In silico Antibody Mutagenesis for Optimizing its Binding to the Spike Protein of SARS-CoV-2

Binquan Luan* and Tien Huynh

Computational Biological Center, IBM Thomas J. Watson Research, Yorktown Heights, New York 10598, USA

E-mail: bluan@us.ibm.com

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, developing 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 last 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 protein (Spro) are in clinical trials. With known atomic structures of mAb-Spro complex, it becomes possible to *in silico* investigate the molecular mechanism of mAb's binding with Spro and design more potent mAbs through protein mutagenesis studies, complementary to existing experimental efforts. Leveraging superb computing power nowadays, we propose a fully automated *in silico* protocol for quickly identifying possible mutations in a mAb (e.g. CB6) to enhance its binding affinity with Spro 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 as well as birds, and are implicated as a major cause of viral respiratory infections worldwide.[1] SARS-CoV-2 is the pathogen that causes the outbreak of the 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 COVID-19 has a significantly lower fatality rate than SARS and MERS, it is highly contagious with the underlying SARS-CoV-2 virus spreading more easily among people, resulting in the current pandemic worldwide with nearly 21 million people infected and at least 700 thousand deaths globally as of August 21st, 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 recovery[3] 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 on 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 about 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

2

SARS-CoV-2 as a potential for both therapeutic and prophylactic applications. Indeed, protruding from the spherical surface of SARS-CoV-2 particles (Fig. 1a), the spike glycoprotein (Spro) 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. Majority of the recently isolated neutralizing high-potency mAbs have been shown to target various epitopes on the receptor-binding domain (RBD),[9–11] a crucial and stable (with no mutation identified so far) region of the Spro that facilitates the contact of SARS-CoV-2 with the ACE2 receptor.[12–21] Some other high-potency neutralizing mAbs have also been found to target the N-terminal domain of Spro.[22]

Complementary to experimental efforts, *in silico* approaches such as the all-atom molecular dynamics (MD) simulation have been widely used to investigate the molecular mechanism of proteins and proven to produce results consistent with experimental ones.[23–25] 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 *in silico* protocol for quickly identifying possible protein mutations that can enhance the binding between designed mAbs and the RBD of Spro. More importantly, this approach can accelerate the searching of new corresponding mAbs to neutralize mutated SARS-CoV-2 when needed. Among recently discovered human neutralizing mAbs, the IC50 of CB6[15] is larger than those of many others (such as BD-368-2,[5] P2C-1F11,[18] H4+B38[19] and S309[12]), therefore the binding affinity of CB6 to Spro is relatively weaker. Due to the binding competition with ACE2, it is desirable to obtain mAbs with their potency as high as possible. Here, we demonstrate that our *in silico* protocol can be utilized to improve the potency of the mAb CB6 that recognizes an epitope site in the RBD overlapping the binding site of ACE2.

