Molecular Recognition of SARS-CoV-2 Spike Glycoprotein: Quantum Chemical Hot Spot and Epitope Analyses

Chiduru Watanabe†‡*, Yoshio Okiyama§, Shigenori Tanaka¶, Kaori Fukuzawa†, Teruki Honma†

†Center for Biosystems Dynamics Research, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
‡JST PRESTO, 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan
§Division of Medicinal Safety Science, National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa 210-9501, Japan
¶Department of Computational Science, Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan
§School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan

KEYWORDS: fragment molecular orbital method; intermolecular interaction, inter-fragment interaction energy; protein-protein interaction; SARS-CoV-2; coronavirus disease 2019; spike protein; angiotensin-converting enzyme 2

ABSTRACT: Due to the COVID-19 pandemic, researchers have attempted to identify complex structures of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike glycoprotein (S-protein) with angiotensin-converting enzyme 2 (ACE2) or a blocking antibody. However, the molecular recognition mechanism—critical information for drug and antibody design—has not been fully clarified at the amino acid residue level. Elucidating such a microscopic mechanism in detail requires a more accurate molecular interpretation that includes quantum mechanics to quantitatively evaluate hydrogen bonds, XH/π interactions (X = N, O, and C), and salt bridges. In this study, we applied the fragment molecular orbital (FMO) method to characterize the SARS-CoV-2 S-protein binding interactions with not only ACE2 but also the B38 Fab antibody involved in ACE2-inhibitory binding. By analyzing FMO-based interaction energies along a wide range of binding interfaces carefully, we identified amino acid residues critical for molecular recognition between S-protein and ACE2 or B38 Fab antibody. Importantly, hydrophobic residues that attribute to weak interactions such as CH-O and XH/π interactions, as well as polar residues that construct conspicuous hydrogen bonds, play important roles in molecular recognition and binding ability. Moreover, through these FMO-based analyses, we also clarified novel hot spots and epitopes that had been overlooked in previous studies by structural and molecular mechanical approaches. Altogether, these hot spots/epitopes identified between S-protein and ACE2/B38 Fab antibody may provide useful information for future antibody design and small or medium drug design against the SARS-CoV-2.

1. INTRODUCTION

Coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has been declared a pandemic by the World Health Organization and has caused worldwide social and economic problems1. Despite its significant infectious strength and the worldwide research efforts, to date, no specific treatment has been established against this new virus. The SARS-CoV-2 is composed of a variety of proteins, including the spike glycoprotein (S-protein),2 which is believed to promote the invasion of host cells and the proliferation of the virus by binding to human angiotensin-converting enzyme 2 (ACE2) (Figure 1A). The SARS-CoV-2 S-protein monomer comprises nine domains/regions: N-terminal domain (NTD), receptor-binding domain (RBD), subdomain 1 (SD1), subdomain 2 (SD2), fusion peptide (FP), heptad repeat 1 (HR1), heptad repeat 2 (HR2), transmembrane region (TM), and intracellular domain (IC) (Figure 1B). The RBD domain plays a key role in cell infection via ACE2 recognition. Therefore, the SARS-CoV-2 S-protein is one of the target proteins used for antibody-drug design to prevent virus invasion, including S-protein neutralizing antibodies.3-5 To tackle the SARS-CoV-2 pandemic, researchers are extracting and designing neutralizing antibodies to block the binding of S-protein to ACE2.2-5 Understanding the binding mode of S-protein with ACE2 and existing neutralizing antibodies will help design potent neutralizing antibodies. Moreover, a better understanding of the molecular recognition differences of ACE2 between the SARS-CoV-2 and other SARS-CoV can improve anti-
body-drug design. Thus, certain X-ray crystallographic and electron microscopic analyzes have already been performed, and geometry-based interactions have been discussed in previous studies.  

Analyses of molecular recognition hot spots between the SARS-CoV-2 S-protein and ACE2 were conducted by several computational simulations, including electrostatic potential, molecular mechanics (MM)-based interaction energy analysis with classical force field, and quantum mechanics (QM)-based interaction energy analysis with evaluation of semiempirical dispersion energy. Molecular dynamic (MD) simulations for the complex formed by the S-protein and ACE2 were conducted to better understand their association. Approaches for drug-neutralizing antibody design targeting S-protein have been reported, such as a de novo design peptide inhibitors for SARS-CoV-2, a virtual screening of antiviral compounds for the SARS-CoV-2 and sequence-based epitope analysis for the SARS-CoV-2, among others. These previous studies have been based on the molecular recognition of S-protein and ACE2/antibody and were highly beneficial. These studies had focused on more on hydrogen bonds, which certainly play key roles in determining binding poses of a protein-protein interaction (PPI) system, than on hydrophobic interactions, which influence the binding affinity. Even when hydrophobic interactions were considered, they were limited by the non-quantitative or inaccurate evaluation using a structure-based, MM-based, or semi-empirical based analysis. Moreover, in the prediction of binding ability, the enthalpic interactions, such as electrostatic and dispersion interactions, as well as the effects of solvation and sugar chain have not been examined sufficiently. Therefore, we used ab initio electron-correlated QM theory that yields rigorous dispersion energy in order to quantitatively and accurately evaluate hydrophobic interactions such as CH/π interactions.

The QM calculation of whole protein system consisting of hundreds to thousands of residues can be conducted using the fragment molecular orbital (FMO) approach. The hydrogen bonds, electrostatic interactions, salt bridges, and hydrophobic interactions were quantitatively analyzed using the interfragmentary interaction energy (IFIE) and its energy decomposition analysis (PIEDA) based on FMO calculations. In recent years, the FMO method has been widely used as a drug discovery tool. To date, the FMO method has been successfully used for identifying key amino acid residues and sugar chains for PPI, such as molecular recognition for an influenza hemagglutinin of a fragment antigen-binding (Fab) antibody and a measles hemagglutinin of a signaling lymphocytic activation molecule. In addition, it is essential to combine desolvation energy and enthalpic binding energy to predict binding ability. Therefore, the present study aimed to address additional calculations that could shed light on the molecular process underlying the SARS-CoV-2 and SARS-CoV S-protein recognition by ACE2 or B38 Fab antibody, which remains poorly understood. In particular, the study explored the amino acid residues involved in, as well as types of interactions that make, hot spots for S-protein binding to ACE2. This FMO-based hot spot and epitope analysis can be applied to other PPI systems in drug-antigen design as well.

Figure 1. Human host cell infection via ACE2 recognition by SARS-CoV-2 S-protein. (A) ACE2 is the host cell receptor responsible for mediating the SARS-CoV-2 infection, the new coronavirus responsible for coronavirus disease 2019 (COVID-19). (B) Overall topology of the SARS-CoV-2 S-protein monomer that comprises an N-terminal domain (NTD), receptor-binding domain (RBD), subdomain 1 (SD1), subdomain 2 (SD2), fusion peptide (FP), heptad repeat 1 (HR1), heptad repeat 2 (HR2), transmembrane region (TM), and an intracellular domain (IC). A complex between the SARS-CoV-2 S-protein and ACE2 is shown in cyan and pink ribbon models.

2. COMPUTATIONAL METHODS

To clarify the important interactions between the SARS-CoV-2 S-protein and ACE2, interaction energy analysis was performed according to the procedures described below.

2.1 Preparation of the SARS-CoV-2 S-protein and ACE2 complex

We retrieved the crystal structures of the complexes formed by the SARS-CoV-2 S-protein with ACE2 and B38 Fab antibody from the PDB (PDB ID: 6LZG and 7BZ5, respectively) for the FMO calculation. To compare the epitope candidate of S-proteins related to the SARS-CoV-2, FMO calculations were also performed for complexes between SARS-CoV-2 chimeric and SARS-CoV S-proteins with ACE2 (PDB ID: 6VW1 and 2AJF, respectively). The structures of the different viral protein/ACE2 or B38 Fab antibody complexes are shown in Figure 2. All sugar chains and crystal water molecules were retained. Since
Cl⁻ and Zn²⁺ ions existed far away from the binding surface between the SARS-CoV-2 S-protein and ACE2, they were deleted to simplify the interaction analysis. The missing atoms and missing residues were complemented by structure preparation with the Molecular Operating Environment (MOE) graphical software package (Chemical Computing Group, Montreal, QC, Canada).²⁵ Hydrogen atoms were added to each complex using Protonate 3D with MOE, and sequences termini were capped with amine (-NH₂) and carboxylic acid (-COOH) groups. These molecular changes were optimized concurrently using the Amber:EHT force field with MOE.

2.2 FMO method and intermolecular interaction energy analysis

The ab initio FMO method used was as follows. Briefly, a large molecule or molecular cluster was divided into small fragments and the molecular orbital (MO) calculations for each fragmented monomer and dimer were performed to obtain the properties of the entire system. The many-body effects were considered through the environmental electrostatic potentials. The total energies of the FMO calculations were given by:

\[ E_{\text{total}} = \sum_{i} E'_i + \sum_{ij} \Delta E_{ij} \]  

where \( E'_i \) represents the monomer energy without the environmental electrostatic potential (\( \Delta E_{ij} \)) of the IFIE; and \( i \) and \( j \) are the fragment indices. In addition, PIEDA was used to decompose IFIE into its energy components \( \Delta E_{ij} \) including the electrostatic (ES), exchange (EX), dispersion interaction (DI), and charge transfer with others (CT+mix), according to:

\[ \Delta E_{ij} = \Delta E_{ij}^{\text{ES}} + \Delta E_{ij}^{\text{EX}} + \Delta E_{ij}^{\text{CT+mix}} + \Delta E_{ij}^{\text{DI}} \]  

By analyzing IFIE-sum over a target protein, one can extract amino acid residues that specifically interact with the target protein, on its binding protein. Moreover, summing over a set of fragments \( (A) \) as

\[ \Delta E_{ij}^{A} = \sum_{i \in A} \Delta E_{ij} \]  

equation 3

one can obtain the interaction energy between \( A \) and \( B \) sets which is used as the binding energy between two proteins in this study.

All FMO calculations were performed using the second-order Møller–Plesset perturbation theory (MP2) with 6-31G* basis set using the ABINIT-MP program.²⁷,²⁸ The MP2 electron-correlation energy corresponds to the DI term and can accurately and quantitatively evaluate CH/π and π–π interactions that cannot be sufficiently evaluated using the MM and semi-empirical methods. A Cholesky decomposition integral approximation²⁷ was applied to speed up the MP2 calculation while keeping the accuracy. The fragment unit for proteins was each amino acid residue and each sugar chain. During the standard fragmentation process it is desirable to cleave at a sp³ hybridized single bond; therefore, single amino acid residues were cleaved between Cα and C=O in the main chain rather than at the sp² hybridized amide bond. Interaction energy between amino acid residues was evaluated using IFIE analysis with BioStation Viewer.²⁵ The FMO calculation results of the PDB ID entries 6LZG, 6VV1, 2AJF, and 7BZ5 are registered

Figure 2. Complexes of several S-proteins with ACE2 or B38 Fab antibody. The complexed structure of ACE2 (pink ribbon; Chain A: seqs#19-615, 19-614, and 19-615 in (A), (B), and (C), respectively) with each S-protein is depicted using a ribbon model, and residues whose sequences differ between the three S-proteins are shown in space-filling (CPK) model. The S-protein of (A) SARS-CoV-2 (Chain B: seq#333-527), (B) SARS-CoV-2 chimera (Chain E: seq#334-527), and (C) SARS-CoV (Chain E: seq#323-502) are illustrated in the cyan, blue, yellow ribbon model (PDB IDs: 6LZG, 6VV1, and 2AJF), respectively. (D) A complex of B38 Fab antibody (orange ribbon, Chain h, seq#40-217; green ribbon, Chain L, seq#0-215) with SARS-CoV-2 S-protein (cyan ribbon, Chain A, seq#334-528) is shown. (E) In sequence alignment of the SARS-CoV-2, SARS-CoV-2 chimeric, and SARS-CoV S-proteins, the well-aligned and poorly aligned residues are shown in color gradation as blue and red, respectively, evaluated using the BLOSUM62 scoring matrix.⁵⁵ A receptor binding motif on the binding surface with ACE2 is shown by a red belt.

Considering the common sequence of the SARS-CoV-2 (C361–E516) and the SARS-CoV (C348–E502), 50 and 36 amino acid mutations were identified in the common sequence and the receptor binding motif, respectively (Figure 2E). Sequence homology between the SARS-CoV-2 and the SARS-CoV S-proteins was 84.4 % for the whole Pro337–Phe515 and 47.2 % for Tyr453–Tyr505. The sugar chain located on the SARS-CoV-2 chimeric/SARS-CoV S-protein and ACE2 binding surface was the only BMA sugar chain and was the terminal chain of the NAG-NAG-MGA sugar chain extending from the side chain of N90 on ACE2 (Figures 2B and 2C). These X-ray structures of ACE2 have a common sequence (S19–A614).
in the FMO database (https://drugdesign.riken.jp/FMODB/) as FMDB ID entries XXXX, XXXX, XXXX, and XXXX, respectively.

2.3 Geometric interaction fingerprint analysis

The hydrogen-bond, ion pair, and π-orbital interactions, such as XH/π, cation/π, and π–π interactions (Figure 3), were detected by geometric interaction fingerprint (IFP) analysis based on the detected interatomic distance and angle using the “prolig_Calculate” function of the MOE software. MM-based energy restrictions defined by the “prolig_Calculate” function were employed to extract all possible atom pairs.

![Figure 3](image)

Figure 3. Types of intermolecular and intramolecular interactions by geometric interaction fingerprint (IFP). Schematic diagrams of the (A) XH-Y hydrogen bond, (B) ion pair interaction, (C) XH-π interaction, (D) cation/π interaction, and (E) π-π interaction, respectively. Characteristic distances of each interaction were measured between detected heavy atoms or centroids of aromatic rings.

3. RESULTS AND DISCUSSION

3.1. How does the SARS-CoV-2 S-protein recognize ACE2 in human host cells?

In a first approach, a detailed structural analysis was performed to elucidate which key amino acid residues are important for the molecular recognition of ACE2 by the SARS-CoV-2-S-protein, as well as which types of interactions were involved. Such information could help identify epitope candidates for novel antibodies to prevent the binding between the S-protein and ACE2. FMO-based interaction energy and geometric IFP analyses of the S-protein and ACE2 were performed.

To determine key residues for molecular recognition, it is necessary to identify amino acid residue pairs with attractive interactions between proteins. IFIE and IFIE-sum analyses and their energy decomposition analyses with PIEDA are useful for quantitatively evaluating interaction energies such as hydrogen bonds and CH-π interactions. In 3.1, 3.2, and 3.3, IFIE-sums over S-protein and ACE2 were calculated using a common amino acid sequence aligned to compare among complexes because the X-ray crystal structures were composed of different lengths of amino acid sequences. Besides, the IFIE-sums did not include IFIES of sugar chains and water molecules. Figure 4 shows IFIE-sums of the SARS-CoV-2 S-protein (C361–E516) and ACE2 (S19–A614) using PIEDA. The data revealed that the ES components were dominant in their binding. Six acidic residues (D30, E35, E37, D38, E329, and D355) of ACE2 and five basic residues of the S-protein (R403, R408, K417, R457, and K458) showed a remarkable attractive interaction. These results were considered to be caused by total charges of molecular systems on ACE2 and S-protein of −26e and +3e, respectively. In the CT+mix component, attractive interactions were shown by 10 fragment residues of ACE2 (Q24, D30, K31, H34, D38, Y41, Q42, Y83, K353, and G354) and nine residues of the S-protein (K417, Y449, F456, F486, N487, Q493, T500, N501, and G502). In the case of the DI components, characteristic interactions were identified on 15 residue fragments of ACE2 (Q24, T27, F28, D39, K31, H34, E35, D38, Y41, Q42, M82, Y83, K353, G354, and D355) and 14 residue fragments of the S-protein (K417, Y449, Y453, K455, F456, G476, F487, Y489, Q493, Q498, T500, N501, G502, and Y505). Since these amino acid residues are considered to play key roles in the molecular recognition between ACE2 and the S-protein, the origin of the interaction energy of each fragment pair was further explored (Figure 5, section 3.1.1). Nonetheless, it is important to note that in FMO analysis, in general, the ES component is effective over a long distance and tends to be overestimated as compared with the CT and DI ones. Thus, in IFIE-sums with a charged fragment, the weak CT and DI components that show hydrogen bonding and CH-π interaction tend to be hidden by the ES one. Moreover, a direct comparison of the interaction strengths of the charged fragment pair and the neutral one can be challenging using IFIEs that include the ES components. Therefore, herein, the study focused on the amino acid residues contributing for attractive interactions in the CT and DI components caused by short-range interactions, such as hydrogen bonds and XH/π interactions, to reveal the molecular recognition mechanism on the interface between the SARS-CoV-2 S-protein and ACE2.

3.1.1 Hot spots of the SARS-CoV-2 S-protein for binding to ACE2

The interaction of the 15 amino acid residues of the SARS-CoV-2 S-protein by IFIE-sums believed to play a key role in ACE2 recognition is summarized in Figure 5. As supplementary information, the results of geometric IFP of hydrogen bond and XH/π interaction are summarized in Tables S1 and S2 with the corresponding IFIES.

- The T500spike formed two hydrogen bonds with ACE2 (Figure 5B and 5C): OH-O hydrogen bond between the hydrogen atom bounded OG1 oxygen atom of the side chain on T500spike and OH oxygen atom of the side chain on Y41ACE2, and the CH-O hydrogen bond of the hydrogen atom bounded CB carbon atom of side chain on T500spike and OD1 oxygen atom of side chain on Y41ACE2.
- The N501spike attractively interacted with Y41ACE2 (Figures 5C) and K353ACE2 (Figure 5D). The NH-O hydrogen bond formed between the hydrogen atom...
bound ND2 nitrogen atom of side chain on N501\textsubscript{Spike} and the OH oxygen atom of the side chain; NH/π interaction formed between the hydrogen atom bound ND2 nitrogen atom on side chain on N501\textsubscript{Spike} and the π-orbital of phenol ring on Y441\textsubscript{ACE2}. The hydrogen atoms bound CA and CB carbons on N501\textsubscript{Spike} interacted with O oxygen atom on the main chain of K353\textsubscript{ACE2}, which was contained within the G354\textsubscript{ACE2} fragment, because of CH-O hydrogen bonds.

- In the case of G502\textsubscript{Spike}, there was an NH-O hydrogen bond between the hydrogen atom bound N atom of the main chain on G502\textsubscript{Spike} and the O oxygen atom of the main chain on K353\textsubscript{ACE2}, which was contained within the G354\textsubscript{ACE2} fragment (Figure 5D).

- The Q498\textsubscript{Spike} was formed with an CH-O hydrogen bond between OE1 oxygen atom of the side chain on Q498\textsubscript{Spike} and the hydrogen atom bound CD2 carbon atom of the side chain on L45\textsubscript{ACE2} (Figure 5E). Moreover, attractive ES and DI energy components of Q498\textsubscript{Spike} with D38\textsubscript{ACE2} and Y41\textsubscript{ACE2} were interpreted as an NH-O hydrogen bond between the hydrogen atom bound NE2 nitrogen atom of the side chain on Q498\textsubscript{Spike} and the OD2 oxygen atom of the side chain on D38\textsubscript{ACE2}, and a CH/π interaction between the hydrogen atoms bound CD2 and CE2 carbon atoms on Y41\textsubscript{ACE2} and π-orbital in the amide group of the side chain on Q498\textsubscript{Spike}, respectively (Figure 5E).

- In the case of Y505\textsubscript{Spike}, there were CH-O hydrogen bonds and two CH/π interactions. Figures 5D and 5F show the CH-O hydrogen bond between the hydrogen atom bound CB carbon atom of the side chain on Y505\textsubscript{Spike} and the O oxygen atom of the main chain on K353\textsubscript{ACE2}, which was contained within the G354\textsubscript{ACE2} fragment, and OH-O hydrogen bonds between the hydrogen atoms bound OH oxygen atom of the phenol on Y505\textsubscript{Spike} and two oxygen atoms (OE1 and OE2) of side chain on E37\textsubscript{ACE2}, respectively. Also, in Figure 5F, CH/π interaction is presented in conformation between π-orbital in phenol ring on Y505\textsubscript{Spike}, and two the hydrogen atoms bound carbons atom (CA and CG) on K353\textsubscript{ACE2}.

- The Y449\textsubscript{Spike} attractively interacted with D38\textsubscript{ACE2} through OH-O and CH-O hydrogen bonds between two hydrogen atoms bound OH oxygen atom and CE2 carbon atom of phenol ring on Y449\textsubscript{Spike} with OD1 oxygen atom of side chain on D38\textsubscript{ACE2}, respectively (Figure 5G).

- The Y453\textsubscript{Spike} was conformationally stabilized by an OH-N hydrogen bond between the hydrogen atom bound OH oxygen atom of phenol ring on Y453\textsubscript{Spike} and ND1 nitrogen atom of imidazole ring on H34\textsubscript{ACE2} (Figure 5H); whereas, the OH/π interaction was formed between the hydrogen atom bound OH oxygen atom of phenol on Y453\textsubscript{Spike} and π-orbital in imidazole ring on H34\textsubscript{ACE2}.

- The molecular recognition of Q493\textsubscript{Spike} by ACE2 (Figure 5I) comprised NH-O and OH-N hydrogen bonds, OH/π interaction, and electrostatic interaction. Q493\textsubscript{Spike} formed an NH-O hydrogen bond between the hydrogen atom bound NE2 nitrogen atom of the side chain and the OE1 and OE2 oxygen atoms of the side chain on E35\textsubscript{ACE2}. The O oxygen atom of the main chain on H34\textsubscript{ACE2}, which was contained within the E35\textsubscript{Spike} fragment, attractively interacted because of the NH-O hydrogen bond with the hydrogen atom bound NE2 nitrogen atom of the side chain on Q493\textsubscript{Spike}. Although the absence of directed hydrogen bond and XH/π interaction, K31\textsubscript{ACE2} showed stronger attractive interaction with Q493\textsubscript{Spike} (−23.4 kcal/mol) than E35\textsubscript{ACE2} and H34\textsubscript{ACE2}. This interaction was primarily caused by the ES energy component (Figure 5I).

- The K417\textsubscript{Spike} attractively interacted with E30\textsubscript{ACE2} by a salt bridge between cation of NZ nitrogen atom and anions of OD1 and OD2 oxygen atoms on E30\textsubscript{ACE2} in the presence of two NH-O hydrogens bonds (Figure 5J). The IFIE between K417\textsubscript{Spike} and E30\textsubscript{ACE2} was the most attractive interaction (−116.35 kcal/mol; Table S1).

- The L455\textsubscript{Spike} attractively interacted with H34\textsubscript{ACE2} because of CH/π interaction between the hydrogen atoms bound CD1 and CD2 carbon atoms on the side chain of L455\textsubscript{Spike} and the π-orbital in imidazole ring on H34\textsubscript{ACE2} (Figure 5H).

- The F456\textsubscript{Spike} was structurally stabilized by a CH-O hydrogen bond between the hydrogen atom bound CZ carbon atom in the benzene ring on F456\textsubscript{Spike} and the O oxygen atom of the main chain on T27\textsubscript{ACE2}, which was contained within the F28\textsubscript{ACE2} fragment (Figure 5K). The F456\textsubscript{Spike} formed CH/π interaction between π-orbital in benzene ring on F456\textsubscript{Spike} and the hydrogen atom bound CG2 carbon atom of side chain on T27\textsubscript{ACE2} (Figure 5K).

- In the case of Y489\textsubscript{Spike} (Figure 5L), π-orbital in benzene ring on formed CH/π interactions with the hydrogen atoms bound CG2 and CB carbon atoms of the side chains on T27\textsubscript{ACE2} and K31\textsubscript{ACE2}, respectively. Moreover, the π-orbital in the benzene ring on Y489\textsubscript{Spike} showed potential to also interact with the hydrogen atom bound CA carbon atom of the main chain on F28\textsubscript{ACE2} in terms of the DI energy component and structural conformation (Figure 5L).

- The OD1 oxygen atom of the side chain on N487\textsubscript{Spike} makes the NH-O hydrogen bond with the hydrogen atom bound NE2 nitrogen atom of side chain on Q24\textsubscript{ACE2}, as well as the CH-O hydrogen bond with the hydrogen atom bound CG carbon atom of the side chain on Q24\textsubscript{ACE2} (Figure 5M and Table S1).

- F486\textsubscript{Spike} attractively interacted with L79\textsubscript{ACE2}, Y83\textsubscript{ACE2}, and M82\textsubscript{ACE2} due to a CH-O hydrogen bond and two CH/π interactions (Figure 5N). The CH-O hydrogen bond was formed between the hydrogen atom bound CE1 carbon atom of a benzene ring on F486\textsubscript{Spike} and the O oxygen atom of the main chain on L79\textsubscript{ACE2}, which was included within the A80\textsubscript{ACE2} fragment. The hydrogen atoms bound CE1 and CZ carbon atoms of the benzene ring on F486\textsubscript{Spike} also formed CH/π interactions with the π-orbital in a phenol ring on Y83\textsubscript{ACE2}, and the F486\textsubscript{Spike} was also stabilized by CH/π interaction between the π-orbital in the
benzene ring on F486\text{Spike} and the hydrogen atoms bounded CB carbon atom of side chain on M82\text{ACE2}.

- The DI energy component of G476\text{Spike} with Q24\text{ACE2} can be interpreted as indication that they formed CH/π interactions between the hydrogen atom bounded CA carbon atom of the main chain on G476\text{Spike} and the π-orbital in the amide group of the side chain on Q24\text{ACE2} (Figure 5O).

Next, the relationship between the energy intensity and the observed interaction was described based on the above the interaction analysis between the SARS-CoV-2 spike and ACE2.

**Figure 4.** Intermolecular interaction energies between the SARS-CoV-2 S-protein and ACE2 in host cells. The IFIE-sums over (A, B) ACE2 (S19–A614) and (C, D) S-protein (C361–E516) with each amino acid residue near the binding surface of the S-protein and ACE2 are shown. Visualization of IFIE-sums over the fragments of ACE2 (A) and SARS-CoV-2 S-protein (C) with attractive and repulsive interactions are represented by red and blue, respectively. Sugar chains are depicted by the ball and stick model. Figures (B) and (D) illustrate the IFIE-sums over ACE2 and SARS-CoV-2 S-protein, respectively, total, electrostatic (ES), exchange (EX), charge transfer with others (CT+mix), and dispersion interaction (DI) components at the amino acid residue. The residue names of the corresponding fragments with interaction energies of ES, CT+mix, and DI components lower than −30, −3, and −3 kcal/mol, respectively, are shown.
Figure 5. 3D visualization of the interaction energies between the SARS-CoV-2 S-protein amino acid residues near the contact surface with ACE2 (A). The fragments of T500Spike (B), Y41ACE2 (C), G354ACE2 (D), Q498Spike (E), Y505Spike (F), Y449Spike (G), H34ACE2 (H), Q493Spike (I), K417Spike (J), F456Spike (K), Y489Spike (L), N487Spike (M), F486Spike (N), and G476Spike (O) are shown in yellow. The main components of the attractive and repulsive interaction energies are represented by the following color scheme: ES component, red and blue; DI component, green and white. The SARS-CoV-2 S-protein and ACE2 are shown using stick and ball model and stick model, respectively. Hydrogen bonds, ion pair interactions, and XH/π interactions are shown by the cyan, red, and pink dotted lines.
### 3.1.2 XH-Y hydrogen bonds

The heavy atom pairs detected as hydrogen bonds in the geometric IFP and their FMO-based interaction energies are summarized in Table S1. It was confirmed that the fragments showing attractive interaction energies of ES and CT+mix terms less than −30 and −3 kcal/mol (Figures 4B and 4D), respectively, were mainly associated with NH-O and OH-O hydrogen bonds with a distance between heavy atoms within 3 Å. The hydrogen bonds between the O oxygen atom of K353ACE2, which was contained within the G354ACE2 fragment, and the hydrogen atom bonded N nitrogen atom of G502Spike (Figure 5D), the OD1 oxygen atom of D38ACE2 and the hydrogen atom bonded OH oxygen atom of Y449Spike (Figure 5G), the ND1 nitrogen atom of H34ACE2 and the hydrogen atom bonded OH oxygen atom of Y453Spike (Figure 5H), and the OD1 oxygen atom of N487Spike and the hydrogen atom bonded NE2 nitrogen atom of Q24ACE2 (Figure 5M) were typical examples if such attractive interactions. The interaction energy between D30ACE2 and K417Spike was the largest at −116.35 kcal/mol, where the distance between the heavy atoms of OD1 and OD2 oxygen atoms of D30ACE2 and NZ nitrogen atom of K417Spike (Figure 5J) were 2.863 and 3.761 Å, respectively. The atom pairs were also the only salt bridge pair between the S-protein and ACE2 that was detected from the geometric IFP. It was confirmed that these amino acid residues forming hydrogen bonds had attractive interaction energy lower than −5 kcal/mol, even in the ES component.

On the other hand, although weaker than the interaction energies of the NH-O and OH-O hydrogen bonds, CH-O hydrogen bonds also showed an attractive interaction energy that contributed to the molecular recognition of the SARS-CoV-2 S-protein and ACE2, such as the oxygen and carbon atom pairs for the O oxygen atom of T27ACE2, which was contained within the F28ACE2 fragment, and the hydrogen atom bonded CZ of F456Spike (Figure 5K), as well as the OD1 oxygen atom of N487Spike with the hydrogen atom bonded CG carbon atom of Gln24ACE2 (Figure 5M).

### 3.1.3 XH/π interactions

Besides the π-orbital of the aromatic ring, the heavy atom pairs detected as XH/π in the geometric IFP and their FMO-based interaction energies are summarized in Table S2. It was confirmed that the fragments showing attractive interaction energy of DI component lower than −3 kcal/mol (Figures 4B and 4D) were mainly associated with XH/π interactions with a distance between the centroid of the aromatic ring and heavy atom of X within 4 Å. For example, there were CH/π interactions via π and σCH orbitals on the imidazole ring of H34ACE2 and the hydrogen atoms bonded CD1 and CD2 carbon atoms of L455Spike (Figure 5H), the phenol ring of Y83ACE2 and the hydrogen atoms bonded CE1 and CZ carbon atoms of F486Spike (Figure 5N), and the benzene ring of F486Spike and the hydrogen atom bonded CB carbon atom of M82ACE2 (Figure 5N). Moreover, it was clarified that not only the CH of the hydrophobic amino acid residues but also CH/π interaction between the phenol ring of Y505Spike and the hydrogen atom bonded CA carbon atom of K353ACE2 (Figure 5F) contributed to the SARS-CoV-2 and ACE2 binding.

In addition to the CH/π interactions described, OH/π interaction between the imidazole ring of H34ACE2 and the hydrogen atom bonded OH oxygen atom of Y453Spike (Figure 5H), and the NH/π interaction between the phenol ring of Y41ACE2 and the hydrogen atom bonded ND2 nitrogen atom of N501Spike (Figure 5C), were also found. It was confirmed that interaction energies of DI component for H34ACE2 with Y453Spike and Y41ACE2 with N501Spike had stable interaction energies of −3.44 and −4.56 kcal/mol, respectively.

### 3.1.4 Ion pair, cation/π, and π-π interactions

Based on PIEDA and geometric IFP analyses, ion pair interaction was detected only on the NZ nitrogen atom of the side chain on K417Spike and between OD1 and OD2 oxygen atoms of the side chain on E30ACE2. In addition, there were no cation/π and π-π interactions between the S-protein and ACE2.

Taken together, the FMO data revealed that interaction networks were formed between the SARS-CoV-2 S-protein and ACE2 via hydrogen bond, XH/π, and salt bridge interactions spanning multiple residues. The 15 residues of ACE2 selected by CT and DI energy analyses play key roles in the recognition of the SARS-CoV-2 S-protein, representing hot spot residues for inhibiting the binding to ACE2 in the context of drug and antibody design. Although the XH-Y hydrogen bond and XH/π interactions were detected based on the geometric IFP analysis, this approach may fail to detect some molecular interactions. In particular, PIEDA can detect various types of XH/π interactions and also evaluate them quantitatively, whereas structure-based analysis detects only typical interactions. In fact, the CH/π interaction between the phenol ring of Y489Spike and the CA carbon atom of F28ACE2 (Figure 5E), which was not detected by geometric IFP analysis, was found by interaction energy of DI component with PIEDA. By the geometric IFP, the XH/π interactions with the π-orbital of the amide were not detected because this analysis method targeted the π-orbital of the aromatic ring. By PIEDA; however, it was confirmed that the XH/π interactions with the π-orbital of the amide in the side chain on Q498Spike and Q24ACE2 were also crucial for molecular recognition between the SARS-CoV-2 S-protein and ACE2 (Figures 5E and 5O).

