In Silico Antibody Mutagenesis for Optimizing Its Binding to Spike Protein of Severe Acute Respiratory Syndrome Coronavirus 2

Binquan Luan* and Tien Huynh

ABSTRACT: Coronavirus disease 2019 (COVID-19) is an ongoing global pandemic, and there are currently no FDA-approved medicines for treatment or prevention. Inspired by promising outcomes for convalescent plasma treatment, the development of antibody drugs (biologics) to block SARS-CoV-2 infection has been the focus of drug discovery, along with tremendous efforts in repurposing small-molecule drugs. In the past several months, experimentally, many human neutralizing monoclonal antibodies (mAbs) were successfully extracted from plasma of recovered COVID-19 patients. Currently, several mAbs targeting the SARS-CoV-2’s spike glycoprotein (S-protein) are in clinical trials. With known atomic structures of the mAb and S-protein complex, it becomes possible to investigate in silico the molecular mechanism of mAb’s binding with S-protein and to design more potent mAbs through protein mutagenesis studies, complementary to existing experimental efforts. Leveraging today’s superb computing power, we propose a fully automated in silico protocol for quickly identifying possible mutations in a mAb (e.g., CB6) to enhance its binding affinity for S-protein for the design of more efficacious therapeutic mAbs.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a new member of the broad family of RNA viruses known as coronaviruses that infect a wide range of vertebrates, including mammals and birds, and are implicated as a major cause of viral respiratory infections worldwide.1 SARS-CoV-2 is the pathogen that caused the outbreak of coronavirus disease 2019 (COVID-19) in China.2 Of the seven coronaviruses known to infect humans, HCoV-229E, HCoV-OC43, HCoV-NL63, and HCoV-HKU1 are relatively harmless common cold-causing respiratory pathogens, while the other three, MERS-CoV, SARS-CoV, and SARS-CoV-2, are highly pathogenic and could result in substantial morbidity and mortality. Although the fatality rate of COVID-19 is significantly lower than those of SARS and MERS, it is highly contagious with the underlying SARS-CoV-2 virus spreading more easily among people, resulting in the current worldwide pandemic with nearly 21 million people infected and at least 700000 deaths globally as of August 21, 2020. With limited response time for COVID-19, the only therapeutic approach is by means of repurposing existing medicines for rapid clinical trials. So far, two FDA-approved drugs (of small molecules), remdesivir and dexamethasone, have shown moderate therapeutic effects such as shortening the time to recovery3 and reducing mortality.4

As the COVID-19 pandemic continues, researchers are racing against time to search for new therapeutic treatments as well as preventive vaccines. Besides putting effort continuously into drug repurposing, much work has also focused on studying the antibodies separated from plasma of convalescent COVID-19 patients.5,6 Antibodies are Y-shaped proteins produced by the B lymphocytes (B-cells), one of the most important cells in the adaptive immune system, to fight disease-causing bacteria and viruses (antigen). Antibodies neutralize the pathogens by attaching to the surface of the invading antigen, blocking them from entering host cells and signaling them for destruction by other immune cells. So far, there are ~30 FDA-approved antibody drugs, such as ibalizumab for HIV infection.

Recently, the mechanisms of how SARS-CoV-2 infects the target cells have been reported,7,8 which help shed light on exploring neutralizing monoclonal antibodies (mAbs) to SARS-CoV-2 as a potential for both therapeutic and prophylactic applications. Indeed, protruding from the spherical surface of SARS-CoV-2 particles (Figure 1a), the spike glycoprotein (S-protein) that binds the angiotensin-converting enzyme 2 (ACE2) receptor found on numerous types of host cells as a prelude for viral entry is the main target of neutralizing mAbs. A majority of the recently isolated

Received: September 5, 2020
Accepted: October 28, 2020
Published: November 4, 2020
neutralizing high-potency mAbs have been shown to target various epitopes on the receptor-binding domain (RBD),\textsuperscript{9−11} a crucial and stable (with no mutation identified so far) region of the S-protein that facilitates the contact of SARS-CoV-2 with the ACE2 receptor.\textsuperscript{12−21} Some other high-potency neutralizing mAbs have also been found to target the N-terminal domain of S-protein.\textsuperscript{22}

Complementary to experimental efforts, \textit{in silico} approaches such as all-atom molecular dynamics (MD) simulations have been widely used to investigate the molecular mechanism of proteins and proven to produce results consistent with experimental ones.\textsuperscript{23−28} Given the urgent need for highly potent mAbs that can be used in antibody cocktails for potential treatment of COVID-19, we are motivated to develop an automated \textit{in silico} protocol for quickly identifying possible protein mutations that can enhance the binding between designed mAbs and the RBD of S-protein.\textsuperscript{22}

Figure 1. MD simulation of the Fab–RBD complex. (a) Illustration of a mAb (with two Fab regions and one Fc region) targeting the RBD of S-protein (a trimer colored yellow, green, and purple) on the surface of SARS-CoV-2 (gray). (b) MD simulation system for the antibody CB6’s one Fab in complex with the RBD of S-protein (Protein Data Bank entry 7C01). Proteins are shown in the cartoon representation; Na\textsuperscript{+} (tan) and Cl\textsuperscript{−} (cyan) are shown as van der Waals spheres, and water (transparent) is in the molecular surface representation. The Fab contains one heavy chain (fragment) and one light chain, colored blue and orange, respectively; the RBD of S-protein is colored purple. The heavy (light) chain comprises a variable region V\textsuperscript{H} (V\textsuperscript{L}) and a constant region C\textsuperscript{H} (C\textsuperscript{L}). A buried salt bridge that is composed of D104 in V\textsuperscript{H} and K417 in RBD is highlighted at the interface. (c) Time-dependent distances between the NZ atom in K417 (RBD) and the CG atom in D104 (V\textsuperscript{H}). The inset shows an enlarged view of the stable salt bridge.

Remarkably, at the interface there exists a buried salt bridge formed by K417 in the RBD and D104 in the V\textsuperscript{H} region (Figure 1b and the inset of Figure 1c). Given the relatively low dielectric constant (∼4, generally) inside the protein, the salt bridge yields a strong electrostatic interaction across the Fab–RBD interface. To quantify the stability of this salt bridge, we calculated the distance between the NZ atom in K417 and the CG atom in D104 from the 200 ns simulation trajectory. Figure 1c demonstrates that after ∼50 ns the distance between the pair of atoms saturated around 3.2 Å, confirming the stable salt bridge buried inside the protein complex. This is worth noting because salt bridges on a protein surface generally are much weaker and can break and re-form frequently due to
their exposure to water. Other key interfacial bindings will be discussed further below.

We also modeled the Fab alone in the 0.15 M NaCl electrolyte (Figure S2a), for ~200 ns of equilibration. In the absence of S-protein, the side chains of the interfacial residues of the mAb CB6 were oriented differently. For instance, being exposed to water, D104 of VH can form a hydrogen bond with Y98 of VÎ. From time to time, enhancing internal interactions between VH and VL. Here, a hydrogen bond was defined using the standard cutoff values for the distance (3 Å) and angle (20°). Similarly, the entire Fab structure alone in the simulation was stable, with RMSD values for VÎ and VÎ saturated at 0.8 and 0.9 Å (Figure S2), respectively.

With both the Fab–RBD complex (a bound state) and stand-alone Fab (a free state) structures equilibrated in solution, we applied the free energy perturbation (FEP) method (see Methods for details) to carry out the in silico alanine scan for all interfacial residues in the Fab, aiming to identify key residues for stabilizing the Fab–RBD complex. Here, we define interfacial residues as those in Fab that are within 5 Å of the RBD. These residues comprise V2, G26, F27, T28, S30, S31, N32, Y33, Y52, S53, G54, G55, S56, N76, R97, L99, M101, Y102, and D104 in the VÎ region, along with S28, S30, R31, Y92, and T94 in the VÎ region. Panels a–d of Figure 2 show a designed thermodynamical cycle that is used in the FEP method to calculate free energy difference ΔΔG for the Y33A mutation: ΔGÎ and ΔGÎ are free energy changes for RBD’s binding to the Fab and the mutant, respectively; ΔGÎ and ΔGÎ are free energy changes for annihilating Y33 and simultaneously annihilating A33 in the bound and free states, respectively. Generally, to circumvent the difficulty of directly calculating ΔGÎ and ΔGÎ, we compute ΔGÎ and ΔGÎ for the alchemy processes in the bound and free states, respectively. Details for calculating ΔΔG (=ΔGÎ −ΔGÎ) for each mutation are described in Methods.

Figure 2e summarizes all alanine scan results for interfacial residues in VÎ (blue text) and VÎ (orange text). When ΔΔG > 0 (i.e., ΔGÎ > ΔGÎ), each mutation to alanine is less favorable in the bound state than in the free state. Therefore, these residues before mutation play an important role in stabilizing the Fab–RBD complex. As discussed above, residue D104 in VÎ forms a buried salt bridge with K417 in the RBD and correspondingly D104A gives rise to a large ΔΔG (=11.47 kcal/mol). Typically, the binding free energy for a solvent-exposed salt bridge is only approximately −1 kcal/mol.13 Here, an approximately one order of magnitude larger value of ΔΔG results from the roughly one order of magnitude lower dielectric constant for electrostatic interactions inside a protein than in water. The second largest value of ΔΔG belongs to the Y33A mutation. As shown in Figure 2a, Y33 forms a stable hydrogen bond with L455 in the RBD; the mutation to alanine in the bound state (Figure 2b) reduced the level of interfacial binding and thus is unfavorable. On the contrary, when surrounded by water molecules the hydrophobic Y33 in the free state (Figure 2c) is disadvantageous, and thus, its mutation to alanine becomes favorable in the free state. Altogether, ΔΔG = 8.27 kcal/mol, suggesting that Y33 is also essential in stabilizing the Fab–RBD complex.

Additionally, an alanine scan for N32, Y52, M101, L99, N76, and G54 in VÎ together with Y92, S30, S28, and T94 in VÎ yielded positive values of ΔΔG (Figure 2e). Therefore, all of these residues contribute substantially to the stable interfacial binding observed in the MD simulation. For example, in Figure S3, we described the molecular mechanism of the N32A mutation in how the interfacial interaction was reduced locally after the mutation. When searching for mutations that can enhance Fab–RBD binding, we intentionally keep these key residues intact.

When ΔΔG < 0 (i.e., ΔGÎ < ΔGÎ), each mutation to alanine becomes more favorable in the bound state than in the free state. Therefore, alanine mutations (for V2, G55, R97, S56, T28, and S31 in VÎ, as shown in Figure 2e) with negative values of ΔΔG are possible candidates for increasing the level of Fab–RBD binding. To account for negative ΔΔG values obtained from FEP runs, we unveiled the molecular mechanism from simulation trajectories that several mutations to alanine can eliminate the stable local structure in the free state, which allows nearby residues in the Fab to form stronger interfacial bindings with the RBD in the bound state with weaker internal constraints inside the Fab. For instance, the ΔΔG for V2A is −0.96 kcal/mol (Figure 2e), and from the snapshots taken before and after the mutation (panels a and b of Figure S4, respectively), we observe that originally (before the V2A mutation) Y108 inside VÎ was blocked by V2 (due to the strong hydrophobic interaction) and not close to the interface. However, after the V2A mutation, Y108 made its way to the interface and interacted with N487 in the RBD. The pairwise interaction potential energy for Y108 in the Fab and N487 in the RBD decreased by ~0.5 kcal/mol (Figure S4c).

Figure 2. In silico alanine scan. (a–d) Illustration of the FEP calculations with the Y33A mutation. Protein segments (in cartoon representation) are colored the same as those in Figure 1b. (a) Bound state between the original Fab and RBD. (b) Bound state between the mutated Fab (Y33A) and RBD. (c) Free state of the original Fab in water. (d) Free state of the mutated Fab (Y33A) in water. (e) Alanine scan results. Mutations in VÎ and VÎ are colored blue and orange, respectively.
which suggests an improved interfacial binding (echoing the negative $\Delta\Delta G$ for V2A). Similarly, R97A yielded a $\Delta G_0$ that was 0.83 kcal/mol larger than $\Delta G_0$ in the bound state [i.e., $\Delta\Delta G = -0.83 \text{ kcal/mol}$ (Figure 2e)] because it destabilized the local structure in the free state where R97 formed a salt bridge with D107 and a hydrogen bond with nearby N32 on the surface of V$^H$.