To model the interaction between the mAb CB6 and the Spro of SARS-CoV-2, we focused on the interfacial interactions between one Fab of the antibody CB6 and the RBD of the Spro as shown in Fig. 1b. Detailed simulation protocols are described in the Methods section. Briefly, atomic coordinates for the Fab-RBD complex were taken from the crystal structure
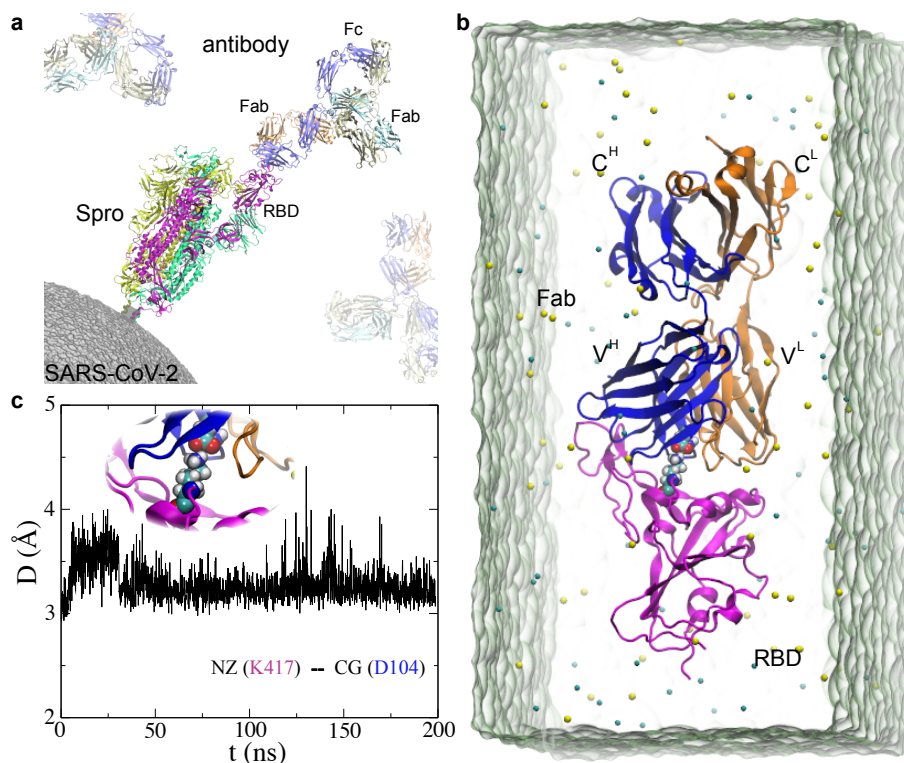
Figure 1: MD simulation of the Fab-RBD complex. a) An illustration of a mAb (with two Fab regions and one Fc region) targeting the RBD of Spro (a trimer colored in yellow, green and purple) on the surface of SARS-CoV-2 (grey). b) The MD simulation system for the antibody CB6's one Fab in complex with the RBD of Spro (PDB number: 7C01). Proteins are in the cartoon representation, K^+ (tan) and Cl^- (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 in blue and orange respectively; the RBD of Spro is in purple. The heavy (light) chain comprises a variable region V^H (V^L) and a constant region C^H (C^L). A buried salt-bridge that is composed of D104 in V^H and K417 in RBD is highlighted at the interface. c) Time-dependent distances between the atom NZ in K417 (RBD) and the atom CG in D104 (V^H). The inset shows the enlarged view of the stable salt-bridge.

(PDB ID: 7C01). The complex was further solvated in a 0.15 NaCl electrolyte. Both variable regions V^H and V^L of the heavy (blue) and the light (orange) chains respectively bind the same epitope of the RBD as the human ACE2 (Fig. 1b) does. During the about 200 ns MD simulations, the crystal structure of the Fab-RBD complex was properly equilibrated in the physiology-like environment. Fig. S1 shows the root mean square deviations (RMSD) for backbone atoms in RBD, V^H and V^L, that all saturated after about 50 ns. Both V^H and V^L regions contain stable secondary structures, as reflected by the saturated RMSD of only about 1 Å. The RBD comprises disordered loops (Fig. 1b) and consequently the saturated RMSD values are larger, around 1.7 Å. Overall, these small RMSD values indicate that the entire complex was stable during the 200 ns equilibration.

Remarkably, at the interface there exists a buried salt-bridge formed by K417 in the RBD and D104 in the V^H region (Fig. 1b and the inset of Fig. 1c). Given the relatively low dielectric constant ($\sim$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 atom NZ in K417 and the atom CG in D104 from the 200 ns simulation trajectory. Fig. 1c demonstrates that after about 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 reform frequently due to their exposure to water. Other key interfacial bindings will be further discussed below.

