

Special Features of COVID-19 in the FMO DB: Fragment Molecular Orbital Calculations and Interaction Energy Analysis of SARS-CoV-2-Related Proteins

Kaori Fukuzawa,* Koichiro Kato,* Chiduru Watanabe,* Yusuke Kawashima, Yuma Handa, Ami Yamamoto, Kazuki Watanabe, Tatsuya Ohyama, Kikuko Kamisaka, Daisuke Takaya, and Teruki Honma*

Cite This: <https://doi.org/10.1021/acs.jcim.1c00694>

Read Online

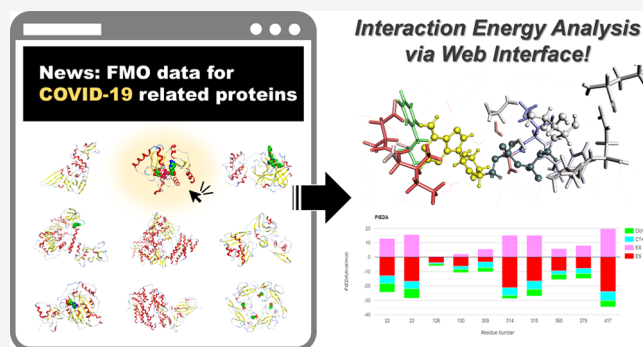
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: SARS-CoV-2 is the causative agent of coronavirus (known as COVID-19), the virus causing the current pandemic. There are ongoing research studies to develop effective therapeutics and vaccines against COVID-19 using various methods and many results have been published. The structure-based drug design of SARS-CoV-2-related proteins is promising, however, reliable information regarding the structural and intra- and intermolecular interactions is required. We have conducted studies based on the fragment molecular orbital (FMO) method for calculating the electronic structures of protein complexes and analyzing their quantitative molecular interactions. This enables us to extensively analyze the molecular interactions in residues or functional group units acting inside the protein complexes. Such precise interaction data are available in the FMO database (FMO DB) (<https://drugdesign.riken.jp/FMO DB/>). Since April 2020, we have performed several FMO calculations on the structures of SARS-CoV-2-related proteins registered in the Protein Data Bank. We have published the results of 681 structures, including three structural proteins and 11 nonstructural proteins, on the COVID-19 special page (as of June 8, 2021). In this paper, we describe the entire COVID-19 special page of the FMO DB and discuss the calculation results for various proteins. These data not only aid the interpretation of experimentally determined structures but also the understanding of protein functions, which is useful for rational drug design for COVID-19.



1. INTRODUCTION

The COVID-19 pandemic has been ongoing since its declaration by WHO in March 2020, and as of June 8, 2021, it had killed 3 718 944 people and infected 172 million people worldwide.¹ COVID-19 has had a devastating impact on global health and economic activities, and the loss of social infrastructure due to urban lockdown has reached unprecedented levels, with no signs of abatement. To fight against COVID-19, there is a need to understand the causative virus, SARS-CoV-2, and develop effective vaccines and therapies. Consequently, structural analyses of SARS-CoV-2-related proteins have been rapidly conducted and are available worldwide in the Protein Data Bank (PDB).² In addition to experimental structures, researchers worldwide collate model structures, molecular dynamics (MD) simulation results, and bioinformatics data and made them available as their data resources.³ The PDB Japan (PDBj)^{4,5} has categorized PDB data of SARS-CoV-2-related proteins into “All entries,” “Representative entries,” and “Latest entries” on the COVID-19 special page, and 1277 structures

are available as of June 8, 2021. The viral proteins include four structural proteins that form virus particles and 16 nonstructural proteins (nsps) intracellularly produced after infecting human cells. The structures of three structural proteins and 13 nsps have already been clarified by cryo-electron microscopy (Cryo-EM), X-ray, and NMR. Though these structural data are very useful information for developing therapeutic agents, there is a need to precisely calculate and clarify how these viral proteins interact with each other and when candidate therapeutic compounds bind strongly to the proteins.

Based on these PDB structures, many molecular simulation studies have been actively carried out, and the most widely

Received: June 18, 2021

performed are classical MD calculations. For example, microsecond MD simulations^{6–8} of SARS-CoV-2-related proteins have been performed using special-purpose computers such as Anton2⁹ and MDGRAPE-4A,¹⁰ and general-purpose software such as GENESIS-2.0,¹¹ NAMD,¹² and GROMACS.¹³ In particular, many dynamical structural analysis studies have been performed, such as the movement of the open structure of the spike (S) protein,⁸ the evaluation of the effect of glycans,^{8,14,15} the binding mechanism of the receptor-binding domain (RBD) of the S protein to the angiotensin-converting enzyme 2 (ACE2) or antibodies,¹⁵ the binding of the main protease (Mpro) to inhibitors,^{6,16} and the apo and the remdesivir-binding structure of RNA-dependent RNA polymerase (RdRp).^{17,18} As for QM calculation, the fragment molecular orbital (FMO) method^{19,20} is one of the methods that can treat the entire protein with uniform accuracy. The FMO calculation for several specific SARS-CoV-2-related proteins, such as the S protein,^{21–25} Mpro,^{26–30} and RdRp,³¹ have been performed using the ABINIT-MP program^{32,33} and the GAMESS program.^{34–36} The ABINIT-MP program used in this paper is capable of performing precise electron correlation calculations at the MP4 (SDQ) level for S proteins of up to 3300 residues using the Fugaku supercomputer.²² On the other hand, the FMO-DFTB method included in GAMESS is suitable for screening owing to its much lower computational cost.³⁷ The FMO-DFTB method has also been employed as a part of the *in silico* drug discovery pipeline using supercomputers.³⁸ In contrast to classical MD calculations, which reveal the dynamic behavior and molecular mechanisms of proteins, QM calculations can reveal the precise electronic structure and interactions in a stationary structure (often an X-ray or cryo-EM structure). These two types of molecular simulations are used in complementary ways.

Here, we have performed several FMO-based^{19,20} quantum chemical calculations on the entire structure of SARS-CoV-2-related proteins, focusing on the representative entries of PDBj, and all of the results have been published in the FMO database (FMOdb)^{39,40} since April 2020. In FMOdb, all FMO calculation results can be downloaded and molecular interaction analysis can be performed through the web interface and BioStation Viewer software.⁴¹ In this paper, we describe the FMO calculation results (681 structures as of June 8, 2021) available on the COVID-19 special page of the FMOdb.