The amino acid residue pairs that formed the NH-O and OH-O hydrogen bonds were in good agreement with similar interactions reported in previous studies. Herein, it was the first report describing that the formation of CH-O hydrogen bonds with weak interaction energy are important for molecular recognition of ACE2 by the SARS-CoV-2. In addition, although it was described in reports of X-ray crystallography and electron microscopic analyses that hydrophobic residues are involved in molecular recognition as an effect of Van der Waals interaction with ambiguous contributions, the findings here reported provide further clarification that such hydrophobic residues form XH/π interactions.
### 3.2. What are the differences and similarities in the ACE2 recognition mechanism among three SARS S-proteins?

To clarify the differences in the molecular recognition of ACE2 between the SARS-CoV-2 (Figure 4), the SARS-CoV-2 chimeric (Figure S1), and the SARS-CoV S-proteins (Figure S2), the FMO-based interaction energies of all three structures were analyzed and compared. Furthermore, the difference in the binding ability of ACE2 (S19–A614) between the SARS-CoV-2 (C348–E502), the SARS-CoV-2 chimeric (C348–E502), and the SARS-CoV S-proteins was also investigated by quantum mechanics (QM)-based interaction energy analysis (Figures 6 and 7). This approach revealed hot spot residues of ACE2 for designing low, medium, and peptide inhibitors, as well as epitope candidates of S-protein, to prevent S-protein and ACE2 binding.

#### 3.2.1 Hot spot analysis of ACE2

The three types of S-proteins were compared with each other for their ability to interact with ACE2, and key amino acid residues that were candidates for hot spot for inhibit-
for amino acid residues of the receptor binding motif were recorded over ACE2 (S19

The number of ‘+’ shows importance of hot spots between three complexes. *Count is number of hot spot residues.

3.2.2 Hot spot analysis of S-protein

Similar analyses of potential hot spots for ACE2 on the S-protein (Figure 7) revealed 34, 37, and 34 amino acid residues on the SARS-CoV-2, the SARS-CoV-2 chimeric, and the SARS-CoV S-proteins, respectively, that attractively interacted with ACE2. The SARS-CoV-2 and the SARS-CoV-2 chimeric S-proteins had similar interaction results as the amino acid sequences of the receptor binding motif were almost the same (Figure 2). In addition, 28 amino acid residues showed a common attractive interaction among the three S-proteins, being considered to be essential hot spots for the S-protein and ACE2 recognition regardless of the protein type. It was also found that, similar to the hot spot analysis of ACE2, the complex of the SARS-CoV-2 and the SARS-CoV-2 chimera had more amino acid residues that formed short-range interactions than the complex with the SARS-CoV. The 28 common hot spots including 17 conserved residues were highly homologous among the three S-proteins and can be designated as essential hot spots for recognizing ACE2 in S-proteins (SARS-CoV-2, SARS-CoV, and MARS virus).

As shown in figures 6 and 7, the SARS-CoV-2 S-protein had more hot spots that interacted with ACE2 by CT and DI interaction than the SARS-CoV-2 S-protein; in addition, the SARS-CoV-2 S-protein is considered to bind more firmly to ACE2 than the SARS-CoV-2 S-protein. This result is also consistent with previous findings. Moreover, the essential hot spots of the 15 common amino acid residues on ACE2...
and the 28 common residues on SARS S-proteins identified in the current study are partially in agreement with hot spots on ACE2 identified by Lim et al. via FMO calculations using the self-consistent charge density-functional tight-binding method with third-order expansion using semi-empirical dispersion (DFTB3/D) method. Lim et al. showed two sizable hot regions, where several hot spots were assembled. Their results were almost consistent with our results, showing that the common hot spots between ACE2 and S-protein were as follows: The first hot spot region included Q24ACE2, A25ACE2, F28ACE2, D30ACE2, K31ACE2, E35ACE2, and Y83ACE2, which interacted with K417Spike, L455Spike, S477Spike, E484Spike, F486Spike, N487Spike, Y489Spike, and Q493Spike in the SARS-CoV-2. The second hot spot region included E37ACE2, D38ACE2, Y41ACE2, Q42ACE2, E329ACE2, N330ACE2, K353ACE2, G354ACE2, and D355ACE2, which interacted with R403Spike, R439Spike, Y449Spike, F486Spike, G496Spike, F497Spike, Q498Spike, T500Spike, N501Spike, G502Spike, Y505Spike, and Q506Spike in the SARS-CoV-2. However, in this study, T27ACE2, H34ACE2, A80ACE2, and M82ACE2, V407Spike, R408Spike, Q409Spike, I418Spike, G446Spike, N448Spike, Y453Spike, F456Spike, R457Spike, K458Spike, L461Spike, Y473Spike, G476Spike, T478Spike, S494Spike, V503Spike, and G504Spike were newly identified as hot spots, for which the contributions of CT and DI were important for describing the hydrogen bond and XH/π interaction. We discovered new hot spots that were overlooked in their research for the following reasons: The DFTB3/D method with a partially semi-empirical approach tended to underestimate CT and DI components compared to the MP2 method with ab initio electron-correlated QM theory. In addition, there were several amino acid residues for which the tendency of attractive and repulsive interactions of CT and DI components was different between DFTB3/D and MP2 methods. Therefore, we succeeded in finding hot spots that were structurally and quantum chemically valid.

3.3. Is FMO-based epitope analysis of the SARS-CoV-2 S-protein useful?

Based on the complex between the SARS-CoV-2 S-protein and ACE2, the expected epitope candidates analyzed by IFIE-sum over B38 Fab antibody using PIEDA were further explored as potential neutralizing epitopes of S-protein. Using a complex of the SARS-CoV-2 S-protein (C361–E516) and B38 Fab antibody (D0–S217 on heavy chain and D0–C215 on light chain), PIEDA and geometric IFP analyses were performed similarly as described in the sections 3.1 and 3.2. The actual epitope residues of the B38 Fab antibody on the SARS-CoV-2 S-protein were revealed and compared to the candidate epitope residues upon ACE2 binding.

First, FMO-based interaction energy and geometric IFP analyses of the S-protein and B38 Fab antibody were performed to clarify the amino acid residues and their interaction type that were critical for molecular recognition. FMO-based interaction energy was analyzed by IFIE-sum (Figure S3). Next, the hydrogen bonds, XH/π interactions detected by geometric IFP, and IFIEs of the corresponding fragment pair are assessed (Tables S3 and S4). Based on the IFIE and the geometric IFP analyses, the interaction of 20 amino acid residues in the SARS-CoV-2 S-protein played a key role in B38 Fab antibody recognition.

3.3.1 XH-π hydrogen bonds

IFIEs between fragments that include heavy atom pairs detected as hydrogen bonds in the geometric IFP are summarized in Table S3. It was confirmed that the fragments showing attractive interaction energies of ES and CT+mix components lower than −30 and −3 kcal/mol (Figures S3B and S3D) were mainly associated with NH-O, OH-O hydrogen bonds with a distance between heavy atoms within 3 Å. The hydrogen bonds between the OG oxygen atom of S30E3B-H and NZ nitrogen atom of K458Spike, the OD2 oxygen atom of D429Spike and OG oxygen atom of S56E3B-H, and the O oxygen atom of the main chain on L455Spike (F456Spike fragment) and OH oxygen atom of Y33E3B-H were typical examples of such interaction. There was no interaction between the SARS-CoV-2 S-protein and B38 Fab antibody indicating a salt bridge by IFIE and geometric IFP analyses. Although weaker than the interaction energies of NH-O and OH-O hydrogen bonds, CH-O hydrogen bonds also showed an attractive interaction energy that contributed to the molecular recognition of the SARS-CoV-2 S-protein and B38 Fab antibody, such as oxygen and carbon atom pairs for the O oxygen atom of the main chain on L455Spike (F456Spike fragment) and CE2 carbon atom of Y33E3B-H and the OG oxygen atom of S30E3B-L and CG carbon atom of Q498Spike.

3.3.2 XH/π interactions

IFIEs between fragments that include heavy atom pairs detected as XH/π in the geometric IFP are summarized in Table S4. It was confirmed that the fragments showing attractive interaction energy by DI components lower than −3 kcal/mol (Figure S3B and S3D) were mainly associated with XH/π interactions with a distance between the centroid of the aromatic ring and heavy atom of X within 5 Å. For example, there were CH/π interactions via π and OCB orbitals on the benzene ring of F456Spike and the CE2 carbon atom of Y33E3B-H, the phenol ring of Y32E3B-L and the CD2 and CE2 carbon atoms of Y505Spike and the phenol ring of Y505Spike and the CA and CB carbon atoms of I29E3B-L. It was also clarified that not only the CH of the hydrophobic amino acid residues but also the CH/π interaction between the phenol ring of Y94E3B-L and the CB carbon atom of D405Spike contributed to the SARS-CoV-2 S-protein and B38 Fab antibody binding. Moreover, CH/π and NH/π interactions between the phenol ring of Y505Spike and NE2 nitrogen atom of Q90E3B-L were also identified.

3.3.3 FMO-based epitope analysis

The actual epitope residues of the B38 Fab antibody on the SARS-CoV-2 S-protein were revealed by IFIE-sums. We compared the actual epitope residues of the SARS-CoV-2 S-protein recognized by B38 Fab and the epitope candidate residues bounded to ACE2 (Figure 8). There were 21 common amino acid residues of the SARS-CoV-2 S-protein among the 31 epitope residues bound by B38 Fab antibody and the 34 hot spots bound by ACE2. On the other hand,
the amino acid residues of S-protein in which attractive interaction was observed with ACE2 but not with the B38 Fab antibody were V407, R408, I418, N439, G446, N448, Y449, Y453, R457, T478, Q493, F497, and T500. It was also revealed new amino acid residues through which the S-protein interacted with the B38 Fab antibody (D405, E406, T415, G416, D420, Y421, N460, Y473, A475, E484, and G496). Several amino acid residues of S-protein that had an attractive interaction with ACE2 did not interact with the B38 Fab antibody. Nevertheless, the B38 Fab antibody showed more CT and DI interactions with the S-protein amino acid residues that did not interact with ACE2.

| Residue | ES (kcal/mol) | EX (kcal/mol) | CT (kcal/mol) | DI (kcal/mol) | *Hot spot |
|---------|---------------|---------------|---------------|---------------|-----------|
| ARG403  | -265.3        | 0.0           | 0.0           | 0.0           | +         |
| ASP405  | -254.8        | 0.0           | 0.0           | 0.0           | +         |
| GLU406  | -234.6        | 0.0           | 0.0           | 0.0           | +         |
| VAL407  | -6.8          | 0.0           | 0.0           | 0.0           | +         |
| ARG408  | -22.6         | 0.0           | 0.0           | 0.0           | +         |
| GLN409  | -6.9          | 0.0           | 0.0           | 0.0           | +         |
| THR415  | -1.1          | 0.0           | 0.0           | 0.0           | +         |
| GLY416  | -3.6          | 0.0           | 0.0           | 0.0           | +         |
| LYS417  | -3.4          | 5.1           | -4.5          | -4.2          | +         |
| LEU418  | -6.1          | 0.0           | 0.0           | 0.0           | +         |
| ASP420  | 197.5         | 0.0           | 0.0           | 0.0           | +         |
| TYR421  | -1.5          | 0.0           | 0.0           | 0.0           | +         |
| ASN439  | -5.0          | 0.0           | 0.0           | 0.0           | +         |
| VAL445  | -6.2          | 0.0           | 0.0           | 0.0           | +         |
| GLY446  | -7.8          | 0.0           | 0.0           | 0.0           | +         |
| GLY447  | 2.4           | 1.4           | -1.3          | -1.5          | +         |
| GLN448  | -11.8         | 0.0           | 0.0           | 0.0           | +         |
| TYR449  | -31.8         | 12.1          | -7.4          | -6.9          | +         |
| TYR453  | -15.0         | 3.7           | -2.6          | -3.5          | +         |
| LEU455  | -0.5          | 3.9           | 2.9           | -6.0          | +         |
| PHE456  | 7.1           | 6.1           | -3.4          | -8.4          | +         |
| ARG457  | 18.9          | 0.0           | 0.0           | 0.0           | +         |
| LYS458  | 10.3          | 0.0           | 0.0           | 0.0           | +         |
| SER459  | -0.3          | 0.0           | 0.0           | 0.0           | +         |
| ASN460  | 5.1           | 0.0           | 0.0           | 0.0           | +         |
| LEU461  | -5.2          | 0.0           | 0.0           | 0.0           | +         |
| TYR473  | -3.8          | 0.0           | 0.0           | 0.0           | +         |
| ALA475  | 0.2           | -0.4          | -1.5          | +           |
| GLY476  | 7.4           | 2.5           | -2.1          | -3.2          | +         |
| SER477  | -3.8          | 0.0           | 0.0           | 0.0           | +         |
| THR478  | -5.6          | 0.0           | 0.0           | 0.0           | +         |
| CYS480  | -2.7          | 0.0           | 0.0           | 0.0           | +         |
| GLU484  | 19.0          | 0.0           | 0.0           | 0.0           | +         |
| GLY485  | -0.4          | 0.0           | 0.0           | 0.0           | +         |
| PHE486  | 0.4           | 6.6           | 3.5           | -10.6         | +         |
| ASN487  | -20.1         | 0.0           | 0.0           | 0.0           | +         |
| TYR489  | 0.9           | 5.3           | 2.5           | -9.4          | +         |
| PHE490  | 1.7           | 0.0           | 0.0           | -0.6          | +         |
| PRO491  | -1.0          | 0.0           | 0.0           | -0.4          | +         |
| GLN493  | -27.5         | 3.1           | -2.7          | -6.3          | +         |
| SER494  | -4.0          | 0.0           | 0.0           | 0.0           | +         |
| TYR495  | 8.4           | 0.0           | 0.0           | -0.1          | +         |
| GLY496  | 0.9           | 2.0           | -2.8          | +           |
| PHE497  | -9.5          | 1.3           | -1.3          | -2.4          | +         |
| GLU498  | -4.9          | 3.2           | -2.6          | -6.5          | +         |
| PRO499  | -1.4          | 0.3           | 0.0           | 0.0           | +         |
| THR500  | -19.9         | 15.9          | -8.3          | -9.1          | +         |
| ASN501  | -20.4         | 5.6           | -6.2          | -10.8         | +         |
| GLY502  | -21.0         | 13.9          | -5.9          | -5.7          | +         |
| VAL503  | 10.5          | 0.6           | -0.1          | -1.2          | +         |
| GLY504  | -9.4          | 0.0           | 0.0           | 0.0           | +         |
| TYR505  | -15.4         | 8.4           | -2.7          | -10.8         | +         |
| GLN506  | -18.4         | 0.0           | 0.0           | -0.2          | +         |

Figure 8. Difference of the IFIE-sums over the SARS-CoV-2 S-protein (C361–E516) between ACE2 and B38 Fab antibody. Interaction of each amino acid residue on S-protein with ACE2 and B38 Fab antibody are listed. Attractive interaction energies of the electrostatic (ES), charge transfer (CT), and dispersion interaction (DI) are identified using red, light blue, and green gradations, respectively, and repulsive interaction energy of the exchange repulsion (EX) is identified using pink gradation. *Hot spots and epitopes with an interaction energy of −3 kcal/mol or less in any of the ES, EX, CT, and DI components, were labeled as "*". **The number of "*" shows importance for hot spots and epitopes. Count is number of hot spot and epitope residues.

The numbers of amino acid residues in the SARS-CoV-2 S-protein showing CT and DI interaction were 9 and 15, respectively, when complexed with ACE2; whereas it was 14 and 20, respectively, when the S-protein was complexed with the B38 Fab antibody. This indicated that B38 Fab acquired more hydrogen bonds and XH/n interactions than ACE2. From these results, the 21 amino acid residues of the SARS-CoV-2 S-protein shared by the ACE2 and B38 Fab antibody interactions would be essential epitopes that directly inhibit ACE2 recognition and binding. This data (Figure 8) is expected to provide useful information for the development of potentially therapeutic antibodies.
example, for residues without attractive interaction with the antibody, the substitution of amino acid residues that mimic the interaction with ACE2 will lead to the design of antibodies with higher binding capacity. These amino acid residues were also consistent with the description of NH-O and OH-O hydrogen bonds reported by a previous study. New NH-O and OH-O hydrogen bonds, not shown by the previous study, were identified and their interaction was clarified by FMO calculations. This study demonstrated that CH-O hydrogen bonds and XH/π interactions are crucial for the molecular recognition between the SARS-CoV-2 S-protein and B38 Fab antibody, with 21 epitope residues of the SARS-CoV-2 S-protein being critical for antibody design.

The difference in the interaction energy of each amino acid residue of the S-protein between ACE2 and B38 Fab antibody will provide worthwhile information for improving B38 Fab and other antibodies. Lim et al. has shown two large hot spot regions from FMO-based interaction analysis of ACE2 and S-proteins of the SARS-CoV-2. Since the hot spots of S-protein with the several neutralizing antibodies of the SARS-CoV coincided with the 2nd hot spot region, the authors speculated that the 2nd hot spot region was crucial for drug design against the SARS-CoV-2. However, our results suggest that all key hot spots for molecular recognition between the SARS-CoV-2 and S protein and ACE2 may play a significant role in the design of neutralizing antibodies. This is because the first and the second sizable hot spot regions of S protein, defined by Lim et al., were both found as epitopes of the complex between S protein and B38 Fab antibody.

As Figure 8 shows, in ACE2 binding, we confirmed that the primary contribution of local hot spots was long-distance electrostatic interaction, such as that of R403 and K417; however, in B38 Fab binding, the local hot spots (e.g., R403 and K417) and their surrounding amino acid residues acquired a more complex and robust interaction network by short-distance interaction such as hydrogen bond and XH/π interactions. We also confirmed that short-range interactions, such as hydrogen bond and XH/π interactions, form a complex interaction network. That is, the critical amino acid residues in ACE2 binding determined, using FMO-based interaction analysis as hot spots, have the potential to become epitopes of the antibody; hence, Figures 6, 7, and 8 may be useful as drug-antibody design guidelines.

### 3.4. Is the binding potential predictable by FMO-based binding energy?

The ability of ACE2 to bind to the three S-proteins has been previously reported as $K_d$ values (SARS-CoV-2: 44.2 nM, SARS-CoV chimera: 23.2 nM, SARS-CoV: 185 nM). It can also be assumed that B38 Fab antibody binds to the SARS-CoV-2 S-protein more strongly than ACE2 because the B38 Fab antibody has been reported to strongly inhibit the SARS-CoV-2 S-protein. Therefore, these binding abilities were further examined in light of the FMO calculation results. We evaluated the binding energy between the SARS-CoV-2 S-protein and ACE2/B38 Fab antibody with and without sugar moieties.

### 3.4.1 FMO-based binding energies

The predicted binding energies using summation of the IFIEs (total, ES, EX, CT, and DI) between S-proteins and ACE2/B38 antibody with sugar moieties are listed in Table 1. The total binding energy using IFIE approach was evaluated by:

$$\Delta E_{\text{Bind}} = \sum \Delta E_{ij}^{\text{ES}} + \Delta E_{ij}^{\text{EX}} + \Delta E_{ij}^{\text{CT+mix}} + \Delta E_{ij}^{\text{DI}}$$

where $I$ is the fragment of the S-protein and $J$ is the fragment of the ACE2/B38 Fab antibody.

### Table 1: Predicted binding energies between the S-protein and ACE2/B38 Fab antibody.

| Bonded protein | CoV-2 | CoV-2 chimera | CoV | CoV-2 B38 Fab antibody |
|----------------|-------|---------------|----|------------------------|
| $K_d$ (nM)     | 44.2  | 23.2          | 185| -                      |
| p$K_d$         | 7.35  | 7.63          | 6.73| -                      |
| ES             | -887.15 | -645.09    | -818.50 | -432.38        |
| EX             | 122.71 | 121.73      | 110.62 | 269.35         |
| CT+mix         | -70.93 | -67.92      | -56.88 | -130.86        |
| DI             | -125.36 | -123.48    | -105.42 | -179.50        |
| Total          | -960.73 | -714.75    | -870.19 | -473.38        |
| SC-ES          | -612.88 | -491.07    | -549.43 | -498.64        |
| Total (SC)     | -686.46 | -560.74    | -601.11 | -539.65        |
| Desolv         | 592.53 | 395.91      | 531.82 | 98.78          |
| Total (Desolv) | -368.20 | -318.84    | -338.36 | -374.61        |
| Total (SC+Desolv) | -93.92 | -164.83   | -69.29 | -440.87        |

**Abbreviations:** CoV, coronavirus; CT+mix, charge transfer with others; Desolv, desolvation; DI, dispersion interaction; ES, electrostatic; EX, exchange; $K_d$, binding dissociation constant; SC, statistically correction; SC-ES, statistically corrected electrostatic.

In the case of the three S-proteins with ACE2, the main component of the binding energies was electrostatic interaction energy. This is probably due to the positive charge (SARS-CoV-2: $+2e$, SARS-CoV: $+1e$) of the viral S-proteins and the highly negative charge of ACE2 ($-26e$). The IFIEs of total and ES components did not correlate with p$K_d$ values. On the other hand, quantum chemical short-range interactions, such as the CT and DI energies, were aligned in the same magnitudes as p$K_d$ for the order of p$K_d$ for the SARS-CoV-2 and the SARS-CoV S-proteins were regarded as almost the same (7.35 and 7.63, respectively).