We also modeled the Fab alone in the 0.15 M NaCl electrolyte (Fig. S2a), for about 200 ns of equilibration. Without the presence of Spro, the side chains of the interfacial residues of the mAb CB6 were orientated differently. For instance, being exposed to water, D104 of the V^H can form a hydrogen bond with Y98 of the V^L from time to time, enhancing internal interactions between V^H and V^L. Similarly, the entire Fab structure alone in simulation was stable, with RMSD values for V^H and V^L saturated at 0.8 Å and 0.9 Å (Fig. S2), respectively.
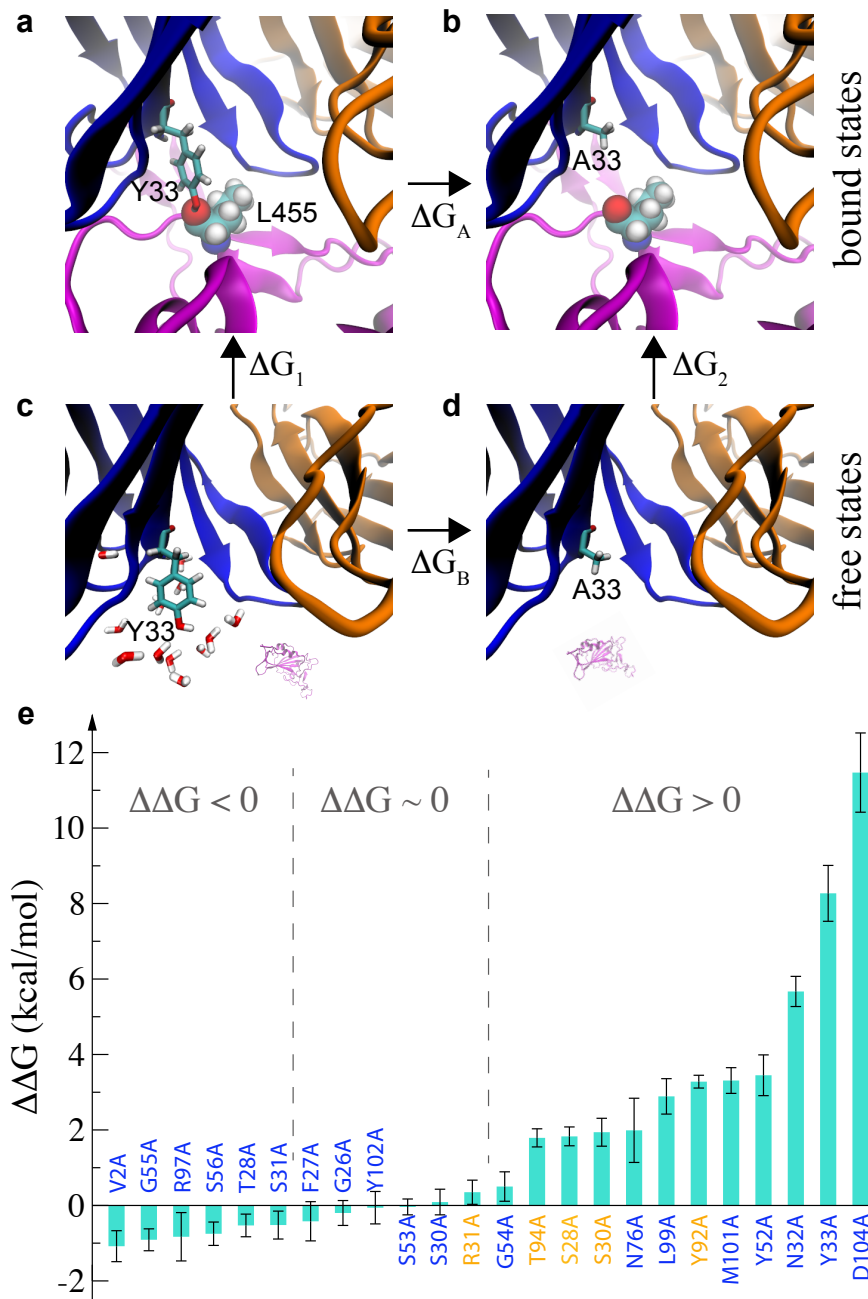
Figure 2: *in silico* alanine scan. a-d) Illustration of the FEP calculations with the mutation Y33A. Protein segments (in cartoon representation) are colored the same as those in Fig. 1b. a) The bound state between the original Fab and RBD. b) The bound state between the mutated Fab (Y33A) and RBD. c) The free state of the original Fab in water. d) The free state of the mutated Fab (Y33A) in water. e) Alanine scan results. Mutations occurred in V^H and in V^L are colored in blue and orange, respectively.

6

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 section 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 the ones in Fab and 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 V^H region, along with S28, S30, R31, Y92 and T94 in the V^L region. Figures 2a-2d show a designed thermodynamical cycle that is used in the FEP method to calculate the free energy difference $\Delta\Delta G$ for the mutation Y33A: ΔG_1 and ΔG_2 are free energy changes for RBD's binding to the Fab and the mutated one, respectively; ΔG_A and ΔG_B are free energy changes for annihilating Y33 and simultaneously exnihilating A33 in the bound and free states, respectively. Generally, to circumvent the difficulty of directly calculating ΔG_1 and ΔG_2, we compute ΔG_A and ΔG_B for the alchemy processes in the bound and free states respectively. Details for calculating $\Delta\Delta G$ ($=\Delta G_A - \Delta G_B$) for each mutation are described in the Methods section.

Figure 2e summaries all alanine-scan results for interfacial residues in V^H (blue text) and in V^L (orange text). When $\Delta\Delta G$ >0 (i.e. $\Delta G_A > \Delta G_B$), 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, the residue D104 in V^H forms a buried salt-bridge with the K417 in the RBD and correspondingly D104A gave rise to a large $\Delta\Delta G$ (=11.47 kcal/mol). Typically, the binding free energy for a solvent-exposed salt-bridge is only about -1 kcal/mol.[26] Here, about an order of magnitude larger value of $\Delta\Delta G$ results from the roughly an order of magnitude less dielectric constant for electrostatic interactions inside a protein than in water. The second largest value of $\Delta\Delta G$ belongs to the Y33A mutation. As shown in Fig. 2a, Y33 forms a stable hydrogen bond with L455 in the RBD, the mutation to alanine in the bound state (Fig. 2b) reduced the interfacial binding and thus is unfavorable. On the other hand, when surrounded by water

7

molecules the hydrophobic Y33 in the free state (Fig. 2c) is disadvantageous and thus its mutation to alanine becomes favorable in the free state. Taken all together, $\Delta\Delta G$= 8.27 kcal/mol, suggesting that Y33 is also essential in stabilizing the Fab-RBD complex.

Additionally, alanine scan for N32, Y52, M101, L99, N76 and G54 in V^H together with Y92, S30, S28, T94 in V^L yielded positive values of $\Delta\Delta G$ (Fig. 2e). Therefore, all these residues contribute substantially for the stable interfacial binding observed in MD simulation. For example, in Fig. S3 (in Supporting Information), we described the molecular mechanism of the mutation N32A on how the interfacial interaction was reduced locally after the mutation. When searching for mutations that can enhance the Fab-RBD binding, we intentionally keep these key residues intact.

When $\Delta\Delta G$ <0 (i.e. $\Delta G_A < \Delta G_B$), each mutation to alanine become more favorable in the bound state than in the free state. Therefore, alanine mutations (for V2, G55, R97, S56, T28 and S31 in V^H, as shown in Fig. 2e) with negative values of $\Delta\Delta G$ are possible candidates for increasing the Fab-RBD binding. To account for negative $\Delta\Delta G$ values, 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 less internal constraints inside the Fab. For instance, the $\Delta\Delta G$ for V2A is -0.96 kcal/mol (Fig. 2e) and from the snapshots taken before and after the mutation (Figs. S4a and S4b in Supporting Information), we observe that originally (before the V2A mutation) Y108 inside V^H 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 about 0.5 kcal/mol (Fig. S4c), which suggests an improved interfacial binding (echoing with the negative $\Delta\Delta G$ for V2A) . Similarly, R97A yielded a ΔG_B of 0.83 kcal/mol larger than ΔG_A in the bound state (i.e., $\Delta\Delta G$= -0.83 kcal/mol, Fig. 2e) since 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 (Fig. 2e), their mutations to alanine were fruitless with negligible values (~ 0) of $\Delta\Delta G$, which indicates that these residues, despite being close to the interface, are dispensable in the Fab-RBD binding. Indeed, these residues located at peripheral areas of the interfacial contact and were more exposed to water than contacted by residues in RBD. Nevertheless, mutations of these residues to other ones might offer unforeseen opportunities for enhancing the Fab-RBD binding.