2. METHODS

2.1. Molecular Modeling. We performed FMO calculations for 681 protein structures, mainly from the representative entries published on the COVID-19 special page of the PDBj to reveal their precise interaction energies. MOE⁴² was employed for molecular modeling, and structural refinement was performed according to the resolution of the registered structures following the procedure shown in Figure 1. Here, missing atoms in the PDB structure are complemented, and appropriate structural optimization is performed with constraints to avoid large deviations from the experimental structure. The level of structural optimization is changed according to the experimental method and its resolution but, for some data, exceptions are made, such as performing structural optimization for nonhydrogen atoms, even when the resolution is less than 2.0.

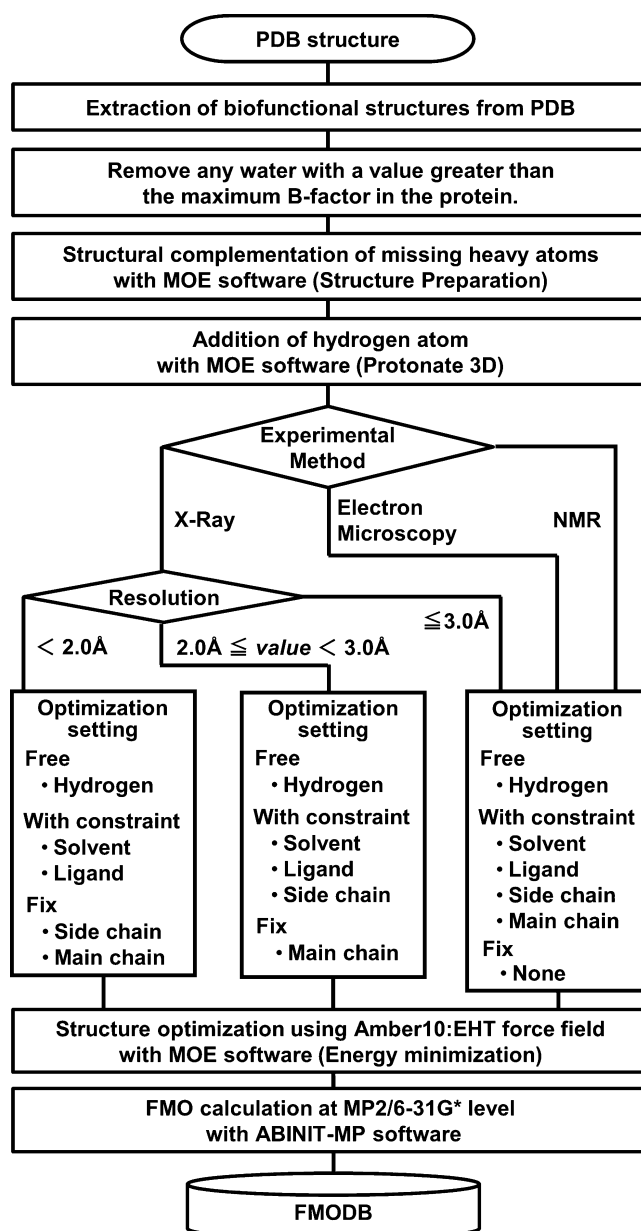


Figure 1. Sequence of workflow from downloading the PDB structure to FMOdb registration.

2.2. FMO Calculations. The basic fragmentation method for FMO calculations is as follows: proteins are fragmented into amino acid residue units, nucleic acids are fragmented into backbone and base units, and ligands are fragmented into one or several fragments.⁴³ Quantum chemical calculations were performed at the theoretical level of FMO2-MP2/6-31G*^{44,45} to calculate the total energy (E_{total}), atomic charge, and interfragment interaction energy (IFIE; $\Delta\tilde{E}_{IJ}$).

$$E_{\text{total}} = \sum_I E'_I + \sum_{I>J} \Delta\tilde{E}_{IJ} \quad (1)$$

The IFIE can be further decomposed into four energy components, including electrostatic (ES), exchange repulsion (EX), charge transfer with higher-order terms (CT+mix, hereafter simply referred to as CT in the text), and dispersion (DI) terms. This is known as the pair interaction energy

decomposition analysis (PIEDA),^{46,47} and is expressed as follows

$$\Delta\tilde{E}_{IJ} = \Delta\tilde{E}_{IJ}^{\text{ES}} + \Delta\tilde{E}_{IJ}^{\text{EX}} + \Delta\tilde{E}_{IJ}^{\text{CT+mix}} + \Delta\tilde{E}_{IJ}^{\text{DI}} \quad (2)$$

PIEDA is vital in applying the FMO method to drug discovery because it gives information about the characteristics of the interaction, in addition to their magnitude (stable or unstable). For interpreting PIEDA, refer to FMO books^{20,34} and the original reference.⁴⁶ Hydrogen bonding is mainly detected as the stabilization energy for the ES and CT terms, and the hydrophobic type of interactions, such as CH/ π and π - π , are mainly detected as the stabilization energy for the DI term.

Among the data registered in the FMO DB, for entries for which the binding energy can be defined, such as protein–ligand, protein–protein, and protein–RNA, the binding energy is calculated from the sum of IFIEs. The binding energy between molecules A and B (ΔE_{AB}) is expressed as follows:

$$\Delta E_{\text{AB}} = \sum_{\substack{I \in \text{A} \\ J \in \text{B}}} \Delta\tilde{E}_{IJ} \quad (3)$$

All calculations in the FMO DB were performed using ABINIT-MP,^{32,33} the FMO calculation software. We used supercomputers, including TSUBAME3.0 at the Tokyo Institute of Technology, Oakforest-PACS at JCAHPC, and Fugaku and HOKUSAI at RIKEN.

