. J.; Lessene, G.; Davis, T. N.; Stayton, P. S.; Stoddard, B. L.; Fairlie, W. D.; Hockenbery, D. M.; Baker, D. Computationally Designed High Specificity Inhibitors Delineate the Roles of BCL2 Family Proteins in Cancer. *Elife* **2016**, *5*, No. e20352.
(34) Koday, M. T.; Nelson, J.; Chevalier, A.; Koday, M.; Kalinoski, H.; Stewart, L.; Carter, L.; Nieusma, T.; Lee, P. S.; Ward, A. B.; Wilson, I. A.; Dagley, A.; Smee, D. F.; Baker, D.; Fuller, D. H. A Computationally Designed Hemagglutinin Stem-Binding Protein Provides In Vivo Protection from Influenza Independent of a Host Immune Response. *PLoS Pathog.* **2016**, *12*, No. e1005409.
(35) Barlow, K. A.; Ó Conchúir, S.; Thompson, S.; Suresh, P.; Lucas, J. E.; Heinonen, M.; Kortemme, T. Flex DdG: Rosetta Ensemble-Based Estimation of Changes in Protein-Protein Binding Affinity upon Mutation. *J. Phys. Chem. B* **2018**, *122*, 5389−5399.
(36) Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX Web Server: An Online Force Field. *Nucleic Acids Res.* **2005**, *33*, W382−W388.
(37) Huang, X.; Zheng, W.; Pearce, R.; Zhang, Y. SSIPe: Accurately Estimating Protein−Protein Binding Affinity Change upon Mutations Using Evolutionary Profiles in Combination with an Optimized Physical Energy Function. *Bioinformatics* **2020**, *36*, 2429−2437.
(38) 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.
(39) Smith, C. A.; Kortemme, T. Backrub-Like Backbone Simulation Recapitulates Natural Protein Conformational Variability and Improves Mutant Side-Chain Prediction. *J. Mol. Biol.* **2008**, *380*, 742−756.
(40) Davis, I. W.; Arendall, W. B.; Richardson, D. C.; Richardson, J. S. The Backrub Motion: How Protein Backbone Shrugs When a Sidechain Dances. *Structure* **2006**, *14*, 265−274.
(41) Keedy, D. A.; Georgiev, I.; Triplett, E. B.; Donald, B. R.; Richardson, D. C.; Richardson, J. S. The Role of Local Backrub Motions in Evolved and Designed Mutations. *PLoS Comput. Biol.* **2012**, *8*, No. e1002629.
(42) Shapovalov, M. V.; Dunbrack, R. L. A Smoothed Backbone-Dependent Rotamer Library for Proteins Derived from Adaptive Kernel Density Estimates and Regressions. *Structure* **2011**, *19*, 844−858.
(43) Song, Y.; Tyka, M.; Leaver-Fay, A.; Thompson, J.; Baker, D. Structure-guided Forcefield Optimization. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 1898−1909.
(44) O'Meara, M. J.; Leaver-Fay, A.; Tyka, M. D.; Stein, A.; Houlihan, K.; DiMaio, F.; Bradley, P.; Kortemme, T.; Baker, D.; Snoeyink, J.; Kuhlman, B. Combined Covalent-Electrostatic Model of Hydrogen Bonding Improves Structure Prediction with Rosetta. *J. Chem. Theory Comput.* **2015**, *11*, 609−622.
(45) Hastie, T. J.; Tibshirani, R. J. *Generalized Additive Models*; CRC press, 1990; Vol. *43*.
(46) Delgado, J.; Radusky, L. G.; Cianferoni, D.; Serrano, L. FoldX 5.0: Working with RNA, Small Molecules and a New Graphical Interface. *Bioinformatics* **2019**, *35*, 4168−4169.
(47) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural Basis of Receptor Recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221−224.
(48) Wang, Y.; Liu, M.; Gao, J. Enhanced Receptor Binding of SARS-CoV-2 through Networks of Hydrogen-Bonding and Hydrophobic Interactions. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 13967−13974.
(49) Zondlo, N. J. Aromatic−Proline Interactions: Electronically Tunable CH/π Interactions. *Acc. Chem. Res.* **2013**, *46*, 1039−1049.
(50) Pal, D.; Chakrabarti, P. Cis Peptide Bonds in Proteins: Residues Involved, Their Conformations, Interactions and Locations 1 1Edited by J. M. Thornton. *J. Mol. Biol.* **1999**, *294*, 271−288.
(51) Bhattacharyya, R.; Chakrabarti, P. Stereospecific Interactions of Proline Residues in Protein Structures and Complexes. *J. Mol. Biol.* **2003**, *331*, 925−940.
(52) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Designing a 20-Residue Protein. *Nat. Struct. Biol.* **2002**, *9*, 425−430.
(53) Zarrinpar, A.; Bhattacharyya, R. P.; Lim, W. A. The Structure and Function of Proline Recognition Domains. *Sci. Signaling* **2003**, *2003*, re8.
(54) Kay, B. K.; Williamson, M. P.; Sudol, M. The Importance of Being Proline: The Interaction of Proline-rich Motifs in Signaling Proteins with Their Cognate Domains. *FASEB J.* **2000**, *14*, 231−241.
(55) Wang, C.; Li, W.; Drabek, D.; Okba, N. M. A.; van Haperen, R.; Osterhaus, A. D. M. E.; van Kuppeveld, F. J. M.; Haagmans, B. L.; Grosveld, F.; Bosch, B.-J. A Human Monoclonal Antibody Blocking SARS-CoV-2 Infection. *Nat. Commun.* **2020**, *11*, 2251.
(56) Westendorf, K.; Žentelis, S.; Foster, D.; Vaillancourt, P.; Wiggin, M.; Lovett, E.; Hendle, J.; Pustilnik, A.; Sauder, J. M.; Kraft, L.; Hwang, Y.; Siegel, R. W.; Chen, J.; Heinz, B. A.; Higgs, R. E.; Kalleward, N.; Jepson, K.; Goya, R.; Smith, M. A.; Collins, D. W.; Pellacani, D.; Xiang, P.; de Puyraimond, V.; Ricicova, M.; Devorkin, L.; Pritchard, C.; O'Neill, A.; Cohen, C.; Dye, J.; Huie, K. I.; Badger, C. V.; Kobasa, D.; Audet, J.; Freitas, J. J.; Hassanali, S.; Hughes, I.; Munoz, L.; Palma, H. C.; Ramamurthy, B.; Cross, R. W.; Geisbert, T. W.; Borisevich, V.; Lanz, I.; Anderson, L.; Sipahimalani, P.; Corbett, K. S.; Wang, L.; Yang, E. S.; Zhang, Y.; Shi, W.; Graham, B. S.; Mascola, J. R.; Fernandez, T. L.; Hansen, C. L.; Falconer, E.; Jones, B. E.; Barnhart, B. C. LY-CoVI404 Potentyl Neutralized SARS CoV-2 Variants. **2021**. bioRxiv.
(57) Hurlburt, N. K.; Seydoux, E.; Wan, Y.-H.; Edara, V. V.; Stuart, A. B.; Feng, J.; Suthar, M. S.; McGuire, A. T.; Stamatatos, L.; Pancera, M. Structural Basis for Potent Neutralization of SARS-CoV-2 and Role of Antibody Affinity Maturation. *Nat. Commun.* **2020**, *11*, 5413.
(58) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM Structure of the 2019-NCoV Spike in the Prefusion Conformation. *Science* **2020**, *367*, 1260−1263.
(59) Yao, H.; Cai, H.; Li, T.; Zhou, B.; Qin, W.; Lavillette, D.; Li, D. A High-Affinity RBD-Targeting Nanobody Improves Fusion Partner's Potency against SARS-CoV-2. *PLoS Pathog.* **2021**, *17*, No. e1009328.