Among these residues, we chose S30 and G26 in V^H for performing enumerated mutations to other residues (Figs. 3a and 3b). Majority of mutations for S30 (such as S30L and S30 T) produced positive values of $\Delta\Delta G$ (Fig. 3a), i.e. weakening the Fab-RBD binding. For the mutation S30G, $\Delta\Delta G \sim 0$ because G30 is even smaller than alanine and thus became more trivial in the 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 (Fig. 4a). Due to the geometric constraints, in the mutation of S30E we did not observed its salt-bridge formation between E30 and K458, and correspondingly $\Delta\Delta G$ for S30E is 0.38 kcal/mol (i.e. unfavorable). While the advantage of S30D can be easily recognized, the molecular mechanism for S30M is not intuitive. As shown in Fig. 4b, the direct interaction between the charged K458 residue and the hydrophobic M30 one 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 one. Overall, S30M stabilized the local structure in the bound state and resulted in a negative $\Delta\Delta G$ value.
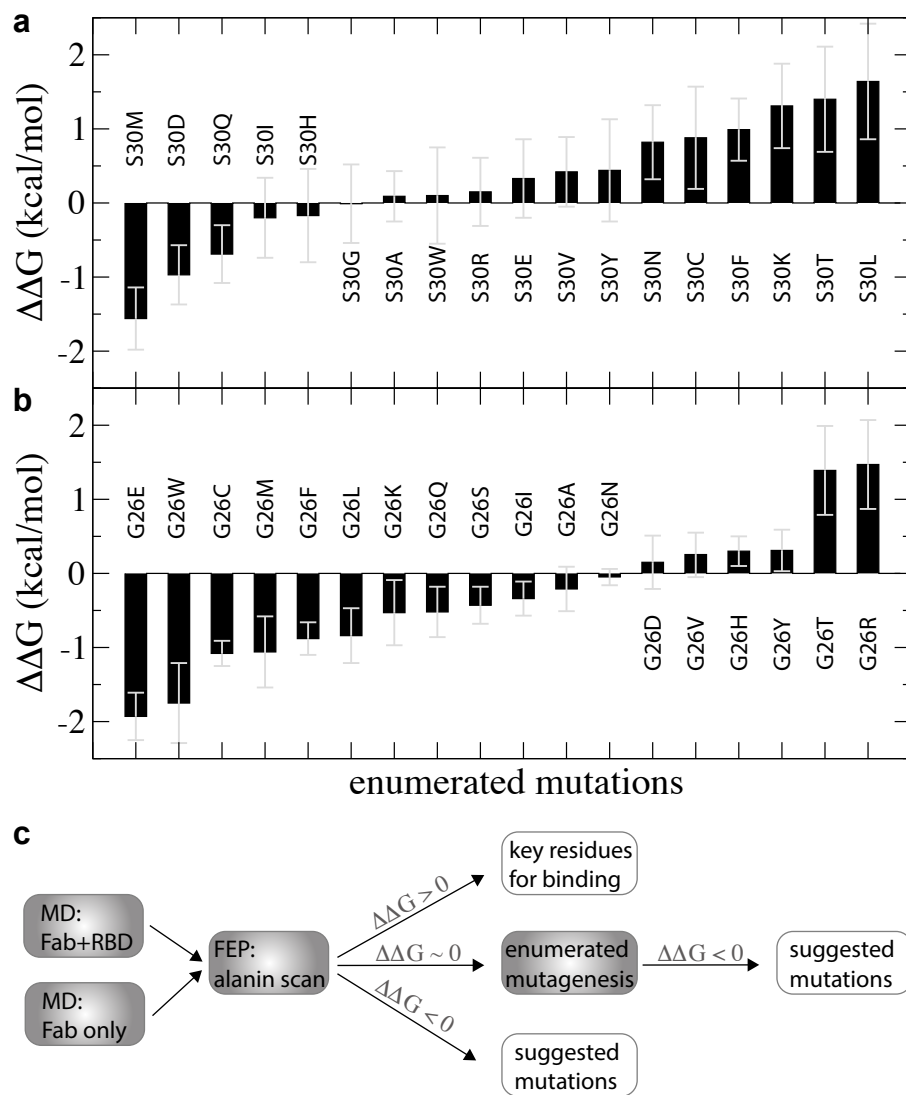
Figure 3: FEP calculations for non-essential residues. a) Enumerated mutations for S30 in the V^H domain. b) Enumerated mutations for G26 in the V^H domain. c) The *in silico* workflow for identifying possible mutations that enhance the antibody-Spro binding.

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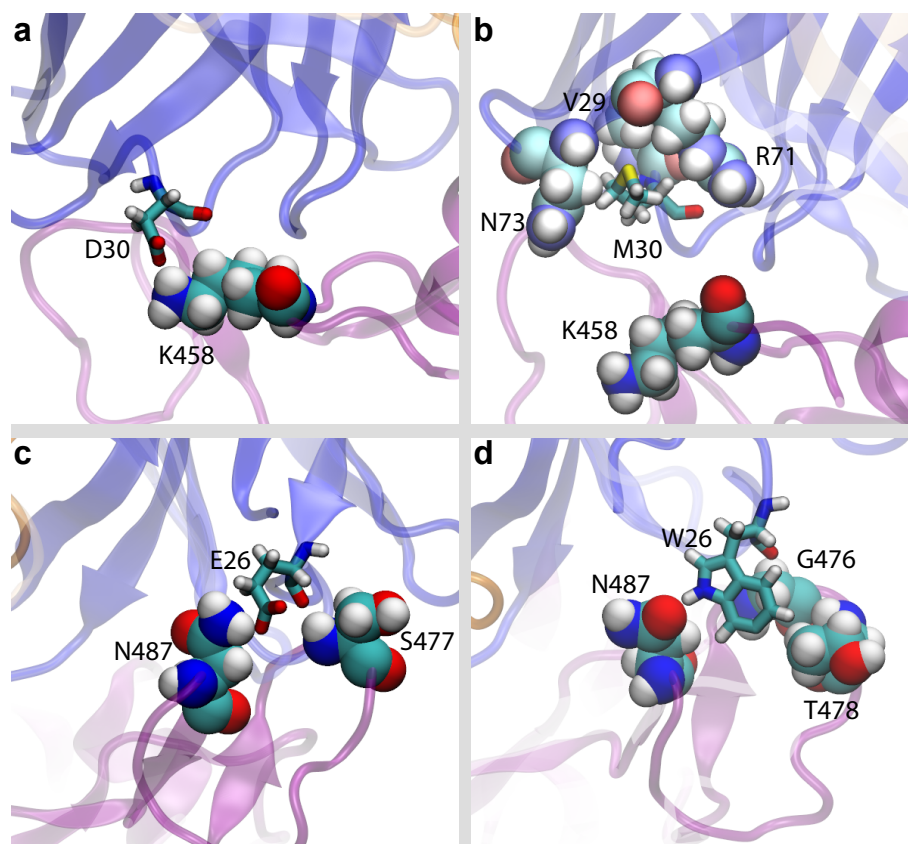

Figure 4: Illustrations of possible molecular mechanisms of mutations that enhance the Fab-RBD binding. a) S30D. b) S30M. c) G26E. d) G26W.

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 (Fig. 3b). Since G26 (comprising only one hydrogen atom in its side chain) is the smallest one among all amino acids, it only interacted weakly

with surface residues in the RBD as indicated in the alanine scan. When mutating G26 into other more bulky residues, it is foreseeable to strengthen the interfacial Fab-RBD interaction. Particularly, the $\Delta\Delta G$ values for G26E and G26W are respectively -1.93 and -1.75 kcal/mol, significantly improving the local interfacial binding. The molecular mechanism for G26E is illustrated in Fig. 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 bulky one) can form a hydrogen bond between the carboxamide group in the side chain of N487 (in the RBD) with its indole nitrogen (-NH-) group. Additionally, W26 was in contact with G476 and T478 (in the RBD) via the hydrophobic interaction (Fig. 4d). On the other hand, it is expected that the exnihilation of the hydrophobic W26 in the free state is energetically unfavorable.

In case that we had not obtained any new mutations for G26 and S30 to enhance the Fab-RBD binding, other residues with their $\Delta\Delta G$ close to zero or even negative from the alanine scan (Fig. 2e) should be further explored with enumerated mutations. In summary of the strategy discussed above for identifying possible mutations for the more efficacious Fab-RBD binding, we illustrated the entire *in silico* workflow in Fig. 3c. Overall, using this workflow, we identified several encouraging mutations such as V2A, G55A, R97A, S30M, S30D, G26E, G26W, G26C, G26M and so on. Furthermore, it is possible to combine two or three of these favorable mutations together to further enhance the binding affinity.

In conclusion, we proposed an *in silico* approach for optimizing the binding between a designed antibody and Spro (particularly the RBD). Taking advantages of ever increasing computing power, we performed all-atom MD simulations as well as FEP calculations for alanine scan and enumerated mutations, which yielded several promising candidates for optimizing the mAb CB6. Different from previous studies (e.g. ref.[27] that relies on human expertise to select possible mutations), here we demonstrated that the entire workflow (as shown in Fig. 3c) can be easily automated on high-performance clusters (HPC) or super-

computers (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 future.