2.3. FMO DB-Registered Information and Web Interface. The FMO calculation results of COVID-19-related proteins are available from the top page of the FMO DB or the COVID-19 special page.³⁹ Each FMO DB ID page contains information on the used PDB structures and calculation conditions, and the entire dataset, including ABINIT-MP input and output files, can be downloaded. In addition, the binding energies and IFIE/PIEDA energy lists are provided to analyze the interaction data using the web interface. BioStation Viewer⁴¹ can be used to perform a detailed analysis using the downloaded data. As an example, Figure 2 shows a detailed view of the FMO DB for the complex structure of the main protease and ligand. The upper part shows the structural data and calculation conditions, and the lower part shows the interaction energies with amino acid residues near the ligand by the PIEDA energy component. The bottom chart shows the interaction energies with amino acid residues near the ligand by the PIEDA energy component. Conditions, such as the range to be displayed, can be specified.

3. RESULTS AND DISCUSSION

SARS-CoV-2 is a single-strand positive-sense RNA virus composed of four structural proteins and RNA genes. After infecting a host cell, the virus synthesizes 16 types of nsps. S protein, a type of structural protein that protrudes from the viral surface, plays an important role in the process of viral infection of the hosts, making it an important target for vaccines and therapeutic agents.^{48,49} For nsps, protease and RdRp are also promising drug targets. Two proteases of SARS-CoV-2 are involved in the cleavage of polyproteins synthesized from RNA to produce various proteins. The Mpro cleaves the polyprotein at 11 sites, and the papain-like protease (PLpro) cleaves the polyprotein at three specific sites.⁵⁰ RdRp and helicase are the key enzymes in gene

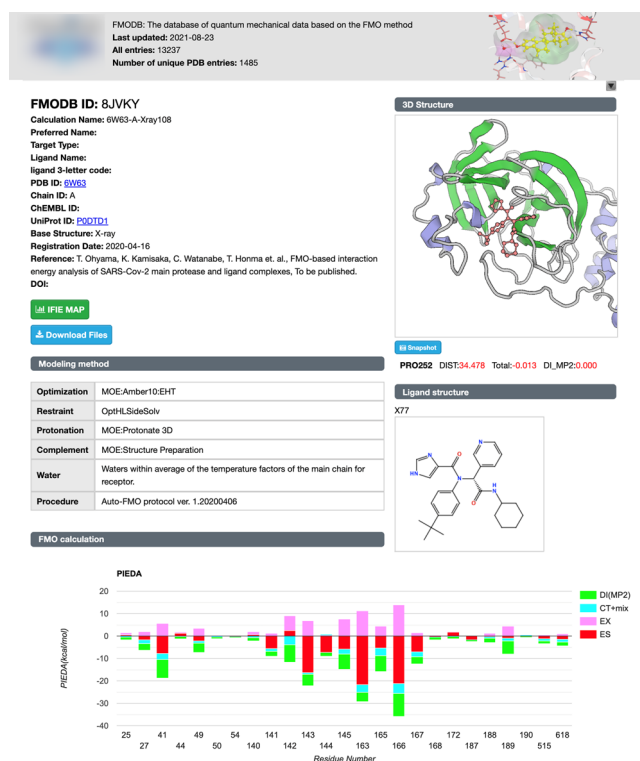


Figure 2. Web interface of the FMO DB.

duplication by creating a replication and transcription complex (RTC).⁵¹

The number of registered structures on June 8, 2021 is shown in Figure 3. About 50% of the 681 structures registered in the FMO DB are the Mpro, followed by S protein and PLpro. In addition, ADP ribose phosphatase, endoribonuclease, 2'-O-ribose methyltransferase, helicase, etc., have been widely registered. The experimentally determined structures of these SARS-CoV-2-related proteins are shown also in Figure S1. The number of entries per protein depends on the number of representative entries in the PDBj site, and the FMO DB preferentially collects the structures with the highest resolution among the PDB entries with 100% identical amino acid sequences, i.e., representative structures in the PDBj. The FMO DB also preferentially collects the structures whose primary citation information is publicly available from the PDB site.

The following is a typical example of registered structures.

3.1. Papain-like Protease (nsp3). PLpro is an essential coronavirus enzyme required for the cleavage of viral polyproteins at three sites, and the catalytic residues for processing are Cys111 and His272.^{52,53} The substrate-binding site of PLpro is exposed to solvent and contains a flexible β -hairpin loop (G266–G271) called the blocking loop 2 (BL2), which changes its structure significantly depending on the substrate and inhibitor configuration.⁵⁴ PLpro is also involved in the cleavage of post-translational modifications of proteins on host proteins (deubiquitination) as an evasion mechanism against the host antiviral immune responses.^{55,56} Some examples of covalent and noncovalent ligand complexes registered in the FMO DB are as follows.

As an example of FMO calculations, we show the results for the X-ray crystal structure of the tetrapeptide Ac-hTyr-Dap-Gly-Gly-VME (VIR251) complex, which is one of the