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(60) Koenig, P.-A.; Das, H.; Liu, H.; Kümmerer, B. M.; Gohr, F. N.; Jenster, L.-M.; Schiffelers, L. D. J.; Tesfamariam, Y. M.; Uchima, M.; Wuerth, J. D.; Gatterdam, K.; Ruetalo, N.; Christensen, M. H.; Fandrey, C. I.; Normann, S.; Tödtmann, J. M. P.; Pritzl, S.; Hanke, L.; Boos, J.; Yuan, M.; Zhu, X.; Schmid-Burgk, J. L.; Kato, H.; Schindler, M.; Wilson, I. A.; Geyer, M.; Ludwig, K. U.; Hällberg, B. M.; Wu, N. C.; Schmidt, F. I. Structure-Guided Multivalent Nanobodies Block SARS-CoV-2 Infection and Suppress Mutational Escape. *Science* **2021**, *371*, No. eabe6230.

(61) Pymm, P.; Adair, A.; Chan, L.-J.; Cooney, J. P.; Mordant, F. L.; Allison, C. C.; Lopez, E.; Haycroft, E. R.; O'Neill, M. T.; Tan, L. L.; Dietrich, M. H.; Drew, D.; Doerflinger, M.; Dengler, M. A.; Scott, N. E.; Wheatley, A. K.; Gherardin, N. A.; Venugopal, H.; Cromer, D.; Davenport, M. P.; Pickering, R.; Godfrey, D. I.; Purcell, D. F. J.; Kent, S. J.; Chung, A. W.; Subbarao, K.; Pellegrini, M.; Glukhova, A.; Tham, W.-H. Nanobody Cocktails Potently Neutralize SARS-CoV-2 D614G N501Y Variant and Protect Mice. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2101918118.