In the case of the SARS-CoV-2 S-protein with B38 Fab antibody, the main component of the binding energies was electrostatic interaction energy. The binding energies of the SARS-CoV-2 S-protein with B38 Fab antibody at the
total and ES components were weaker than those for the three S-proteins with ACE2, because the B38 Fab antibody has a positive charge (+6e). In contrast, the binding energies for the DI and CT components were the strongest of the four complexes. As described in sections 3.1–3.3, it is considered that contributions of the CT+mix and the DI components came from the XH-Y hydrogen binding involving charge transfer, the XH/r interaction at the contact surface, and orbital interaction of the aromatic rings. The overall CT and DI energies for the four complexes seem to comply with the strength of the binding affinity; however, the total interaction energy including ES does not work well. Therefore, the solvation effect was next considered for evaluating the binding energy.

### 3.4.2 Solvation effect

To account for the solvation effect, the statistically corrected IFIE (SCIFIE) and FMO method combined with molecular mechanics Poisson–Boltzmann surface area (FMO+MM-PBSA) approaches were used for predicting the binding energy (Table 1). The SCIFIE approach is a method to consider electrostatic and solvation shielding effects. In this research, a statistically corrected ES interaction energy (ΔESC-ES), which felt the shielding effect by both the solvent and the inside of the molecule, was used to analyze the ES component. Then the modified IFIE with the SC-ES component is given by

\[
\Delta E_{\text{IFIE}} = \Delta E_{\text{IFIE}} - \Delta E_{\text{SC-ES}} + \Delta E_{\text{ES}} + \Delta E_{\text{CT+mix}} + \Delta E_{\text{DI}}
\]

(6)

and the total binding energy using SCIFIE approach was evaluated by:

\[
\Delta E_{\text{Bind}} = \sum_{i,j} \Delta E_{ij}^{\text{ES}}
\]

(7)

where \(I\) is the fragment of the S-protein and \(J\) is the fragment of the ACE2/B38 Fab antibody. On the other hand, the FMO+MM-PBSA approach is a method used to evaluate binding energy incorporating the solvation effect by combining the desolvation energy with the sum of IFIEs as follows:

\[
\Delta G_{\text{Desolv}} = G_{\text{Solv}} - G_{\text{Solv}} - G_{\text{Bolv}}
\]

(8)

where Δ\(G_{\text{Desolv}}\) is desolvation energy, and \(G_{\text{Solv}}\) is solvation energy with \(A\) representing the S-protein and \(B\) the ACE2/B38 Fab antibody. The total binding energy was predicted using FMO+MM-PBSA approach as follows:

\[
\Delta E_{\text{Bind}} = \sum_{i,j} \Delta E_{ij} + \Delta G_{\text{Desolv}}
\]

(9)

Lastly, the total binding energy was predicted using SCIFIE combined with the desolvation evaluated using MM-PBSA method as follows:

\[
\Delta E_{\text{Bind}} = \sum_{i,j} \Delta E_{ij}^{\text{SC}} + \Delta G_{\text{Desolv}}
\]

(10)

These total binding energies are shown in Table 1. While \(\Delta E_{ij}^{\text{SC-ES}}\) gives a weak contribution by shielding effect, the binding energy alone using the SCIFIE approach for the four complexes was not in agreement with the experimental binding affinity. By adding the desolvation energies from the MM-PBSA approach, the excessive electrostatic interaction energies were suppressed, and the predicted binding energies approximate to reproduce the experimental binding ability under the physiological condition. Moreover, the binding energy including both statistical correction and desolvation (SC+Desolv) for the three complexes, excluding the SARS-CoV-2 chimeric S-protein and ACE2 complex, seemed to be in agreement with the order of the binding affinity. The large difference seen in the interaction energy due to the electric charge of the molecules was eliminated by adding shielding and the desolvation effects, and the binding energies became comparable to that of the experimental value. Thus, when the PPI binding ability is estimated by FMO calculation among differently charged molecular systems, it may be useful to incorporate both the shielding and desolvation effects.

### 3.4.3 Sugar chain effect

Lastly, the role of sugar chains in the molecular recognition of S-protein and ACE2/B38 Fab antibody was investigated. The sugar chain located on the SARS-CoV-2/SARS-CoV S-protein and ACE2 binding surface was the only BMA sugar chain, and was the terminal sugar chain of the NAG-NAG-MGA sugar chain extending from the side chain of N90 on ACE2 (Figures 2B and 2C). In all complexes between the SARS-CoV-2 S-protein and ACE2 (PDB IDs: 6LZG, 6M0L, and 6M17) a sugar chain was not observed, in which a sugar chain consisting of three sugars from N90 of ACE2 was likely to reach and interact with the S-protein. There was one sugar chain (NAG) from N90 of ACE2, which did not directly interact with the SARS-CoV-2 S-protein (PDB ID: 6LZG). Moreover, no sugar chains were identified on the binding surface between the SARS-CoV-2 S-protein and B38 Fab antibody.

Table S5 lists the predicted binding energies using the IFIEs between the S-protein and ACE2/B38 Fab antibody without interaction energy of the sugar chain fragment. By comparing Tables 1 and S5, the predicted binding energies of IFIEs between the SARS-CoV-2 chimeric S-protein and ACE2 was stabilized at −15.5 kcal/mol by the sugar chains, whereas that of the SARS-CoV S-protein and ACE2 was stabilized at −2.8 kcal/mol. In the SARS-CoV-2 chimeric S-protein/ACE2 complex, the sugar chain on ACE2 attractively interacted with R408 via hydrogen bond and was critical for the molecular recognition of the S-protein. Moreover, the IFIEs between the BMA703 sugar chain on ACE2 and R408 of the SARS-CoV-2 S-protein showed an attractive interaction of −16.2 kcal/mol (ES: −14.4 kcal/mol, EX: 1.5 kcal/mol, CT+mix: −1.3 kcal/mol, DI: −2.0 kcal/mol), which accounts for most of the summation of IFIEs in sugar chain effects. In the complex between the SARS-CoV S-protein and ACE2, weak attractive interaction between the BMA1092 sugar chain on ACE2 with the S-protein was confirmed, representing a repulsive interaction energy of +3.0 kcal/mol; however, several residues (e.g., D392, D393, and T402) of the S-protein interacted with the BMA1092 sugar chain with IFIEs of −1.0 kcal/mol. In this study, only three sugar chains from N90 on ACE2 played key roles in the molecular recognition between the SARS-CoV-2 chimeric S-protein and ACE2. Since sugar chains are generally present on the molecular surface, the electron density tends to be difficult to see in the experimental setting. Thus, it will be necessary to discuss not
only one X-ray crystal structure but also other X-ray crys-
tal structures and structural fluctuations by MD simula-
tions in the future.14

Taken together, it remains challenging to predict the
strength of the binding ability of the differently charged
proteins, such as ACE2, having highly negative charges,
and of the B38 Fab antibody, having highly positive charg-
es with a receptor such as S-protein using bare IFIE. How-
ever, incorporating desolvation energy and SCIPIE demon-
strates that the predicted binding energies between differ-
ently charged proteins could be improved for suppressing
the overestimated ES interactions. Meanwhile, ACE2 com-
prises large amount of Asp, Glu, and His that can have mul-
tiple protonated states. Therefore, while the calculation
was performed using only one protonation state in this
study, it will be necessary to examine multiple plausible
protonation states that may exist in vivo and investigate
the predicted binding energy of PPI.

4. CONCLUSION

The amino acid residues that are the key to the molecu-
lar recognition of three SARS S-proteins and ACE2/B38
Fab antibody were revealed by FMO-based interaction
energy analysis with ab initio electron-correlated theory.
The collected data provided new insights in the impor-
tance of forming a complex interaction network for the
molecular recognition between the S-protein and
ACE2/B38 Fab antibody, not only via NH-O and OH-O hy-
drogen bonds and salt bridges, but also via the CH-O hy-
drogen bonds and XH/π interactions. Since the XH/π in-
teraction is specifically found by the interaction energy of
DI components, it is difficult to accurately understand
XH/π interaction solely through MM-based electrostatic
interaction analysis and structure-based geometry analysis.
Moreover, QM-based hot spot and epitope analyses by
FMO calculations were useful in clarifying the type and
strength of molecular interactions, such as hydrogen bond
and XH/π interactions. Prediction of the binding ability
between the three types of SARS S-proteins and
ACE2/antibody was performed by FMO-based interaction
energy where incorporation of the shielding and desolva-
tion effects was an essential factor. Since sugar chains are
generally present on the molecular surface and have a dis-
ordered structure, the role of sugar chains in molecular
recognition between the S-protein and ACE2 should be ex-
amined using structural fluctuation sampling, such as
MD simulation, in future investigations.

We plan to release all the FMO data in this study on the
public database, FMODDB59,60, so that all researchers can
access it and utilize it for designing effective antibody-
drugs. Also, our group has recently performed over 336
FMO calculations for the COVID-19-related proteins such
as S-protein, main pro tease, RNA dependent RNA55, based
on the representative PDB structures selected in the PDB
Japan database in line with this global fight effort against
the coronavirus epidemic. These data have already been
published on FMODDB and can easily be analyzed IFIEs for
inter- and intramolecular interactions on the Web inter-
fase. In addition, several studies have also provided de-
tailed information on the underlying antigen-antibody or
ligand-binding characteristics of several drug-targeted
proteins in COVID-19.13,66-68 Finally, we expect that these
findings of novel hot spots/epitopes between the SARS-
CoV-2 S-protein and ACE2/B38 Fab antibody will provide
useful information for future antibody design, and small or
medium drug design that overcome COVID-19.

ASSOCIATED CONTENT

Supporting Information

Table S1. IFIE and geometric IFP of X-H hydrogen bonds
between the SARS-CoV-2 S-protein and ACE2.
Table S2. IFIE and geometric IFP of X-H/π interactions
between the SARS-CoV-2 S-protein and ACE2.
Table S3. IFIE and geometric IFP of X-H-Y hydrogen bonds
between the SARS-CoV-2 S-protein and B38 Fab antibody.
Table S4. IFIE and geometric IFP of X-H/π interactions
between the SARS-CoV-2 S-protein and B38 Fab antibody.
Table S5. Summary of the predicted binding energies using
IFIEs without IFIEs of sugar chain fragments.

Figure S1. IFIE-sums over ACE2 and the SARS-CoV-2 chimeric
S-protein.
Figure S2. IFIE-sums over ACE2 and the SARS-CoV S-protein.
Figure S3. IFIE-sums over B38 FAB antibody and the SARS-
CoV-2 S-protein.

This material is available free of charge via the Internet at
http://XXXXXXXXX.

AUTHOR INFORMATION

Corresponding Author

*Chiduru Watanabe, Center for Biosystems Dynamics Re-
search, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
230-0045, Japan. Tel: +81-45-503-9551; Fax: +81-
45-503-9432, Email: chiduru.watanabe@riken.jp.

Author Contributions

The manuscript was written by CW. The FMO calculations
were performed and analyzed by CW. All authors discussed
the results and have given approval to the final version of the
manuscript.

Funding Sources

This study was partially supported by the Platform Project for
Supporting Drug Discovery and Life Science Research (Basis
for Supporting Innovative Drug Discovery and Life Science
Research (BINDS) from the Japan Agency for Medical Re-
search and Development (AMED) (grant number
JP20am01011113). CW was the receiver of a JST PRESTO grant
(IPMIPR18GD).

Acknowledgements

This research was performed in the activities of the FMO drug
design consortium (FMODD). The FMO calculations were par-
tially performed using the HOKUSAII supercomputer (RIKEN
Advanced Center for Computing and Communications, Yokoh-
ama, Japan) and the TSUBAME3.0 supercomputer (Tokyo
Institute of Technology, Tokyo, Japan). PIEDA calculations
were performed with the MIZUHO/BioStation software pack-
age (version 4.0). We also thank Editage (www.editage.jp) for
English language editing.

ABBREVIATIONS

Supporting Information

Table S1. IFIE and geometric IFP of X-H-Y hydrogen bonds
between the SARS-CoV-2 S-protein and ACE2.
Table S2. IFIE and geometric IFP of X-H/π interactions
between the SARS-CoV-2 S-protein and ACE2.
Table S3. IFIE and geometric IFP of X-H-Y hydrogen bonds
between the SARS-CoV-2 S-protein and B38 Fab antibody.
Table S4. IFIE and geometric IFP of X-H/π interactions
between the SARS-CoV-2 S-protein and B38 Fab antibody.
Table S5. Summary of the predicted binding energies using
IFIEs without IFIEs of sugar chain fragments.

Figure S1. IFIE-sums over ACE2 and the SARS-CoV-2 chimeric
S-protein.
Figure S2. IFIE-sums over ACE2 and the SARS-CoV S-protein.
Figure S3. IFIE-sums over B38 FAB antibody and the SARS-
CoV-2 S-protein.

This material is available free of charge via the Internet at
http://XXXXXXXXX.

AUTHOR INFORMATION

Corresponding Author

*Chiduru Watanabe, Center for Biosystems Dynamics Re-
search, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
230-0045, Japan. Tel: +81-45-503-9551; Fax: +81-
45-503-9432, Email: chiduru.watanabe@riken.jp.

Author Contributions

The manuscript was written by CW. The FMO calculations
were performed and analyzed by CW. All authors discussed
the results and have given approval to the final version of the
manuscript.

Funding Sources

This study was partially supported by the Platform Project for
Supporting Drug Discovery and Life Science Research (Basis
for Supporting Innovative Drug Discovery and Life Science
Research (BINDS) from the Japan Agency for Medical Re-
search and Development (AMED) (grant number
JP20am01011113). CW was the receiver of a JST PRESTO grant
(IPMIPR18GD).

Acknowledgements

This research was performed in the activities of the FMO drug
design consortium (FMODD). The FMO calculations were par-
tially performed using the HOKUSAII supercomputer (RIKEN
Advanced Center for Computing and Communications, Yokoh-
ama, Japan) and the TSUBAME3.0 supercomputer (Tokyo
Institute of Technology, Tokyo, Japan). PIEDA calculations
were performed with the MIZUHO/BioStation software pack-
age (version 4.0). We also thank Editage (www.editage.jp) for
English language editing.

ABBREVIATIONS

http://XXXXXXXXX.

AUTHOR INFORMATION

Corresponding Author

*Chiduru Watanabe, Center for Biosystems Dynamics Re-
search, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
230-0045, Japan. Tel: +81-45-503-9551; Fax: +81-
45-503-9432, Email: chiduru.watanabe@riken.jp.

Author Contributions

The manuscript was written by CW. The FMO calculations
were performed and analyzed by CW. All authors discussed
the results and have given approval to the final version of the
manuscript.

Funding Sources

This study was partially supported by the Platform Project for
Supporting Drug Discovery and Life Science Research (Basis
for Supporting Innovative Drug Discovery and Life Science
Research (BINDS) from the Japan Agency for Medical Re-
search and Development (AMED) (grant number
JP20am01011113). CW was the receiver of a JST PRESTO grant
(IPMIPR18GD).

Acknowledgements

This research was performed in the activities of the FMO drug
design consortium (FMODD). The FMO calculations were par-
tially performed using the HOKUSAII supercomputer (RIKEN
Advanced Center for Computing and Communications, Yokoh-
ama, Japan) and the TSUBAME3.0 supercomputer (Tokyo
Institute of Technology, Tokyo, Japan). PIEDA calculations
were performed with the MIZUHO/BioStation software pack-
age (version 4.0). We also thank Editage (www.editage.jp) for
English language editing.

ABBREVIATIONS

http://XXXXXXXXX.
COVID-19, coronavirus disease 2019; ACE2, angiotensin-converting enzyme 2; CPK, space-filling; Dsolv, desolvation; DFTB3/D, self-consistent charge density-functional tight-binding method with the third-order expansion using semi-empirical dispersion; FMO, fragment molecular orbital; FMODB, FMO database; FP, fusion peptide; HR1, heptad repeat 1; HR2, heptad repeat 2; IC, intracellular domain; IFIE, intermediate fragment interaction energy; IFP, interaction fingerprint; Kd, binding dissociation constant; MD, Molecular dynamics; MM, molecular mechanics; MM-PBSA, molecular mechanics Poisson–Boltzmann surface area; MO, molecular orbital; MP2, second order Møller–Plesset perturbation theory; NTD, N-terminal domain; PDB, Protein Data Bank; PIEDA, Pair interaction energy decomposition analysis; PPI, protein–protein interaction; QM, quantum mechanics; RBD, receptor-binding domain; S-protein, spike glycoprotein; SARS-CoV-2, acute respiratory syndrome coronavirus 2; SC, statistically correction; SC-ES, statistically corrected electrostatic; SCIFIE, statistically corrected IFIE; SD2, subdomain 2; SD2, subdomain 1; TM, transmembrane region.

REFERENCES

(1) Coronavirus disease (COVID-19) – World Health Organization. https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed Oct 2, 2020).

(2) 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 (2), 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058.

(3) Wu, Y.; Wang, F.; Shen, C.; Peng, W.; Li, D.; Zhao, C.; Li, Z.; Li, S.; Bi, Y.; Yang, Y.; Gong, Y.; Xiao, H.; Fan, Z.; Tan, S.; Wu, G.; Tan, W.; Lu, X.; Fan, C.; Wang, Q.; Liu, Y.; Zhang, C.; Qi, J.; Gao, G. F.; Gao, F.; Liu, L. A Noncompeting Pair of Human Neutralizing Antibodies Block COVID-19 Virus Binding to Its Receptor ACE2. 2020, 6.

(4) Huo, J.; Le Bas, A.; Ruza, R. R.; Duyvesteyn, H. M. E.; Mikolajek, H.; Malinuskas, T.; Tan, T. K.; Rijal, P.; Dumoux, M.; Ward, P. N.; Ren, J.; Zhou, D.; Harrison, P. J.; Weckener, M.; Clare, D. K.; Vogirala, V. K.; Radecke, J.; Moynié, L.; Zhao, Y.; Gilbert-Jaramillo, J.; Knight, M. L.; Tree, J. A.; Buttigieg, K. R.; Coombes, N.; Elmore, M. J.; Carroll, M. W.; Carrique, L.; Shah, P. N. M.; James, W.; Townsend, A. R.; Stuart, D. I.; Owens, R. J.; Naismith, J. H. Neutralizing Nanobodies Bind SARS-CoV-2 Spike RBD and Block Interaction with ACE2. Nat Struct Mol Biol 2020, 27 (9), 846–854. https://doi.org/10.1038/s41594-020-0469-6.

(5) Pinto, D.; Park, Y.-J.; Beltramello, M.; Walls, A. C.; Tortorici, M. A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; De Marco, A.; Peter, A.; Giarino, B.; Spereaffo, R.; Cameroni, E.; Case, J. B.; Chen, R. E.; Havenar-Daughton, C.; Snell, G.; Tetlenti, A.; Virgin, H. W.; Lanzavecchia, A.; Diamond, M. S.; Fink, K.; Veesler, D.; Corti, D. Cross-Naturalization of SARS-CoV-2 by a Human Monoclonal SARS-CoV Antibody. Nature 2020, 583 (7815), 290–295. https://doi.org/10.1038/s41586-020-2349-y.

(6) 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 USA 2020, 117 (25), 13967–13974. https://doi.org/10.1073/pnas.2008209117.

(7) 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. https://doi.org/10.1038/s41586-020-2180-5.

(8) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural Basis for the Recognition of SARS-CoV-2 by Full-Length Human ACE2. Science 2020, 367 (6485), 1444–1448. https://doi.org/10.1126/science.abb2762.

(9) 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. https://doi.org/10.1038/s41586-020-2179-y.

(10) Amin, M.; Sorour, M. K.; Kasry, A. Comparing the Binding Interactions in the Receptor Binding Domains of SARS-CoV-2 and SARS-CoV. J. Phys. Chem. Lett. 2020, 11 (12), 4897–4900. https://doi.org/10.1021/acs.jpclett.0c01064.

(11) Hassanzadeh, K.; Perez Pena, H.; Dragotto, J.; Buccarelli, L.; Iorio, F.; Pieraccini, S.; Sancini, G.; Feligioni, M. Considerations around the SARS-CoV-2 Spike Protein with Particular Attention to COVID-19 Brain Infection and Neurological Symptoms. ACS Chem. Neurosci. 2020, 11 (15), 2361–2369. https://doi.org/10.1021/acschemneuro.0c00373.

(12) Spinello, A.; Saltalamacchia, A.; Magistrato, A. Is the Rigidity of SARS-CoV-2 Spike Receptor-Binding Motif the Hallmark for Its Enhanced Infectivity? Insights from All-Atom Simulations. J. Phys. Chem. Lett. 2020, 11 (12), 4785–4790. https://doi.org/10.1021/acs.jpclett.0c01148.

(13) Lim, H.; Baek, A.; Kim, J.; Kim, M. S.; Liu, J.; Num, K.-Y.; Yoon, J.; No, K. T. Hot Spot Profiles of SARS-CoV-2 and Human ACE2 Receptor Protein Protein Interaction Obtained by Density Functional Tight Binding Fragment Molecular Orbital Method. Sci Rep 2020, 10 (1), 16862. https://doi.org/10.1038/s41598-020-73820-8.

(14) Shaw, D. E. Molecular Dynamics Simulations Related to SARS-CoV-2 https://www.deshawresearch.com/downloads/download_trajectory_sarscov2.cgi (accessed Sep 15, 2020).