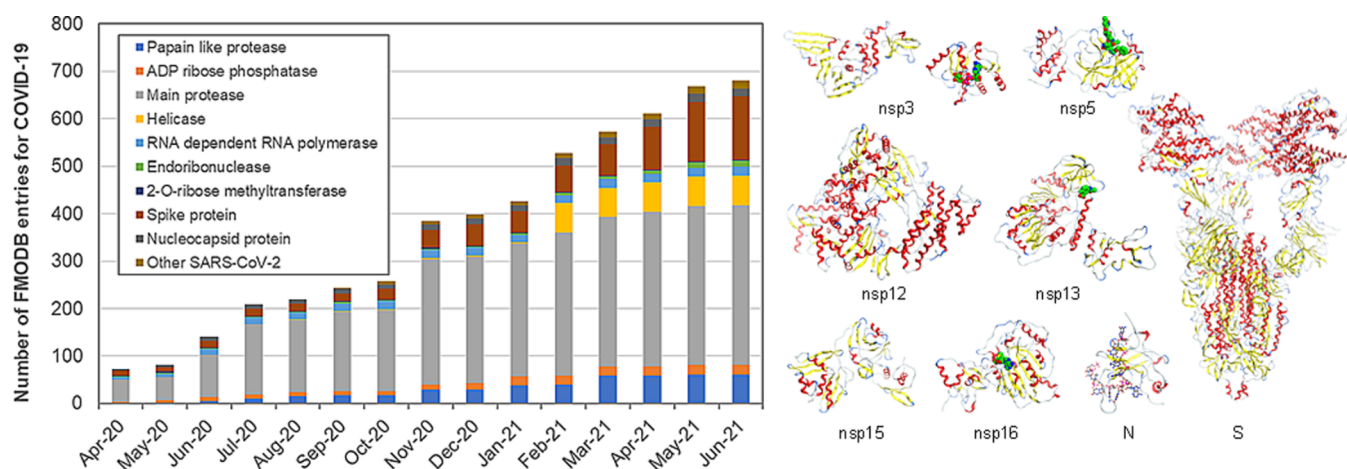


Figure 3. Number of registered structures in the FMOB; papain-like protease (nsp3), ADP ribose phosphatase (nsp3), main protease (nsp5), RNA-dependent RNA polymerase (nsp12), helicase (nsp13), endoribonuclease (nsp15), 2'-O-ribose methyltransferase (nsp16), spike (S) protein, nucleocapsid (N) protein, and others.

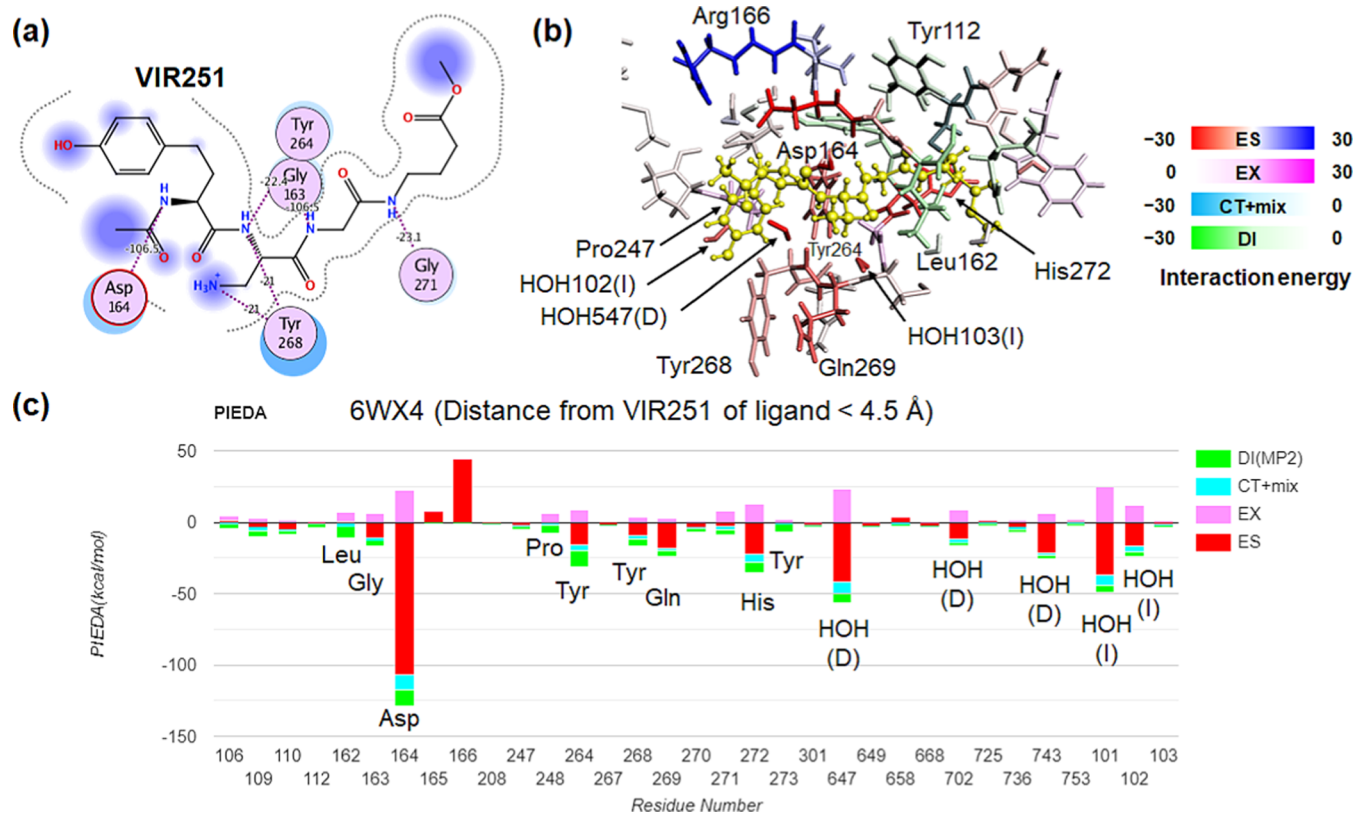


Figure 4. Interaction between papain-like protease (PLpro) and its peptide inhibitor, VIR251 (PDB ID: 6WX4, FMOB ID: YQG52). (a) The interaction site between the entire ligand molecule and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the ligand (yellow) and a residue fragment within 4.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMOB web interface. Energy values are in kcal/mol.

peptide inhibitors of PLpro and is covalently bound to PLpro at Cys111 (PDB ID: 6WX4,⁵⁷ FMOB ID: YQG52). VIR251 strongly inhibits the activity of both SARS-CoV PLpro and SARS-CoV-2 PLpro.⁵⁷ Although VIR251 is a four-residue ligand, we treated the entire ligand as a single fragment (the residue names of the ligands are registered as LIG).