(62) Kondo, T.; Iwatani, Y.; Matsuoka, K.; Fujino, T.; Umemoto, S.; Yokomaku, Y.; Ishizaki, K.; Kito, S.; Sezaki, T.; Hayashi, G.; Murakami, H. Antibody-like Proteins That Capture and Neutralize SARS-CoV-2. *Sci. Adv.* **2020**, *6*, No. eabd3916.

(63) Linsky, T. W.; Vergara, R.; Codina, N.; Nelson, J. W.; Walker, M. J.; Su, W.; Barnes, C. O.; Hsiang, T.-Y.; Esser-Nobis, K.; Yu, K.; Reneer, Z. B.; Hou, Y. J.; Priya, T.; Mitsumoto, M.; Pong, A.; Lau, U. Y.; Mason, M. L.; Chen, J.; Chen, A.; Berrocal, T.; Peng, H.; Clairmont, N. S.; Castellanos, J.; Lin, Y.-R.; Josephson-Day, A.; Baric, R. S.; Fuller, D. H.; Walkey, C. D.; Ross, T. M.; Swanson, R.; Bjorkman, P. J.; Gale, M.; Blancas-Mejia, L. M.; Yen, H.-L.; Silva, D.-A. De Novo Design of Potent and Resilient HACE2 Decoys to Neutralize SARS-CoV-2. *Science* **2020**, *370*, 1208−1214.

(64) Iwanaga, N.; Cooper, L.; Rong, L.; Beddingfield, B.; Crabtree, J.; Tripp, R. A.; Qin, X.; Kolls, J. K. Novel ACE2-IgG1 Fusions with Improved in Vitro and in Vivo Activity against SARS-CoV2. **2020**, 2020.06.15.152157. bioRxiv.

(65) Monteil, V.; Kwon, H.; Prado, P.; Hagelkrüys, A.; Wimmer, R. A.; Stahl, M.; Leopoldi, A.; Garreta, E.; Hurtado del Pozo, C.; Prosper, F.; Romero, J. P.; Wirnsberger, G.; Zhang, H.; Slutsky, A. S.; Conder, R.; Montserrat, N.; Mirazimi, A.; Penninger, J. M. Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human ACE2. *Cell* **2020**, *181*, 905−913.

(66) Leaver-Fay, A.; Jacak, R.; Stranges, P. B.; Kuhlman, B. A Generic Program for Multistate Protein Design. *PLoS One* **2011**, *6*, No. e20937.

(67) Guerois, R.; Nielsen, J. E.; Serrano, L. Predicting Changes in the Stability of Proteins and Protein Complexes: A Study of More than 1000 Mutations. *J. Mol. Biol.* **2002**, *320*, 369−387.

(68) Tiberti, M.; Terkelsen, T.; Cremers, T. C.; Di Marco, M.; da Piedade, I.; Maiani, E.; Papaleo, E. MutateX: An Automated Pipeline for in-Silico Saturation Mutagenesis of Protein Structures and Structural Ensembles. **2019**, 824938. bioRxiv.

(69) Huang, X.; Pearce, R.; Zhang, Y. EvoEF2: Accurate and Fast Energy Function for Computational Protein Design. *Bioinformatics* **2020**, *36*, 1135−1142.

(70) Case, D. A.; Iii, T. E. C.; Darden, T. O. M.; Gohlke, H.; Jr, K. M. M.; Onufriev, A.; Simmerling, C.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comb. Chem.* **2005**, *26*, 1668.

(71) 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.

(72) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(73) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405−413.

(74) Adelman, S. A. Generalized Langevin Equation Approach for Atom/Solid-Surface Scattering: General Formulation for Classical Scattering off Harmonic Solids. *J. Chem. Phys.* **1976**, *64*, 2375.

(75) Axelsen, P. H.; Li, D. Improved Convergence in Dual-Topology Free Energy Calculations through Use of Harmonic Restraints. *J. Comput. Chem.* **1998**, *19*, 1278.

(76) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(77) Le Grand, S.; Götz, A. W.; Walker, R. C. SPFP: Speed without Compromise - A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374−380.

(78) Miller, B. R.; McGee, T. D.; 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.

(79) Case, D. A.; Belfon, K.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kovalenko, A.; Krasny, R.; York, D. M.; Kollman, P. A. *AMBER 2020*; University of California: San Francisco, 2020.

(80) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins: Struct., Funct., Genet.* **2004**, *55*, 383−394.

(81) Wang, E.; Weng, G.; Sun, H.; Du, H.; Zhu, F.; Chen, F.; Wang, Z.; Hou, T. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 10. Impacts of Enhanced Sampling and Variable Dielectric Model on Protein-Protein Interactions. *Phys. Chem. Chem. Phys.* **2019**, *21*, 18958−18969.

(82) Weiser, J. r.; Shenkin, P. S.; Still, W. C. Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). *J. Comput. Chem.* **1999**, *20*, 217−230.

(83) Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *J. Mol. Biol.* **2003**, *330*, 891−913.

https://doi.org/10.1021/acs.jcim.1c00783
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX