All-Atom Simulations and Free-Energy Calculations of Antibodies Bound to the Spike Protein of SARS-CoV-2: The Binding Strength and Multivalent Hydrogen-Bond Interactions

Hwankyu Lee

1. Introduction

The coronavirus disease 2019 (COVID-19), which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has been a serious pandemic all over the world.\cite{1} 110 million people have been infected for last 14 months, leading to deaths of more than 2.5 million people as of February/2021. In fact, this pandemic nowadays gets worse, since the winter season is coming on the northern hemisphere region. To end this disease pandemic, several vaccine candidates have been proposed and even tested on clinical trials.\cite{2} In particular, the receptor-binding domain (RBD) of the spike (S) protein, which is the major surface antigen of SARS-CoV-2, is known to bind to the host receptor angiotensin-converting enzyme 2 (ACE2) that is a type I membrane protein expressed in lungs, heart, kidneys, and intestine, leading to viral infection.\cite{3-6} Therefore, RBD-targeting antibodies have been developed to neutralize SARS-CoV-2 by inhibiting the binding between RBD and ACE2.

Yan et al. determined cryo-electron microscopy structures of full-length human ACE2 bound to the RBD of the S protein of SARS-CoV-2,\cite{5} which has motivated many experimental studies on the development of SARS-CoV-2 RBD-targeting antibodies since early this year.\cite{7-15} For instance, Yuan et al. determined the crystal structure of an antibody CR3022, which is isolated from an SARS patient, complexed with the RBD of SARS-CoV or SARS-CoV-2, showing a highly conserved epitope that enables cross-reactive binding between SARS-CoV and SARS-CoV-2.\cite{13} Ju et al. also isolated 206 antibodies from SARS-CoV-2 patients and characterized crystal structures of antibodies complexed with SARS-CoV-2 RBD, showing successful neutralization of those antibodies to SARS-CoV-2.\cite{15} Rogers et al. developed rapid platforms to evaluate plasma neutralization activity against SARS-CoV-2 and then isolated highly potent neutralizing antibodies.\cite{16} Recently, Yuan et al. found that immunoglobulin G heavy-chain variable 3-53 (IGHV3-53) is the most popular IGHV gene among 294 SARS-CoV-2 RBD-targeting antibodies and determined crystal structures of RBDs bound to IGHV3-53-neutralizing antibodies, CC12.1 and CC12.3.\cite{17} In particular, they analyzed hydrogen-bond interactions of RBD with CC12.1 and CC12.3, which were further compared with those of RBD with ACE2 and another antibody, B38.\cite{17} Although these experiments have provided vital information on the structure of SARS-CoV-2 RBDs and their interactions with antibodies, results from these experiments are not always easy to interpret at the level of specific interactions between individual amino acids of proteins under the explicit solvent condition, which is necessary to develop vaccine or drug molecules targeting the S protein of SARS-CoV-2. An accurate accounting of this requires simulations at nearly atomistic resolution which typically require all-atom simulations.

As a further step toward understanding specific interactions controlling the binding between SARS-CoV-2 RBD and antibodies, here we report all-atom molecular dynamics (MD) simulations of antibodies (CC12.1, CC12.3, and B38) and
ACE2 bound to SARS-CoV-2 RBD. Binding free energies are calculated from umbrella sampling simulations, showing different binding strengths between RBDs and antibodies, which are further confirmed by calculating solvent accessible surface areas (SASAs). In particular, the number and strength of hydrogen bonds between RBDs and antibodies are quantified, showing the importance of multivalent hydrogen-bond interactions and suggesting key amino acids of RBD for the strong RBD-antibody binding, which helps in the rational design of vaccine and drug molecules targeting the S protein of SARS-CoV-2.