(15) Pach, S.; Nguyen, T. N.; Trimpert, J.; Kunec, D.; Osterrieder, N.; Wolber, G. ACE2-Variants Indicate Potential SARS-CoV-2-Susceptibility in Animals: An Extensive Molecular Dynamics Study; preprint; Pharmacology and Toxicology, 2020. https://doi.org/10.1101/2020.05.14.092767.

(16) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; Fadda, E.; Amaro, R. E. Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. ACS Cent. Sci. 2020, 6 (10), 1722–1734. https://doi.org/10.1021/acscentsci.0c00105.

(17) Cao, L.; Goreshnik, L.; 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, eabdb9909. https://doi.org/10.1126/science.abdb9909.

(18) Panda, P. K.; Arul, M. N.; Patel, P.; Vera, S. K.; Luo, W.; Rubahn, H.-G.; Mishra, Y. K.; Suar, M.; Ahuja, R. Structure-Based Drug Designing and Immunoinformatics Approach for SARS-CoV-2. Sci. Adv. 2020, 6 (28), eabb8097. https://doi.org/10.1126/sciadv.abb8097.
Zheng, M.; Song, L. Novel Antibody Epitopes Dominate the Antigenicity of Spike Glycoprotein in SARS-CoV-2 Compared to SARS-CoV. *Cell Mol Immunol* 2020, 17 (5), 536–538. https://doi.org/10.1038/s41423-020-0385-z.

Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist’s Guide to Molecular Interactions. *J. Med. Chem. 2010*, 53 (14), 5061–5084. https://doi.org/10.1021/jm101112j.

Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. The Magnitude of the CH/π Interaction between Benzene and Some Model Hydrocarbons. *J. Am. Chem. Soc. 2000*, 122 (15), 3746–3753. https://doi.org/10.1021/ja993972j.

Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of the Attraction and Directionality of the NH/π Interaction: Comparison with OH/π and CH/π Interactions. *J. Am. Chem. Soc. 2000*, 122 (46), 11450–11458. https://doi.org/10.1021/ja001901a.

Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of Attraction and Directionality of the π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *J. Am. Chem. Soc. 2002*, 124 (1), 104–112. https://doi.org/10.1021/ja0105212.

Watanabe, C.; Fukuzawa, K.; Okiyama, Y.; Tsukamoto, T.; Kato, A.; Tanaka, S.; Mochizuki, Y.; Nakano, T. Three- and Four-Body Corrected Fragment Molecular Orbital Calculations with a Novel Subdividing Fragmentation Method Applicable to Structure-Based Drug Design. *Journal of Molecular Graphics and Modelling 2013*, 41, 31–42. https://doi.org/10.1016/j.jmgm.2013.01.006.

Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules. *Chemical Physics Letters 1999*, 313 (3–4), 701–706. https://doi.org/10.1016/S0009-2614(99)00874-X.

Fedorov, D.; Kitaura, K. The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems. https://www.routledge.com/The-Fragment-Molecular-Orbital-Method-Practical-Applications-to-Large-Molecular-Fedorov-Kitaura/p/book/9780367385620 (accessed Sep 16, 2020).

Tanaka, S.; Mochizuki, Y.; Komeiji, Y.; Okiyama, Y.; Fukuzawa, K. Electron-Correlated Fragment-Molecular-Orbital Calculations for Biomolecular and Nano Systems. *Phys. Chem. Chem. Phys. 2014*, 16 (22), 10310–10344. https://doi.org/10.1039/C4CP00316K.

Mochizuki, Yuji, Tanaka, Shigenori, Fukuzawa, Kaori (Eds.). Recent Advances of the Fragment Molecular Orbital Method - Enhanced Performance and Applicability | Yuji Mochizuki | Springer; Springer Singapore.

Fedorov, D. G.; Kitaura, K. Pair Interaction Energy Decomposition Analysis. *J. Comput. Chem. 2007*, 28 (1), 222–237. https://doi.org/10.1002/jcc.20496.

Fedorov, D. G.; Kitaura, K. Energy Decomposition Analysis in Solution Based on the Fragment Molecular Orbital Method. *J. Phys. Chem. A 2012*, 116 (1), 704–719. https://doi.org/10.1021/jp209579w.

Tsukamoto, T.; Kato, K.; Kato, A.; Nakano, T.; Mochizuki, Y.; Fukuzawa K. Implementation of Pair Interaction Energy Decomposition Analysis and Its Applications to Protein–Ligand Systems. *J. Comput. Chem. Jpn. 2015*, 14 (1), 1–9. https://doi.org/10.2477/jcc.2014-0039.

Watanabe, C.; Watanabe, H.; Fukuzawa, K.; Parker, L. J.; Okiyama, Y.; Yuki, H.; Yokoyama, S.; Nakano, H.; Tanaka, S.; Homma, T. Theoretical Analysis of Activity Cliffs among Benzoferuranone-Class Pim1 Inhibitors Using the Fragment Molecular Orbital Method with Molecular Mechanics Poisson–Boltzmann Surface Area (FMO+MM-PBSA) Approach. *J. Chem. Inf. Model. 2017*, 57 (12), 2996–3010. https://doi.org/10.1021/acs.jcim.7b00110.

Fukuzawa, K.; Mochizuki, Y.; Tanaka, S.; Kitaura, K.; Nakano, T. Molecular Interactions between Estrogen Receptor and Its Ligand Studied by the Ab Initio Fragment Molecular Orbital Method. *J. Phys. Chem. B 2006*, 110 (47), 24276–24276. https://doi.org/10.1021/jp065705n.

Amari, S.; Aizawa, M.; Zhang, J.; Fukuzawa, K.; Mochizuki, Y.; Iwasawa, Y.; Nakata, K.; Chuman, H.; Nakano, T. VISCANA: Visualized Cluster Analysis of Protein–Ligand Interaction Based on the Ab Initio Fragment Molecular Orbital Method for Virtual Ligand Screening. *J. Chem. Inf. Model. 2006*, 46 (1), 221–230. https://doi.org/10.1021/ci050262q.

Ozawa, T.; Tsuji, E.; Ozawa, M.; Handa, C.; Mukaiyama, H.; Nishimura, T.; Kobayashi, S.; Okazaki, K. The Importance of CH/π Hydrogen Bonds in Rational Drug Design: An Ab Initio Fragment Molecular Orbital Study to Leukocyte-Specific Protein Tyrosine (LCK) Kinase. *Bioorganic & Medicinal Chemistry 2008*, 16 (24), 10311–10318. https://doi.org/10.1016/j.bmcm.2008.10.041.

Hitaoka, S.; Harada, M.; Yoshida, T.; Chuman, H. Correlation Analyses on Binding Affinity of Sialic Acid Analogues with Influenza Virus Neuraminidase-I Using Ab Initio MO Calculations on Their Complex Structures. *J. Chem. Inf. Model. 2010*, 50 (10), 1796–1805. https://doi.org/10.1021/ci100225b.

Yoshida, T.; Munei, Y.; Hitaoka, S.; Chuman, H. Correlation Analyses on Binding Affinity of Substituted Benzene sulfonamides with Carbonic Anhydrase Using Ab Initio MO Calculations on Their Complex Structures. *J. Chem. Inf. Model. 2010*, 50 (5), 850–860. https://doi.org/10.1021/ci100068w.

Hitaoka, S.; Matoba, H.; Harada, M.; Yoshida, T.; Tsuji, D.; Hirokawa, T.; Itoh, K.; Chuman, H. Correlation Analyses on Binding Affinity of Sialic Acid Analogue and Anti-Influenza Drugs with Human Neuraminidase Using Ab Initio MO Calculations on Their Complex Structures – LERE-QSAR Analysis (IV). *J. Chem. Inf. Model. 2011*, 51 (10), 2706–2716. https://doi.org/10.1021/ci200239S.

Munei, Y.; Shimamoto, K.; Harada, M.; Yoshida, T.; Chuman, H. Correlation Analyses on Binding Affinity of Substituted Benzene sulfonamides with Carbonic Anhydrase Using Ab Initio MO Calculations on Their Complex Structures (II). *Bioorganic & Medicinal Chemistry Letters 2011*, 21 (1), 141–144. https://doi.org/10.1016/j.bmcl.2010.11.050.

Okiyama, Y.; Watanabe, C.; Fukuzawa, K.; Mochizuki, Y.; Nakano, T.; Tanaka, S. Fragment Molecular Orbital Calculations with Implicit Solvent Based on the Poisson–Boltzmann Equation: II. Protein and Its Ligand-Binding System Studies. *J. Phys. Chem. B 2019*, 123 (5), 957–973. https://doi.org/10.1021/acs.jpcb.8b09326.

Heifetz, A.; James, T.; Southey, M.; Bodkin, M. J.; Bromidge, S. Guiding Medicinal Chemistry with Fragment Molecular Orbital (FMO) Method. In *Quantum Mechanics in Drug Discovery*; Heifetz, A., Ed.; Springer US: New York, NY, 2020; pp 37–48. https://doi.org/10.1007/978-1-0716-0282-9_3.
lar Orbital Method. *Chemical Physics Letters* **2013**, *556*, 272–277. https://doi.org/10.1016/j.cplett.2012.11.085.

(65) Kato, K.; Honma, T.; Fukuzawa, K. Intermolecular Interaction among Remdesivir, RNA and RNA-Dependent RNA Polymerase of SARS-CoV-2 Analyzed by Fragment Molecular Orbital Calculation. *Journal of Molecular Graphics and Modelling* **2020**, *100*, 107695. https://doi.org/10.1016/j.jmgm.2020.107695.

(66) Harada, T.; Yamagishi, K.; Nakano, T.; Kitaura, K.; Tokiwa, H. Ab Initio Fragment Molecular Orbital Study of Ligand Binding to Human Progesterone Receptor Ligand-Binding Domain. *Naunyn-Schmied Arch Pharmacol* **2008**, *377* (4–6), 607–615. https://doi.org/10.1007/s00210-008-0268-9.

(67) Nutho, B.; Mahalapbutr, P.; Hengphasatporn, K.; Pat.taranggoon, N. C.; Simanon, N.; Shigeta, Y.; Hannong-bua, S.; Rungrotmongkol, T. Why Are Lopinavir and Ritonavir Effective against the Newly Emerged Coronavirus 2019? Atomistic Insights into the Inhibitory Mechanisms. *Biochemistry* **2020**, *59* (18), 1769–1779. https://doi.org/10.1021/acs.biochem.0c00160.

(68) Akisawa, K.; Hatada, R.; Okuwaki, K.; Mochizuki, Y.; Fukuzawa, K.; Komeiji, Y.; Tanaka, S. Interaction Analyses on SARS-CoV-2 Spike Protein Based on Large-Scale Correlated Fragment Molecular Orbital Calculations. 2020. https://doi.org/10.26434/chemrxiv.12941606.v1.
FMO-based hot spot and epitope analyses

SARS-CoV-2 S-protein RBD

ACE2
B38 Fab antibody

Interaction energy (kcal/mol)