One potential risk of applying mAb drugs for the therapeutic treatment of COVID-19 is the so called antibody-dependent enhancement (ADE),which can be triggered by vaccines as well. 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.[28] 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 Fig. 3c to find residues in the Fc domain with positive $\Delta\Delta G$ 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.

So far, 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, it is hopeful to accelerate the discovery of safe and efficacious mAb drugs for both therapeutic and prophylactic applications. Available in large-scale production, designed mAb drugs (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[15] bound to RBD of Spro) and free (stand alone Fab of mAb CB6) states using the NAMD2.13 package[29] running on the IBM Power Cluster. To model the Fab-RBD complex (a bound state), we obtained the previously resolved crystal structure (PDB code: 7C01)[15] from the protein data bank. After solvating the complex in a rectangular water box that measures about $76 \times 76 \times 134.76$ Å^3, 66 Na$^+$ and 71 Cl$^-$ were added into the system, neutralizing the entire simulation system and setting the ion concentration to be 0.15 M (Fig. 1b). The final system containing 79,466 atoms was first minimized for 10 ps and further equilibrated for 200 ps in the NPT ensemble ($P \sim 1$ bar and $T \sim 300$ K), with atoms in the backbones harmonically constrained (spring constant $k=1$ kcal/mol/Å^2). After removing constraints on the atoms in V^H, V^L and RBD domain, 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 C^L and C^H domains (not close to the Fab-RBD interface) remained to be harmonically restrained (spring constant $k=1$ kcal/mol/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 field[30] was applied for proteins; the TIP3P model[31,32] was chosen for water; the standard force field[33] 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 the grid size about 1 Å in each dimension. The van der Waals (vdW) energies between atoms were calculated using a smooth (10-12 Å) cutoff. The temperature T was kept at 300 K by applying the Langevin thermostat,[34] while the pressure was kept constant at 1 bar using the Nosé-Hoover method.[35] With the SETTLE algorithm[36] enabled to keep all bonds rigid, the simulation time-step was set to be 2 fs for bonded and non-bonded (e.g. vdW, angle and dihedral) interactions, and electric interactions were calculated every 4 fs, with the multiple time-step algorithm.[37]

Free energy perturbation calculations The free energy perturbation (FEP) method[38] has been previously used for *in silico* mutagenesis studies of proteins.[27] After equilibrating the bound and free states in respective MD simulations, we employed the FEP method to calculate the change of the binding free energy for various mutations on the Fab of the antibody CB6.

As shown in the thermodynamic cycle for the Y33A mutation (Figs.2a-2d), the difference between RBD's binding free energies can be calculated by the following equation,

$$\Delta\Delta G = \Delta G_2 - \Delta G_1 = \Delta G_A - \Delta G_B, \tag{1}$$

Generally, direct calculations of ΔG_1 and ΔG_2 are challenging and can be replaced by computing ΔG_A and ΔG_B instead (Eq. 1). From the following ensemble average,[38] ΔG_1 and ΔG_2 can be calculated theoretically,

$$\Delta G_{1,2} = -k_{\mathrm{B}}T\ln < \exp(\frac{H_f - H_i}{k_{\mathrm{B}}T}) >_i, \tag{2}$$

where k_{B} is the Boltzmann constant; T the temperature; H_i and H_f the Hamiltonians at the initial (i) and the 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 the one with its Y33 replaced by A33. Using the perturbation method, many intermediate stages (denoted by λ) whose Hamiltonian $H(\lambda)=\lambda H_f+(1-\lambda)H_i$ are required between initial and final states to improve the accuracy. In FEP calculations of ΔG_A and ΔG_B with the soft-core potential enabled, λ 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 exnihilating 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 up to 10 independent runs for each mutation and accepted the lowest five free energy changes for calculating the mean and the error.

Competing Interests

T. H. and B. L. declare no conflicts of interest.

Acknowledgement

Supporting Information Available

Fig. S1 showing RMSDs for the RBD, V^H and V^L in the Fab-RBD complex; Fig. S2 showing the MD simulation system for the Fab only and RMSDs for V^H and V^L; Fig. S3 showing the molecular mechanism of the mutation N32A; Fig. S4 showing the molecular mechanism of the mutation V2A; Fig. S5 showing the salt-bridge formed by the phosphorylated S30 in the V^H and K458 in the RBD.