Figure 4 shows the interaction between the entire ligand molecule and its surroundings and the interaction energy quantified by PIEDA. The binding energy between VIR251

and SARS-CoV-2 PLpro, according to eq 3, was -170.6 kcal/mol (except for the interaction with Cys111, which is covalently bound to VIR251). ES (-132.9 kcal/mol) contributed most to the binding energy, followed by DI (-84.2 kcal/mol) and CT (-41.3 kcal/mol). Since VIR251 has a positive charge ($+1e$), the ES interaction (-107.2 kcal/mol) with the acidic residue Asp164, which forms a hydrogen bond with the NH of hTyr, contributes significantly to the binding energy. In addition, the ES and CT components of the hydrogen bond-forming residues, such as Gly163, Tyr264,

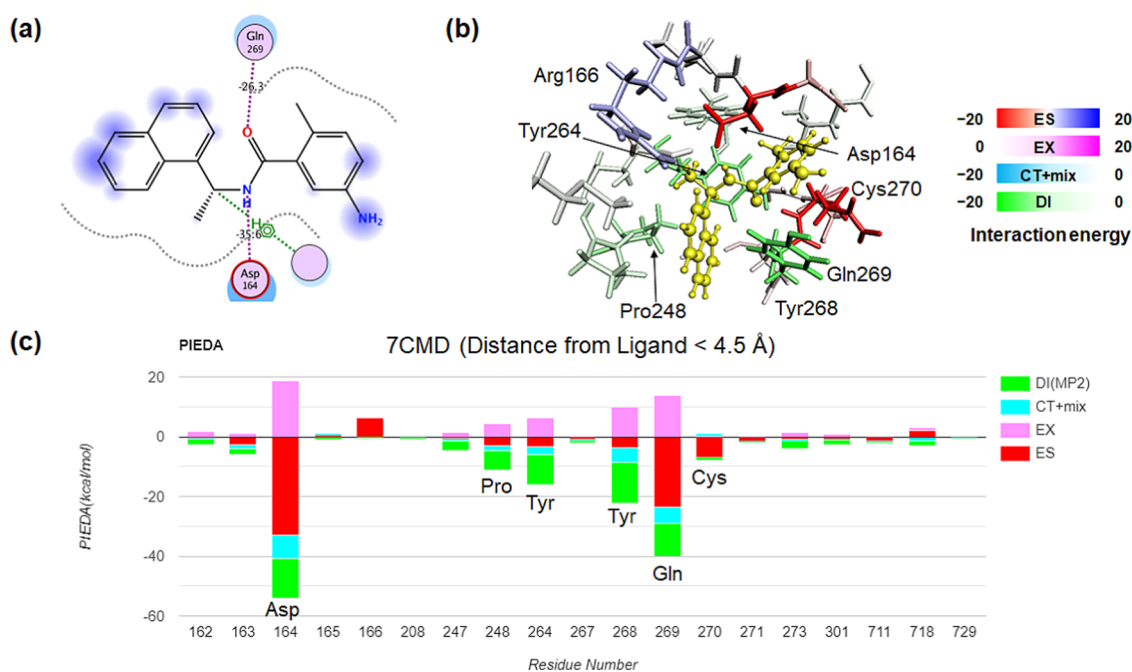


Figure 5. Interaction between PLpro and the noncovalent ligand (PDB ID: 7CMD-A). (a) The interaction site between the entire ligand molecule and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the ligand (yellow) and a residue fragment within 4.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMOODB web interface. Energy values are in kcal/mol.

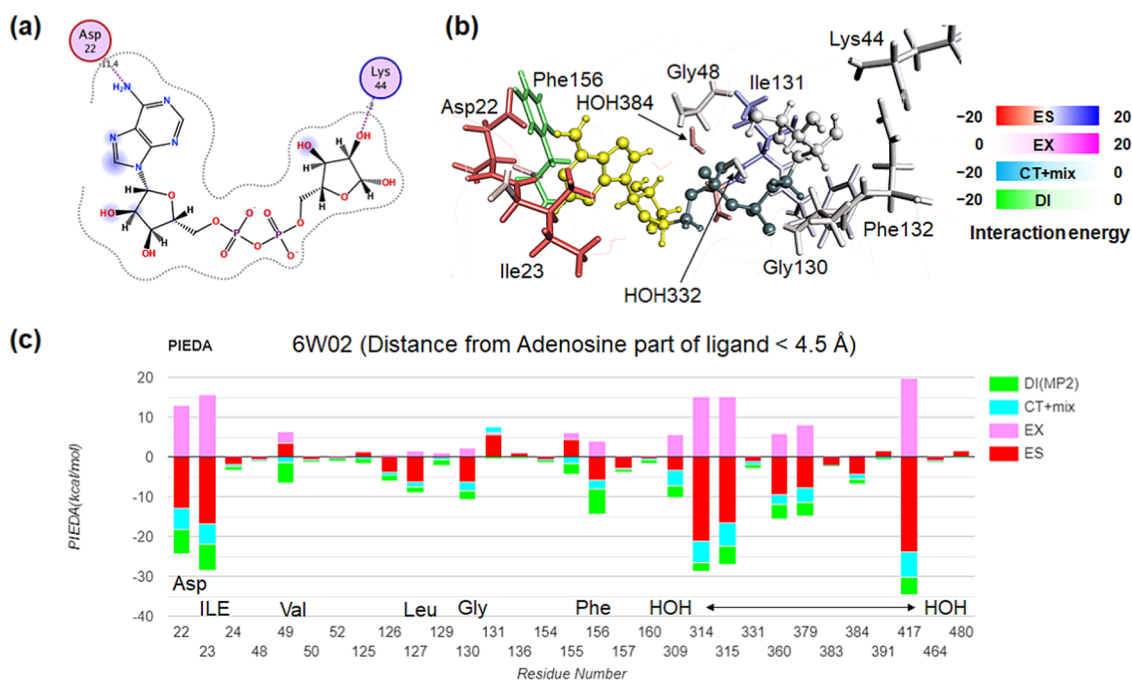


Figure 6. Interaction between ADP ribose phosphatase and ADP ribose as a ligand (PDB ID: 6W02, FMOODB ID: 1JLYZ). (a) The interaction site between the entire ligand molecule and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the adenosine portion of the ligand (yellow) and a residue fragment within 4.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMOODB web interface. Energy values are in kcal/mol.

Tyr268, Gly271, etc., and the DI component due to the CH/ π interactions with surrounding hydrophobic residues, such as Leu162, Pro248, Tyr264, etc., also contributed significantly to the binding energy.

In the interaction with the catalytic His272 fragment, a dispersion interaction derived from the imidazole ring of His272 was observed, in addition to the hydrogen bond with

the C=O derived from Gly271. Note that the C=O from Gly271 belongs to the His272 fragment due to fragmentation. Strong interactions between VIR251 and several water molecules were also detected, suggesting that the role of water molecules in bridging ligands and PLpro is important in designing new inhibitors. Details of these energetic analyses can be performed freely using the FMOODB.