2. Results and Discussion

2.1. Free Energy Calculations: The Binding Strength between RBD and Antibody

To compare the binding strengths of SARS-CoV-2 RBDs complexed with ACE2 and various antibodies such as CC12.1, CC12.3, and B38, potentials of mean force (PMFs) were obtained from umbrella sampling simulations of 30 windows per each system (a total of 2.4 µs for four systems), which were calculated as a function of the distance between RBD and antibody centers, and then used to calculate the binding free energies. In Figure 1, the lowest PMFs are found when RBD and antibody are closely positioned, indicating the strong binding between all simulated antibodies and RBDs, as observed in Yuan et al.’s experiments. However, free energies of CC12.1 and B38 are lower than those of CC12.3 and ACE2, indicating the stronger binding of RBD with CC12.1 and B38 than with CC12.3 and ACE2. To confirm this binding tendency, SASAs of the antibody-binding residues of RBD (406th to 505th amino acids) were calculated, which measures the surface area uncovered by antibodies. Here, the radius of the solvent probe is set to 0.14 nm. Figure 2 shows that when RBDs bind to antibodies, their SASA values significantly decrease, again confirming the tight binding between RBDs and antibodies. In particular, SASAs of RBDs bound to antibodies are lower for CC12.1 and B38 than for CC12.3 and ACE2, indicating the tighter binding of RBD with CC12.1 and B38 than with CC12.3 and ACE2, consistent with free energy calculations in Figure 1. These results indicate that all simulated antibodies bind to SARS-CoV-2 RBD, but their binding strengths slightly differ in the order of B38 > CC12.1 > CC12.3 > ACE2.
2.2. Structures and Binding Sites of SARS-CoV-2 RBD

The strong binding between RBD and antibody may influence the secondary structure of RBD. To understand this, secondary structures of RBDs bound or unbound to antibodies were calculated using the Dictionary of Secondary Structure of Proteins program. Figure 3 shows that RBDs interacting with antibodies predominantly have random coil, turn, and β-sheet structures, regardless of different antibodies. In particular, structures of RBDs bound to antibodies do not significantly differ from those of RBDs unbound to antibodies, indicating that structures of RBDs are not influenced by their interactions with antibodies, although it cannot be ruled out that structural changes might eventually occur for longer simulations.

To understand the interactions between SARS-CoV-2 RBDs and antibodies, minimum distances between individual amino acids of RBD and antibody were calculated. In Figure 4, RBD residues with the minimum distance of less than 0.4 nm are considered to be the binding residues and categorized according to their electrostatics and hydrophobicity. Although there are a few charged or hydrophobic RBD residues bound to antibodies, polar uncharged residues predominantly bind to antibodies, implying that hydrogen-bond interactions of these hydrophilic residues are important for the RBD-antibody binding.

2.3. Multivalent Hydrogen-Bond Interactions of Specific Amino Acids

As discussed above, the RBD-antibody binding may be stabilized by hydrogen-bond interactions. Experimentally, Yuan et al. observed hydrogen bonds between RBDs and antibodies, although their binding strength was not quantified particularly under explicit solvent condition. To resolve this, we calculated numbers and lifetimes of hydrogen bonds between RBDs and antibodies. Here, we assume that a hydrogen-bonding interaction exists when the donor–acceptor distance is <0.35 nm and the angle of the hydrogen-donor–acceptor triplet is <30°. Note
Minimum distances between SARS-CoV-2 RBDs (≈406th–505th residues) and antibodies. For RBD residues with the minimum distance of less than 0.4 nm (dotted red lines), anionic, cationic, polar uncharged, and hydrophobic residues are colored in red, blue, purple, and green, respectively.

Figure 4. Minimum distances between SARS-CoV-2 RBDs and antibodies. For RBD residues with the minimum distance of less than 0.4 nm (dotted red lines), anionic, cationic, polar uncharged, and hydrophobic residues are colored in red, blue, purple, and green, respectively.

that other criteria with the distance from 0.3 to 0.4 nm and the angle from 20° to 40° produce similar qualitative trends, confirming that the analysis does not significantly depend on the distance and angle criteria. Figure 5 shows the number of hydrogen bonds reach steady-state values within the simulated timescale, indicating that the RBD-antibody binding is well equilibrated. More hydrogen bonds are observed for CC12.1 and B38 than for CC12.3 and ACE2, favorably compared with free energy calculations showing the stronger binding of RBD with CC12.1 and B38 than with CC12.3 and ACE2. To quantify the strength of hydrogen-bond interactions, average lifetimes of hydrogen bonds between individual amino acids of RBD and antibody were calculated, which is the inverse of the rate constant for hydrogen-bond kinetics (breaking) described by the autocorrelation function of the hydrogen-bonding existence functions (either 0 or 1) averaged over all hydrogen bonds.\(^{[21,22]}\) If residues of SARS-CoV-2 RBD have lifetimes of longer than 1 ns for their hydrogen bonds with any antibodies, those RBD residues are considered to be key amino acids controlling hydrogen-bond interactions between RBDs and antibodies. In Figure 6, hydrogen bonds with relatively long lifetimes exist in the ≈417th–487th residues of CC12.3 and in the ≈487th–505th residues of ACE2, while CC12.1 and B38 show long hydrogen-bond lifetimes in the whole range of ≈406th–505th residues, indicating that the strong binding between RBDs and antibodies can be achieved by multivalent hydrogen-bond interactions rather than by locally formed hydrogen bonds, as illustrated in Figure 7. These findings indicate that the binding between RBD and antibody is stabilized by multivalent hydrogen-bond interactions of polar uncharged RBD residues, suggesting key amino acids of RBD controlling the RBD-antibody binding, which helps in the rational design

**Figure 5.** Number of hydrogen bonds between SARS-CoV-2 RBDs and antibodies as a function of time.
of vaccine and drug molecules targeting the S protein of SARS-CoV-2.

3. Conclusions

We performed all-atom MD simulations of SARS-CoV-2 RBD bound to ACE2 and IGHV gene-based antibodies such as CC12.1, CC12.3, and B38. Binding free energies were calculated from umbrella sampling simulations, showing that all simulated antibodies and ACE2 bind to SARS-CoV-2 RBD, as observed in experiments, although their binding strengths slightly differ in the order of B38 > CC12.1 > CC12.3 > ACE2. SASAs of RBD bound or unbound to antibodies were calculated, showing the reduced SASA values when bound to antibodies. Polar uncharged residues of RBD more predominantly bind to antibodies than do charged or hydrophobic residues of RBD, showing that the RBD-antibody binding is influenced by hydrophilic interactions rather than by electrostatic or hydrophobic interactions. To understand this, the number and strength of hydrogen bonds between RBDs and antibodies were calculated, showing that the RBD-antibody complex is stabilized by multivalent hydrogen bonds of \( \approx 406\text{th} - 505\text{th} \) RBD residues (CC12.1 and B38) rather than by local hydrogen bonds of \( \approx 417\text{th} - 487\text{th} \) (CC12.3) or \( \approx 487\text{th} - 505\text{th} \) (ACE2) residues. In particular, 17 key residues of RBD for strong hydrogen-bond interactions between RBDs and antibodies are observed, which helps in the rational design of vaccine and drug molecules that can target the S protein of SARS-CoV-2.

4. Computational Methods

All simulations and analyses were performed using the GROMACS-2018.6 simulation package\(^{[23-25]}\) with the optimized potential for liquid simulations (OPLS) all-atom force field (FF) and the TIP4P water model.\(^{[26,27]}\) The coordinates of the RBD complexed with CC12.1, CC12.3, B38, and ACE2 were downloaded from the Protein Data Bank (PDB code: 6XC2,\(^{[17]}\) 6XC4,\(^{[17]}\) 7BZ5,\(^{[12]}\) and 6M0J,\(^{[3]}\) respectively). The SARS-CoV-2 RBD-antibody complex was solvated with \( \approx 107000 \) water molecules in a periodic box of size 15 nm/side, which were neutralized by adding counterions and additional ions of 0.15 M NaCl that mimics the physiological condition. PMFs were calculated using the umbrella sampling algorithm, which is a computational method to overcome the potential barrier and sample all possible configurations by applying harmonic restraint to the molecule of interest via a biased potential with respect to the reference
molecule.[28] In the RBD-antibody complex, RBD was pulled toward the bulk-water region (2.5 nm) and also pulled toward the antibody surface (0.5 nm) with a force constant of 1000 kJ mol⁻¹ nm⁻², which yields a total of 30 sample configurations (called windows) with a window spacing of 0.1 nm in the distance of 3 nm. These 30 windows were equilibrated for 0.1 ns and then used as starting configurations for umbrella sampling simulations. Each window was simulated for 20 ns with a time step of 2 fs (a total of 600 ns for 30 windows per each RBD-antibody complex) on computational facilities supported by the National Institute of Supercomputing and Networking/Korea Institute of Science and Technology Information with supercomputing resources including technical support (KSC-2020-CRE-0202). A temperature of 310 K and a pressure of 1 bar were maintained by applying a velocity-rescale thermostat[30] and Parrinello–Rahman barostat[31] in the NPT ensemble. A real space cutoff of 1 nm was applied for Lennard-Jones and electrostatic forces with the inclusion of particle mesh Ewald summation for long-range electrostatics.[31] The LINCS algorithm was used to constraint the bond length. [32,33] The last 10 ns trajectories were used to unbiased umbrella sampling using the weighted histogram analysis method.[34] Errors were estimated from the bootstrapping analysis, called the Bayesian bootstrapping of complete histograms, where random weights are assigned to all histograms within each bootstrap.[35]

Conflict of Interest
The author declares no conflict of interest.

Data Availability Statement
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords
antibodies, binding free energy, drug discovery, molecular dynamics simulations, protein-protein interaction, SARS-CoV-2

Received: January 14, 2021
Revised: February 19, 2021
Published online: March 26, 2021

[1] S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Science 2020, 368, 860.
[2] N. Lurie, M. Saville, R. Hatchett, J. Halton, N. Engl. J. Med. 2020, 382, 1969.
[3] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, Nature 2020, 581, 215.
[4] M. Letko, A. Marzi, V. Munster, Nat. Microbiol. 2020, 5, 562.
[5] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Science 2020, 367, 1444.
[6] P. Zhou, X. L. Yang, X. G. Wang, B. Hu, L. Zhang, W. Zhang, H. R. Si, Y. Zhu, B. Li, C. L. Huang, H. D. Chen, J. Chen, Y. Luo, H. Guo, R. D. Jiang, M. Q. Liu, Y. Chen, X. R. Shen, X. Wang, X. S. Zheng, K. Zhao, Q. J. Chen, F. Deng, L. L. Liu, B. Yan, F. X. Zhan, Y. Y. Wang, G. F. Xiao, Z. L. Shi, Nature 2020, 579, 270.
[7] Y. Cao, B. Su, X. Guo, W. Sun, Y. Deng, L. Bao, Q. Zhu, X. Zhang, Y. Zheng, C. Geng, X. Chai, R. He, X. Li, Q. Lv, H. Zhu, W. Deng, Y. Geng, Y. Wang, L. Qiao, Y. Tan, L. Song, G. Wang, X. Du, N. Gao, J. Liu, J. Xiao, X. X. Su, Z. Yu, Y. Feng, C. Qin, R. Jin, X. S. Xie, Cell 2020, 182, 73.e16.
[8] X. Chi, R. Yan, J. Zhang, G. Zhang, Y. Zhao, M. Hao, Z. Zhang, P. Fan, Y. Dong, Y. Yang, Z. Chen, Y. Guo, Y. Li, X. Song, Y. Chen, L. Xia, L. Fu, L. Hou, J. Xu, C. Yu, J. Li, Q. Zhou, W. Chen, Science 2020, 369, 650.
[9] D. Pinto, Y. J. Park, M. Beltramello, A. C. Walls, S. A. Rawlings, N. C. Wu, M. Yuan, D. Pinto, Y. J. Park, M. Beltramello, A. C. Walls, S. A. Rawlings, N. C. Wu, M. Yuan, D. F. Robbiani, C. Gaebler, M. C. Nussenzweig, Nature 2020, 584, 437.
[10] R. Shi, C. Shan, X. Duan, Z. Chen, P. Liu, J. Song, T. Song, X. Bi, C. Han, L. Wu, G. Gao, X. Hu, Y. Zhang, Z. Tong, W. Huang, W. J. Liu, G. Wu, B. Zhang, L. Wang, J. Qi, H. Feng, F. S. Wang, Q. Wang, G. F. Gao, Z. Yuan, J. Yan, Nature 2020, 584, 120.
[11] Y. Wu, F. Wang, C. Shen, W. Peng, D. Li, C. Zhao, Z. Li, S. Li, Y. Bi, Y. Yang, X. Cong, H. Xiao, Z. Fan, S. Tan, C. Wu, W. Tan, X. Lu, C. Fan, Q. Pack, Y. Liu, C. Zhang, J. Qi, G. F. Gao, F. G. Gao, L. Liu, Science 2020, 368, 1274.
[12] M. Yuan, N. C. Wu, X. Zhu, C. C. D. Lee, R. T. Y. So, H. Lv, C. K. P. Mok, I. A. Wilson, Science 2020, 368, 630.
[13] S. J. Zost, P. Gulchik, R. E. Chen, J. B. Case, J. X. Reidy, A. Trivette, R. S. Nargi, R. E. Sutton, N. Suryadevara, E. C. Chen, E. Binshtein, S. Shrihari, M. Ostrowski, H. Y. Chu, J. E. Didier, K. W. MacRenaris, T. Jones, S. Day, L. Myers, F. Eun-Hyung Lee, D. C. Nguyen, I. Sanz, D. R. Martinez, P. W. Rothlauf, L. M. Bloyet, S. P. J. Whelan, R. S. Baric, L. B. Thackray, M. S. Diamond, R. H. Carnahan, J. E. Crowe, Jr., Nat. Med. 2020, 26, 1422.
[14] B. Ju, Q. Zhang, J. Ge, R. Wang, J. Sun, X. Ge, J. Yu, S. Shan, B. Zhou, S. Song, X. Tang, J. Yu, J. Lan, J. Yuan, H. Wang, J. Zhao, S. Zhang, Y. Wang, X. Shi, L. Liu, J. Zhao, X. Wang, Z. Zhang, L. Zhang, Nature 2020, 584, 115.
[15] T. F. Rogers, F. Zhao, D. Huang, N. Beutler, A. Burns, W. T. He, O. Limbo, C. Smith, C. Goud, S. Weiloh, L. Yang, R. K. Abbott, S. Callaghan, E. Garcia, J. Hurtado, M. Parren, L. Peng, S. Ramirez, J. Ricketts, M. J. Ricciardi, S. A. Rawlings, N. C. Wu, M. Yuan, D. M. Smith, D. Nemazee, J. R. Teijaro, J. E. Voss, I. A. Wilson, R. Andrahi, B. Briney, E. Landais, D. Sok, J. G. Jardine, D. R. Burton, Science 2020, 369, 956.
[16] M. Yuan, H. Liu, N. C. Wu, C. C. D. Lee, X. Zhu, F. Zhao, D. Huang, W. Yu, Y. Hua, H. Tien, T. F. Rogers, E. Landais, D. Sok, J. G. Jardine, D. R. Burton, I. A. Wilson, Science 2020, 369, 1119.
[17] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graphics 1996, 14, 33.
[18] G. A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structures, Springer-Verlag, Berlin 1991.
[19] M. L. luzar, D. Chandler, Phys. Rev. Lett. 1996, 76, 928.
[20] M. L. luzar, J. Chem. Phys. 2000, 113, 10663.
[21] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, J. Chem. Theory Comput. 2008, 4, 435.
[22] E. Lindahl, B. Hess, D. van der Spoel, J. Mol. Simul. 2001, 7, 306.
[25] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, J. Comput. Chem. 2005, 26, 1701.
[26] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, J. Am. Chem. Soc. 1996, 118, 11225.
[27] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, J. Phys. Chem. B 2001, 105, 6474.
[28] G. M. Torrie, J. P. Valleau, J. Comput. Phys. 1977, 23, 187.
[29] G. Bussi, D. Donadio, M. Parrinello, J. Chem. Phys. 2007, 126, 014101.
[30] M. Parrinello, A. Rahman, J. Appl. Phys. 1981, 52, 7182.
[31] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, J. Chem. Phys. 1995, 103, 8577.
[32] B. Hess, J. Chem. Theory Comput. 2008, 4, 116.
[33] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, J. Comput. Chem. 1997, 18, 1463.
[34] J. S. Hub, B. L. De Groot, D. Van Der Spoel, J. Chem. Theory Comput. 2010, 6, 3713.
[35] M. R. Chernick, Bootstrap Methods: A Guide for Practitioners and Researchers, Wiley-Interscience, Hoboken, NJ 2008, p. 369.