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. *New England Journal of Medicine* **2020**, *382*, 727–733.

3. Beigel, J. H.; Tomashek, K. M.; Dodd, L. E.; Mehta, A. K.; Zingman, B. S.; Kalil, A. C.; Hohmann, E.; Chu, H. Y.; Luetkemeyer, A.; Kline, S. *et al.* Remdesivir for the treatment of Covid-19—preliminary report. *New England Journal of Medicine* **2020**,

4. Villar, J.; Ferrando, C.; Martínez, D.; Ambrós, A.; Muñoz, T.; Soler, J. A.; Aguilar, G.; Alba, F.; González-Higueras, E.; Conesa, L. A. *et al.* Dexamethasone treatment for the acute respiratory distress syndrome: a multicentre, randomised controlled trial. *The Lancet Respiratory Medicine* **2020**, *8*, 267–276.

5. Cao, Y.; Su, B.; Guo, X.; Sun, W.; Deng, Y.; Bao, L.; Zhu, Q.; Zhang, X.; Zheng, Y.; Geng, 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*, 7–9.

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. USA* **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**,

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.; Sakharkar, 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 Straten, K.; Snitselaar, J.; Aldon, Y.; Bangaru, S.; Torres, J.; Okba, 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.; Christian, G.; Frauke, M.; Julio, C.; Zijun, W.; Alice, C.; Marianna, A.; Christopher, O. B.; Anna, G.; Shlomo, F. *et al.* Convergent Antibody Responses to SARS-CoV-2 in Convalescent Individuals. *Nature* **2020**, *584*, 437–442.

14. Yuan, M.; Wu, N. C.; Zhu, 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 Kuppeveld, F. J.; Haagmans, B. L.; Grosveld, F.; Bosch, B.-J. A human monoclonal antibody blocking SARS-CoV-2 infection. *Nature communications* **2020**, *11*, 1–6.

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.; Wloga, 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. Karplus, M.; McCammon, J. Molecular dynamics simulations of biomolecules. *Nature Struct. Biol.* **2002**, *265*, 654–652.

24. Freddolino, P. L.; Harrison, C. B.; Liu, Y.; Schulten, K. Challenges in protein-folding simulations. *Nature 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. Letts* **2020**, *11*, 4413–4420.

26. Horovitz, A.; Serrano, L.; Avron, B.; Bycroft, M.; Fersht, A. R. Strength and co-operativity of contributions of surface salt bridges to protein stability. *J. Mol. Biol.* **1990**, *216*, 1031–1044.

27. Luan, B.; Xu, G.; Feng, M.; Cong, L.; Zhou, R. Combined Computational–Experimental

Approach to Explore the Molecular Mechanism of SaCas9 with a Broadened DNA Targeting Range. *J. Am. Chem. Soc.* **2019**, *141*, 6545–6552.

28. Arvin, A. M.; Fink, K.; Schmid, M. A.; Cathcart, A.; Spreafico, R.; Havenar-Daughton, C.; Lanzavecchia, A.; Corti, D.; Virgin, H. W. A perspective on potential antibody-dependent enhancement of SARS-CoV-2. *Nature* **2020**, 1–14.

29. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comp. Chem.* **2005**, *26*, 1781–1802.

30. MacKerell, A., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

31. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

32. Neria, E.; Fischer, S.; Karplus, M. Simulation of Activation Free Energies in Molecular Systems. *J. Chem. Phys.* **1996**, *105*, 1902–1921.

33. Beglov, D.; Roux, B. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. *J. Chem. Phys.* **1994**, *100*, 9050–9063.

34. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987.

35. Martinetz, T.; Schulten, K. Topology Representing Networks. *Neur. Netw.* **1994**, *7*, 507–522.

36. Miyamoto, S.; Kollman, P. A. SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Molecules. *J. Comp. Chem.* **1992**, *13*, 952–962.

37. Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible multiple time scale molecular dynamics. *The Journal of chemical physics* **1992**, *97*, 1990–2001.

38. Chipot, C.; Pohorille, A. *Free energy calculations*; Springer, 2007.