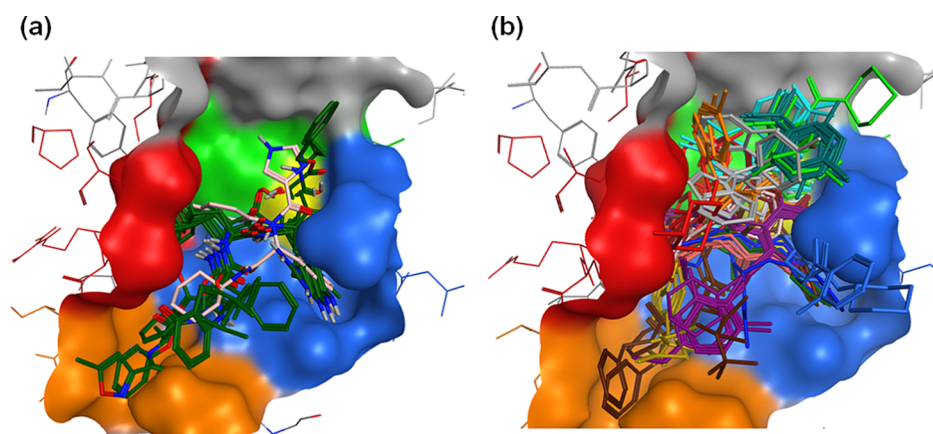


Figure 7. Substrate-binding pocket of the main protease (yellow area: Cys145; green area: S1' subsite; blue area: S1 subsite; red area: S2 subsite; orange area: S4 subsite). (a) Only ligands of complexes classified as representative structures in the PDBj are indicated by sticks (green: 8 covalent ligands; pink: 1 noncovalent ligand). (b) Clustering of 110 complex structures and 14 types of clusters to which four or more ligands belong.

Next, the FMO results for the X-ray crystal structure of the tetramer with the neutral molecule GRL0617 bound as a noncovalent ligand have also been registered for each monomer (PDB ID: 7CMD,⁵⁴ FMOB ID: R1278, Z284N, 6321Z, 2KQVR). GRL0617 has an IC_{50} of $2.2 \pm 0.3 \mu\text{M/L}$ and is considered a promising candidate compound.⁵⁴ The average binding energy of GRL0617 and each monomer registered in FMOB was -112.5 ± 2.7 kcal/mol, which is about two-thirds of that of the positively charged covalent ligand VIR251. The number of interacting residues was also smaller than that of VR251, but the CH/ π interaction with Tyr268 in the BL2 loop, the hydrogen bond with Gln269, and the hydrogen bond and CH/ π interaction with Asp164 in the α 3-to- α 4 loop were large (Figure 5), suggesting that GRL0617 inhibits the inversion of the BL2 loop by interacting with these loops, thereby inhibiting the recognition of the substrate, which is consistent with the structural considerations.⁵⁴

3.2. ADP Ribose Phosphatase (nsp3). Nsp3 encodes the largest multidomain, including the ADP ribose phosphatase domain. Although its function has not yet been elucidated, it is inferred to be related to the ability of proteins to remove ADP-ribose from ADP-ribosylated proteins and RNA.⁵⁸ The X-ray crystal structures of the complexes with ADP, AMP, etc., have been analyzed and they include catalytic water.⁵⁸

The complex of ADP ribose phosphatase and ADP ribose (PDB ID: 6W02, FMOB ID: 1JLYZ, Figure 6a), which is drawn based on the atomic coordinates of the PDB crystal structure, shows that only the interaction between Asp22 and Lys44 is extracted. On the other hand, the PIEDA energies of the adenosine portion of ADP ribose in FMO calculations (Figure 6b,c) show that adenine is stabilized by DI with Phe156 and ES with Asp22 and Ile23 (C=O of the Asp22 main chain in the biochemical description). The side chains of Phe156 and Ile23 interact with the adenine base in a CH/ π interaction, indicating a relatively large DI component. A similar analysis on the diphosphate moiety of the ligand showed strong ES interactions with the two surrounding hydrogen-bonded water molecules, Val49, Gly130, Ile131, and Phe132. The analysis of the ribose moiety showed strong ES interactions with Ala39, Asn40, and Gly48. Lys44 also showed strong ES interactions with the diphosphate moiety,

although they were separated by more than 4.5 Å. The details of the interaction energy values from these single fragments are available on the FMOB web interface. When the ligand consists of multiple fragments, the BioStation Viewer⁴¹ can be used for more detailed ligand interaction energy analysis. Figure S2 shows a mapping of the interaction energies between the entire ligand molecule and the surrounding residues. From the whole ligand molecule, the interactions with two water molecules (HOH332 and HOH384) and Gly130, which are involved in the phosphatase activity, are more clearly shown.

3.3. Main Protease (nsp5). The Mpro is an enzymatic protein encoded in the viral genome, which is one of the two coronavirus proteases responsible for the cleavage of polyproteins translated from genomic RNA and is involved in the maturation of viral proteins.⁵⁰ The Mpro has been studied worldwide as one of the most important drug targets for COVID-19. In structural biology research, the first crystal structure of SARS-CoV-2-related protein was the covalent complex structure of the Mpro and the N3 inhibitor (PDB ID: 6LU7⁵⁹). The structures of the Mpro in the PDB are obtained from X-ray crystallography. Most of the structures are complex with ligands, and the binding mode is either noncovalent or covalent with Cys145. For the complex of the Mpro and the N3 inhibitor (PDB ID: 6LU7), FMO calculations were performed and registered in the FMOB by Hatada et al.²⁶ when the PDB structure was published (FMOB ID: R1GK8). In addition, statistical interaction analyses using FMO calculations for several sampled structures from classical MD calculations have been reported.²⁷ Among the COVID-19-related proteins in the FMOB, the Mpro has the highest number of registered structures.