For the third group containing G26, F27, S30, S53, and Y102 in V$^H$ as well as R31 in V$^L$ (Figure 2e), their mutations to alanine were fruitless with negligible values ($\sim 0$) of $\Delta\Delta G$, which indicates that these residues, despite being close to the interface, are dispensable in Fab–RBD binding. Indeed, these residues located at peripheral areas of the interfacial contact were more exposed to water than contacted by residues in RBD. Nevertheless, mutations of these residues to others might offer unforeseen opportunities for enhancing Fab–RBD binding.

Among these residues, we chose S30 and G26 in V$^H$ to perform enumerated mutations to other residues (Figure 3a,b).

Figure 3. FEP calculations for non-essential residues. (a) Enumerated mutations for S30 in the V$^H$ domain. (b) Enumerated mutations for G26 in the V$^L$ domain. (c) In silico workflow for identifying possible mutations that enhance antibody–S-protein binding.

A majority of mutations for S30 (such as S30L and S30T) produced positive values of $\Delta\Delta G$ (Figure 3a), i.e., weakening Fab–RBD binding. For the S30G mutation, $\Delta\Delta G \sim 0$ because G30 is even smaller than alanine and thus became more trivial in Fab–RBD binding. Fortuitously, two favorable mutations, S30M and S30D, yielded negative values of $\Delta\Delta G$, $-1.56$ and $-0.97$ kcal/mol, respectively. From the trajectory analysis, we found that remarkably the exnihilated side chain of D30 can form a salt bridge with K458 in the RBD, improving the stability of the complex structure (Figure 4a). Due to the geometric constraints, in the S30E mutation, we did not observe the formation of its salt bridge between E30 and K458, and correspondingly, the $\Delta\Delta G$ for S30E is 0.38 kcal/mol (i.e., unfavorable). While the advantage of the S30D mutation can be easily recognized, the molecular mechanism for the S30M mutation is not intuitive. As shown in Figure 4b, the direct interaction between the charged K458 residue and the hydrophobic M30 is energetically forbidden (or effectively repulsive), and consequently, M30 folded itself into a pocket formed by R71, V29, and N73 in V$^H$. During the free state alchemy process, the exnihilated M30 was not in that pocket and was exposed to water instead, indicating that without the effective repulsion from K458 the entropy contribution by M30 to the binding free energy change outweighed the enthalpy contribution. Overall, the S30M mutation stabilized the local structure in the bound state and resulted in a negative $\Delta\Delta G$ value.

It is worth mentioning that besides S30 being mutated to other amino acids S30 can be subject to post-translational modifications, namely phosphorylation. Similar to D30, phosphorylated S30 (S30p) with a net charge of $-2e$ (where $e$ is the elementary charge) can form a salt bridge with K458 in the RBD (Figure S5). From FEP calculations, $\Delta\Delta G = -2.83 \pm 0.73$ kcal/mol, confirming that the stronger electrostatic interaction for S30p than for D30 in the salt bridge with K458 can substantially stabilize interfacial binding. Therefore, as a biologic drug (biologics), the designed mAb can have extra flexibility when being synthesized outside the human body.

For the enumerated mutagenesis of G26, we found that many mutations (into E, W, C, M, F, and L) yielded negative $\Delta\Delta G$ values (Figure 3b). Because G26 (comprising only one hydrogen atom in its side chain) is the smallest among all amino acids, it interacted only weakly with surface residues in the RBD as indicated in the alanine scan. When G26 is mutated into other bulkier residues, one can foresee the strengthening of the interfacial Fab–RBD interaction. In particular, the $\Delta\Delta G$ values for G26E and G26W are $-1.93$ and $-1.75$ kcal/mol, respectively, significantly improving the local interfacial binding. The molecular mechanism for G26E is illustrated in Figure 4c. After the mutation, the exnihilated E26 formed one hydrogen bond with N487 in the RBD via their side chains (namely, the amide group in N487 and the carboxyl group in E26), and the other one with S477 in the RBD via their backbones. For G26W, the exnihilated W26 (a

![Figure 4. Illustrations of possible molecular mechanisms of mutations that enhance Fab–RBD binding: (a) S30D, (b) S30M, (c) G26E, and (d) G26W.](https://dx.doi.org/10.1021/acs.jpclett.0c02706)
bulky one) can form a hydrogen bond between the carboxamide group in the side chain of N487 (in the RBD) and its indole nitrogen (-NH-) group. Additionally, W26 was in contact with G476 and T478 (in the RBD) via hydrophobic interaction (Figure 4d). On the contrary, it is expected that the exnihilation of hydrophobic W26 in the free state is energetically unfavorable.

If we had not obtained any new mutations for G26 and S30 to enhance Fab–RBD binding, other residues with ΔΔG values close to zero or even negative from the alanine scan (Figure 2e) should be further explored with enumerated mutations. In a summary of the strategy discussed above for binding, we illustrate the entire in silico workflow in Figure 3c. Briefly, we first carried out MD simulations for both free (Fab only) and bound (Fab+RBD) states, followed by an alanine scan of all interfacial residues in Fab (FEPS runs). From alanine scan data, we obtained key residues for binding with the RBD (ΔΔG > 0) and the first group of suggested mutations for enhancing the interaction with the RBD (ΔΔG < 0). For mutations for which ΔΔG ∼ 0, we further performed the enumerated mutagenesis (FEPS runs) to search for favorable mutations (ΔΔG < 0), i.e., the second group of suggested mutations for enhancing the interaction with the RBD.

Overall, using this workflow, we identified several encouraging mutations such as V2A, G55A, R97A, S30M, S30D, G26E, G26W, G26C, G26M, etc. Furthermore, it is possible to combine two or three of these favorable mutations to further enhance the binding affinity. For example, we suggest two mutants with triple mutations, CB6-DAW (including S30D, R97A, and G26W) and CB6-AME (including V2A, S30M, and G26E), taking into account the potential interference of encouraging mutations listed above. These designed mAbs with improved binding free energies can help reduce the dosage of mAb, which makes the mAb therapy more affordable.

In conclusion, we proposed an in silico approach for optimizing the binding between a designed antibody and S-protein (particularly the RBD). Taking advantage of ever-increasing computing power, we performed all-atom MD simulations as well as FEPS calculations for alanine scan and enumerated mutations, which yielded several promising candidates for optimizing the mAb CB6, such as CB6-DAW and CB6-AME. On the contrary, we identified several key RBD residues, namely, K417, L455, N487, K458, S477, and G476, that were important for RBD–Fab interactions. It is worth mentioning that among these residues K417, L455, and N487 were also highlighted as being significant for RBD–ACE2 interactions in a previous study. As opposed to previous studies (e.g., ref 30 that relies on human expertise to select possible mutations), here we demonstrated that the entire workflow (as shown in Figure 3c) can be easily automated on high-performance clusters (HPC) or supercomputers (such as IBM’s Summit) without human intervention. Within the accuracy of state-of-the-art force fields used in MD and FEP calculations, we expect that the identified favorable mutations are highly promising for designing more efficacious antibodies and deserve further in vitro or in vivo verification. The feedback from experiments can be further employed to calibrate the simulation protocol, which promotes the synergistic development of mAbs combining in silico and in vitro/in vivo efforts to meet potential challenges of virus mutation in the future. More generally, our designed workflow can be readily applied to optimize any biologic drugs.

One potential risk of applying mAb drugs for the therapeutic treatment of COVID-19 is the so-called antibody-dependent enhancement (ADE), which can also be triggered by vaccines. One possible mechanism of ADE of disease is the potential binding between the mAb’s Fc domain and the FcγRs on myeloid cells that causes the internalization of a mAb-bound virus.31 While this could be detrimental for vaccine-induced mAbs or mAbs from convalescent plasma, for designed mAb drugs it is possible to engineer the Fc domain to avoid its binding with FcγRs. For example, following the workflow shown in Figure 3c to find residues in the Fc domain with a positive ΔΔG in the alanine scan, one can mutate some of those residues to alanine to reduce the binding affinity between the Fc domain and FcγRs, ensuring the safety of mAb drugs.

At present, several mAb drugs (such as LY-CoV555 and REGN-CoV-2) are currently in clinical trials that have already shown propitious outcomes. With collaborative experimental and theoretical efforts, we hope to accelerate the discovery of safe and efficacious mAb drugs for both therapeutic and prophylactic applications. Available in large-scale production, designed mAbs (biologics) are expected to replace natural ones from convalescent plasma for combating COVID-19.

**METHODS**

**MD Simulations.** All-atom MD simulations were carried out for both the bound (Fab of mAb CB6 bound to the RBD of S-protein) and the free (stand-alone Fab of mAb CB6) states using the NAMD2.13 package22 running on the IBM Power Cluster. To model the Fab–RBD complex (a bound state), we obtained the previously resolved crystal structure (PDB entry 7C01) from the PDB. In the crystal structure, a glycan (N-acetyl-d-glucosamine) is present on N343 of the RBD. For a successful binding between the RBD and a mAb, two factors are required. One is the accessibility of the RBD to mAb, which is modulated by glycans on S-protein;27,28,33 and the other is the binding affinity between the RBD (in the “up” state) and a mAb. Therefore, with the focus on the second factor in this work, we did not include the glycan (on the RBD) that is far from the binding interface (Figure S7). After the complex had been solvated in a rectangular water box that measures ~76 Å × 76 Å × 134.76 Å, 66 Na+ and 71 Cl− ions were added to the system, neutralizing the entire simulation system and setting the ion concentration to 0.15 M (Figure 1b). The final system containing 79,466 atoms was first minimized for 10 ps and further equilibrated for 200 ps in the NPT ensemble (P ~ 1 bar, and T ~ 300 K), with atoms in the backbone harmonically constrained (spring constant k = 1 kcal mol−1 Å−2). After constraints on the atoms in VH, VL, and the RBD had been removed, the entire system was equilibrated for additional 1 ns in the NPT ensemble. During the production run in the NVT ensemble, all atoms in the backbones of the C2 and C4 domains (not close to the Fab–RBD interface) remained harmonically restrained (spring constant k = 1 kcal mol−1 Å−2), preventing the whole complex from rotating out of the water box. We followed the same protocol to prepare the free state simulation.

The CHARMM36 force field34 was applied for proteins. The TIP3P model35,36 was chosen for water. The standard force field1 was used for ions. The periodic boundary conditions (PBC) were applied in all three dimensions. Long-range Coulomb interactions were calculated using particle mesh
Ewald (PME) full electrostatics with a grid size of ~1 Å in each dimension. The van der Waals (vdW) energies between atoms were calculated using a smooth (10−12 Å) cutoff. Temperature $T$ was kept at 300 K by applying the Langevin thermostat, while the pressure was kept constant at 1 bar using the Nosé–Hoover method. With the SETTLE algorithm enabled to keep all bonds rigid, the simulation time step was set to 2 fs for bonded and nonbonded (e.g., vdW, angle, and dihedral) interactions, and electric interactions were calculated every 4 fs, with the multiple-time step algorithm.

Free Energy Perturbation Calculations. The free energy perturbation (FEP) method has been previously used for in silico mutagenesis studies of proteins. After equilibrating binding free energy for various mutations on the Fab of the employed the FEP method to calculate the change in the bound and free states in respective MD simulations, we

Generally, direct calculations of the free energy for various mutations on the Fab of the antibody CB6. As shown in the thermodynamic cycle for the Y33A mutation (Figure 2a−d), the difference between RBD’s binding free energies can be calculated by the following equation:

$$\Delta \Delta G = \Delta G_f - \Delta G_i = \Delta G_A - \Delta G_B$$

(1)

Generally, direct calculations of $\Delta G_i$ and $\Delta G_f$ are challenging and can be replaced by computing $\Delta G_A$ and $\Delta G_B$ instead (eq 1). From the following ensemble average, $\Delta G_i$ and $\Delta G_f$ can be calculated theoretically:

$$\Delta G_{i,2} = -k_B T \ln \left( \exp \left( \frac{H_i - H_f}{k_B T} \right) \right)$$

(2)

where $k_B$ is the Boltzmann constant, $T$ is the temperature, and $H_i$ and $H_f$ are the Hamiltonians at the initial (i) and final (f) stages, respectively. For example, for the Y33A mutation, the initial state is the wild-type CB6’s Fab and the final state is that with its Y33 replaced by A33. Using the perturbation method, many intermediate stages (denoted by $\lambda$) whose Hamiltonian

$$H(\lambda) = \lambda H_i + (1 - \lambda) H_f$$

are required between the initial and final states to improve the accuracy. In FEP calculations of $\Delta G_A$ and $\Delta G_B$ with the soft-core potential enabled, $\lambda$ varies from 0 to 1.0 in 20 perturbation windows (lasting 0.3 ns each), yielding gradual annihilation and creation processes for Y33 and A33, respectively. To avoid exhuming a residue’s side chain into an unfavorable location (a high-energy state) during the alchemical process, which is highly possible for a large and flexible side chain such as lysine and arginine, we performed ≤10 independent runs for each mutation and accepted the lowest five free energy changes for calculating the mean and the error. Thus, this approach to analysis can automatically exclude data corresponding to unphysical states of a side chain and is suitable for the fully automated workflow (Figure 3c).

An example of data analysis for G26E is provided in Figure S6.

## ASSOCIATED CONTENT

1. **Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcl.0c02706.

Figure S1 showing RMSDs for the RBD, V{sub}H, and V{sub}L in the Fab–RBD complex, Figure S2 showing the MD simulation system for the Fab only and RMSDs for V{sub}H and V{sub}L, Figure S3 showing the molecular mechanism of the N32A mutation, Figure S4 showing the molecular mechanism of the V2A mutation, Figure S5 showing the salt bridge formed by phosphorylated S30 in V{sup}H and K458 in the RBD, Figure S6 showing the data analysis for G26E, and Figure S7 showing that the glycan on the RBD is not close to the binding interface (PDF).

## AUTHOR INFORMATION

### Corresponding Author

Binquan Luan — Computational Biological Center, IBM Thomas J. Watson Research, Yorktown Heights, New York 10598, United States; orcid.org/0000-0002-9414-5379; Email: bluan@us.ibm.com

### Author

Tien Huynh — Computational Biological Center, IBM Thomas J. Watson Research, Yorktown Heights, New York 10598, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcl.0c02706

### Notes

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

T.H. and B.L. gratefully acknowledge the help from Wendy Cornell and the computing resource from the IBM Cognitive Computing Program.

### REFERENCES

(1) Yin, Y.; Wunderink, R. G. MERS, SARS and other coronaviruses as causes of pneumonia. Respirology 2018, 23, 130−137.

(2) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. 2020, 382, 727−733.

(3) Beigel, J. H.; Tomashek, K. M.; Dodd, L. E.; Mehta, A. K.; Zungman, B. S.; Kalil, A. C.; Holmman, E.; Chu, H. Y.; Luetkemeyer, A.; Kline, S. Remdesivir for the treatment of Covid–19–preliminary report. N. Engl. J. Med. 2020, 383, 992−994.

(4) Villar, J.; Ferrando, C.; Martinez, D.; Ambros, A.; Muñoz, T.; Soler, J. A.; Aguilar, G.; Alba, F.; Gonzalez-Higuera, E.; Conesa, L. A.; et al. Dexamethasone treatment for the acute respiratory distress syndrome: a multicentre, randomised controlled trial. Lancet Respir. Med. 2020, 8, 267−276.

(5) Cao, Y.; Su, B.; Guo, X.; Sun, W.; Deng, Y.; Bao, L.; Zhu, Q.; Zhang, X.; Zheng, Y.; Gen, C.; et al. Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients’ B cells. Cell 2020, 182, 73−84.e16.

(6) Duan, K.; Liu, B.; Li, C.; Zhang, H.; Yu, T.; Qu, J.; Zhou, M.; Chen, L.; Meng, S.; Hu, Y.; et al. Effectiveness of convalescent plasma therapy in severe COVID-19 patients. Proc. Natl. Acad. Sci. U. S. A. 2020, 117, 9490−9496.

(7) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 2020, 181, 271−280.e8.

(8) 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.

(9) Premkumar, L.; Segovia-Chumbez, B.; Jadi, R.; Martinez, D. R.; Raut, R.; Markmann, A.; Cornaby, C.; Bartelt, L.; Weiss, S.; Park, Y.; et al. The receptor binding domain of the viral spike protein is an
immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Science Immunology* 2020, 5, eabc8413.

(10) Wec, A. Z.; Wrapp, D.; Herbert, A. S.; Maurer, D. P.; Haslwanter, D.; Sakharova, M.; Jangra, R. K.; Dieterle, M. E.; Lilov, A.; Huang, D.; et al. Broad neutralization of SARS-related viruses by human monoclonal antibodies. *Science* 2020, 369, 731–736.

(11) Brouwer, P.; Caniels, T.; van der Straten, K.; Smitseelaar, J.; Aldon, Y.; Bangaru, S.; Torres, J.; Okha, N.; Claireaux, M.; Kerster, G.; et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* 2020, 369, 643–650.

(12) Pinto, D.; Park, Y.-J.; Beltramello, M.; Walls, A. C.; Tortorici, M. A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; De Marco, A.; et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* 2020, 583, 290–295.

(13) Robbiani, D. F.; Gaebler, C.; Muecksch, F.; Lorenzo, J. C. C.; Wang, Z.; Cho, A.; Agudelo, M.; Barnes, C. O.; Gazumyan, A.; Finkin, S.; et al. Convergent Antibody Responses to SARS-CoV-2 in Convalescent Individuals. *Nature* 2020, 584, 437–442.

(14) Yuan, M.; Wu, N. C.; Zhou, X.; Lee, C.-C. D.; So, R. T.; Lv, H.; Mok, C. K.; Wilson, I. A. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* 2020, 368, 630–633.

(15) Shi, R.; Shan, C.; Duan, X.; Chen, Z.; Liu, P.; Song, J.; Song, T.; Bi, X.; Han, C.; Wu, L.; et al. A human neutralizing antibody targets the receptor binding site of SARS-CoV-2. *Nature* 2020, 584, 120–124.

(16) Liu, L.; Wang, P.; Nair, M. S.; Yu, J.; Rapp, M.; Wang, Q.; Luo, Y.; Chan, J. F.-W.; Sahi, V.; Figueroa, A.; et al. Potent neutralizing antibodies directed to multiple epitopes on SARS-CoV-2 spike. *Nature* 2020, 584, 450–456.

(17) Wang, C.; Li, W.; Drabek, D.; Okba, N. M.; van Haperen, R.; Osterhaus, A. D.; van Kuppevelt, F. J.; Haagmans, B. L.; Grosfeld, F.; Bosch, B.-J. A human monoclonal antibody blocking SARS-CoV-2 infection. *Nat. Commun.* 2020, 11, 2251.

(18) Ju, B.; Zhang, Q.; Ge, J.; Wang, R.; Sun, J.; Ge, X.; Yu, J.; Shan, S.; Zhou, B.; Song, S.; et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* 2020, 584, 115–119.

(19) Wu, Y.; Wang, F.; Shen, C.; Peng, W.; Li, D.; Zhao, C.; Li, Z.; Li, S.; Bi, Y.; Yang, Y.; et al. A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science* 2020, 368, 1274–1278.

(20) Rogers, T. F.; Zhao, F.; Huang, D.; Beutler, N.; Burns, A.; He, W.-t.; Limbo, O.; Smith, C.; Song, G.; Woehl, J.; et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* 2020, 369, 956–963.

(21) Hansen, J.; Baum, A.; Pascal, K. E.; Russo, V.; Giordano, S.; Wloa, E.; Fulton, B. O.; Yan, Y.; Koon, K.; Patel, K.; et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* 2020, 369, 1010–1014.

(22) Chi, X.; Yan, R.; Zhang, J.; Zhang, G.; Zhang, Y.; Hao, M.; Zhang, Z.; Fan, P.; Dong, Y.; Yang, Y.; et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* 2020, 369, 650–655.

(23) Kaplun, M.; McCammon, J. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* 2002, 9, 646–652.

(24) Fredrickson, P. L.; Harrison, C. B.; Liu, Y.; Schulten, K. Challenges in protein-folding simulations. *Nat. Phys.* 2010, 6, 751–758.

(25) Huynh, T.; Wang, H.; Luan, B. *in silico* Exploration of Molecular Mechanism of Clinically Oriented Drugs for Inhibiting SARS-CoV-2’s Main Protease. *J. Phys. Chem. Lett.* 2020, 11, 4413–4420.

(26) 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.

(27) Taka, E.; Yilmaz, S. Z.; Golcuk, M.; Kilinc, C.; Aktas, U.; Yildiz, A.; Gur, M. Critical Interactions Between the SARS-CoV-2 Spike Glycoprotein and the Human ACE2 Receptor. *bioRxiv* 2020, DOI: 10.1101/2020.09.21.305490.