The Mpro is a three-domain cysteine protease, and the substrate-binding pocket consists of four subsites (S1', S1, S2, and S4) and the Cys145–His41 catalytic dyad.^{60,61} Here, we classify the features of the pockets regarding the interaction of 110 ligand complex structures registered in the FMOB (Figure 7). Figure 7a shows the superposition of eight covalent structures and one noncovalent structure, which are classified as representative structures in the PDBj. The covalent structures, in particular, contact all four subsites. We performed a self-organizing map (SOM) clustering analysis⁶²

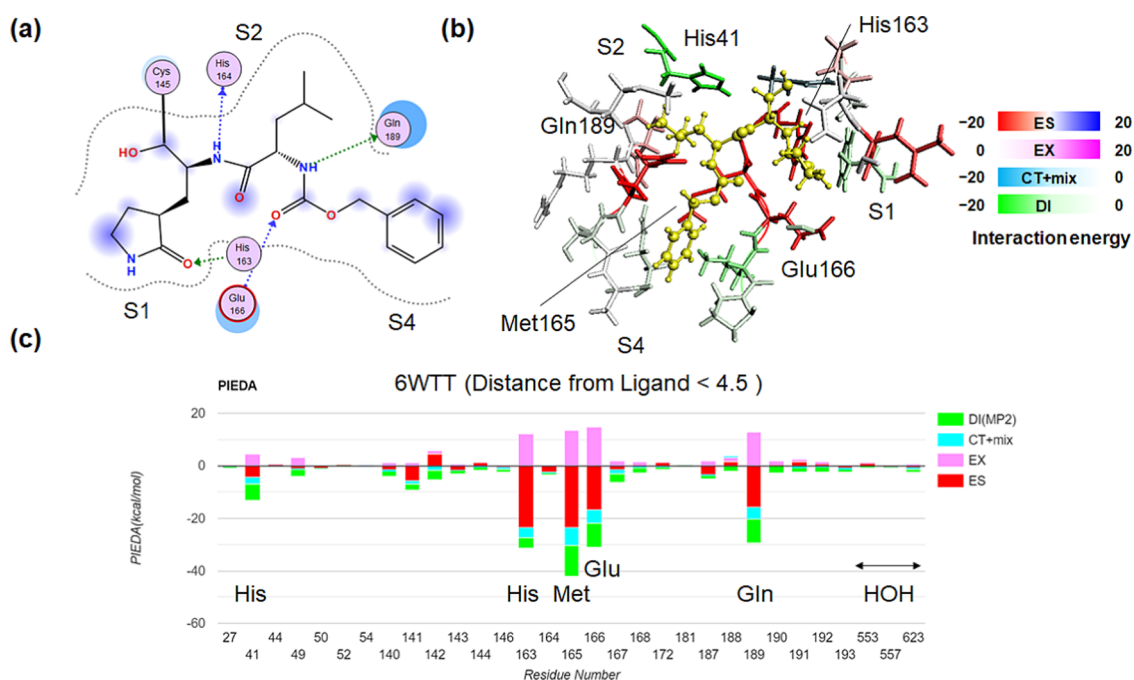


Figure 8. Interaction between the main protease and the covalently bound ligand (PDB ID: 6WTT, FMO DB ID: 4NVVN). (a) The interaction site between the entire ligand molecule and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the ligand (yellow) and a residue fragment within 4.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMO DB web interface. Energy values are in kcal/mol.

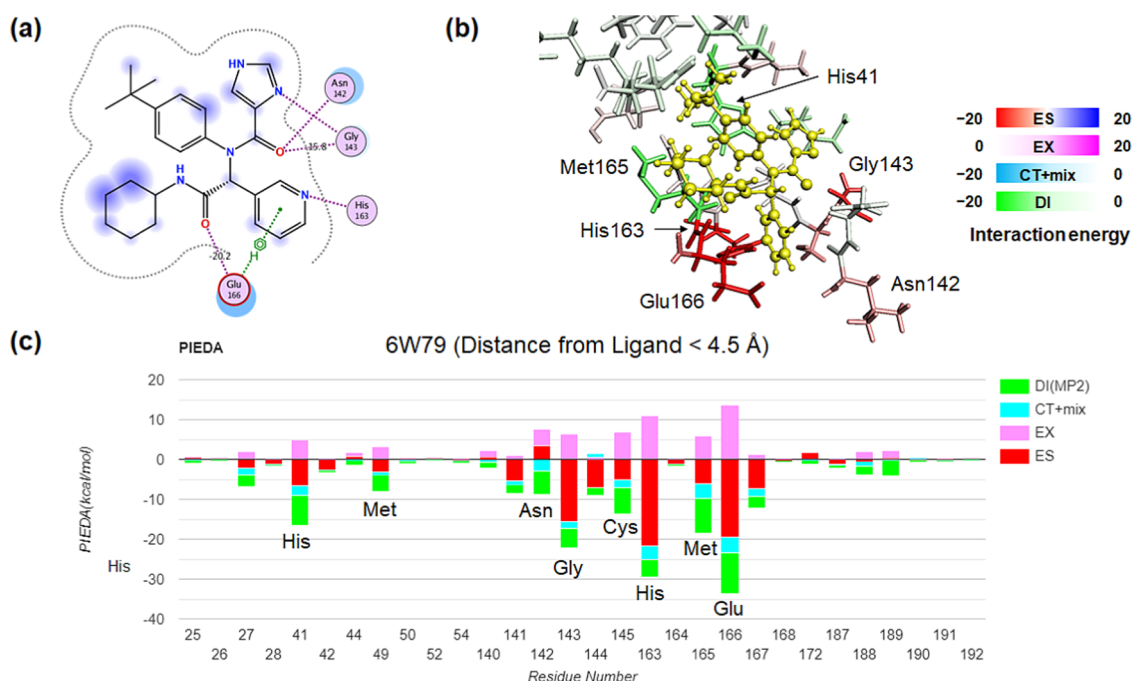


Figure 9. Interaction between the main protease and the noncovalent ligand (PDB ID: 6W79, FMO DB ID: N3QNQ). (a) The interaction site between the entire ligand molecule and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the ligand (yellow) and a residue fragment within 4.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMO DB web interface. Energy values are in kcal/mol.

of the interactions based on PIEDA energies for the 110 ligand complex structures. Figure 7b shows 14 clusters to which four or more ligands belong (figures for each cluster can be found in Figure S3). Most of the clusters have ligands localized to a single site and some ligand clusters were bound across multiple sites.

All FMO calculation results for these structures are registered in the FMO DB and can be accessed by anyone. Here, we discuss one example of each of the covalent and noncovalent ligands shown in Figure 7a. Figure 8 depicts the interaction of an example of covalent ligands (PDB ID: 6WTT,⁶³ FMO DB ID: 4NVVN). The prodrug of peptide

ligands, GC-376, is covalently bound to Cys145 of the Mpro after changing to GC-373. The S1, S2, and S4 binding sites interact with the substituents of the ligand (Figure 8a), and PIEDA analysis showed that the interacting residues are His41, His163, Met165 (His164 C=O), Glu166, and Gln189 (Figure 8b,c). In the S1 site, χ -lactam of the ligand is hydrogen-bonded to His163, the side chain of Glu166, and Met165 (C=O of the His164 main chain in the biochemical notation), resulting in a strong ES + CT interaction. In the S2 site, the isopropyl group of the ligand interacts with the His41 of the dyad in CH/ π . In the S4 site, the amide moiety of the ligand is hydrogen-bonded to the main chains of Glu166 and Gln189, which are stabilized by ES + CT. Gln189 also has a large DI component, which may be attributed to the dispersion interaction between the benzene ring of the ligand and the α and β carbons of the Gln189 side chain. The benzene ring also has a dispersion interaction with the methyl group of Met165 (Figure 8c).⁶³ The binding energy is -130.6 kcal/mol, except for Cys145, which is covalently bound.

In the X-ray structure of the complex of the Mpro with the noncovalent ligand X77 (PDB ID: 6W79, FMO DB ID: N3QNQ), the specific interactions with His41, Asn142, Gly143, His163, Met165, and Glu166 were confirmed by FMO calculations (Figure 9). According to the PIEDA component shown in Figure 9b,c, as in the case of the covalent ligands, the CH/ π interaction (DI) of His41 and the hydrogen bonds (ES + CT) between His163 and Met165 (C=O of His164) are maintained but there is no DI in the Met165 side chain. For Glu166, the hydrogen bond in the main chain (ES + CT) is preserved but the side-chain hydrogen bond is absent, and the CH/ π interaction is formed between the β -carbon and the pyridine ring of the ligand (shown as the DI component). On the other hand, several additional interacting residues are observed around the imidazole ring. In particular, there is a CH/O bond with Asn142, a hydrogen bond with Glu143, and a dispersion interaction with Cys145 since it is a noncovalent ligand. The ligand-binding energy is -149.4 kcal/mol, which is more stabilizing than that of the covalent ligand. A comparison between Figures 8c and 9c shows that the noncovalent ligand has more interactions than the covalent ligand.

3.4. RNA-Dependent RNA Polymerase (nsp12) Complex. RdRp is an RNA synthesis enzyme involved in the viral genome replication, forming RTC with cofactors nsp7–nsp9 and helicase nsp13^{51,64} (Figure 10). RdRp is responsible for replicating the genome, which is most

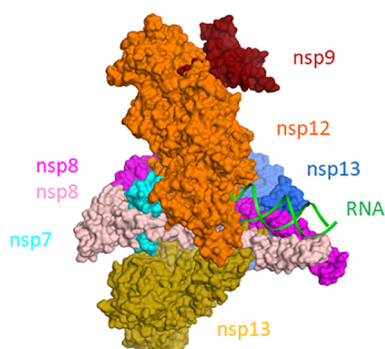


Figure 10. Replication and transcription complex of SARS-CoV-2 (PDB ID: 7CYQ).

important for the survival of the virus. RdRp is an important target of antiviral drugs, including remdesivir and favipiravir, and several ligand-containing PDB structures have been published. The structures registered in the FMO DB are the RdRp–nsp7–nsp8–RNA–remdesivir complex (PDB ID: 7BV2,⁶⁵ FMO DB ID: 1JL3Z³¹), the complex without ligand (PDB ID: 6YYT),⁶⁴ the complex without RNA (PDB IDs: 7BV1, 6M71, 7BW4), nsp7–nsp8 complexes, and nsp9 alone.

The FMO calculations by Kato et al.³¹ for the structure of the complex containing nucleic acid analogue remdesivir show that remdesivir incorporated into the end of the template-primer-RNA duplex retains a nucleotide mimic interaction and is stabilized by hydrogen bonding between base pairs and π – π stacking between inter- and intra-RNA strands (Figure 11). In particular, the stacking interaction is enhanced by the carbon-to-nitrogen substitution introduced into the adenine backbone of remdesivir. In addition, the interaction between remdesivir and RdRp is stabilized by hydrogen bonding with the Asn691 and ES interaction with Asp760. Furthermore, the interaction of remdesivir with Thr687 is stabilized by OH/ π interactions using π electrons of the cyano group introduced into the sugar of remdesivir, in addition to hydrogen bonding. Thus, the interaction of remdesivir in the inhibition process of the strand elongation is clarified. In the interaction between the RNA duplex and the protein, RdRp (nsp12) is mainly responsible for stabilization (about -600 kcal/mol), nsp8 is responsible for about half (-280 kcal/mol), and nsp7 is responsible for -24 kcal/mol. Figure S4 shows a comprehensive interaction-energy diagram, IFIE-MAP.

3.5. Helicase (nsp13). Helicase forms an RTC complex with RdRp, and both RdRp and helicase are essential for viral replication.⁶⁶ Helicase may be useful for smooth replication of the viral genome, such as backtracking RdRp⁶⁷ and unwinding RNA.⁶⁸ The FMO DB has entries on the complex structure with (3-fluoro-4-methylphenyl)methanesulfonamide (PDB ID: SRL6, FMO DB ID: N3YLQ) and about 60 apo structures. The results of the complexes (Figure 12) show strong hydrogen bonding between the amide group of the ligand and Lys192, Tyr224, Thr228, and HOH833. Most of the interactions with the ligand are ES but only Val226 exhibits a CH/ π interaction with the aromatic ring of the ligand, indicating stabilization of the DI component. However, in this calculation, the Zn atom coordinating with helicase is neglected and the FMO calculation that includes Zn is in progress. We shall also perform FMO calculations for the RTC giant complexes (PDB ID: 6XEZ, 7CXN, etc.) consisting of RdRp–helicase–RNA–ADP.

3.6. Endoribonuclease (nsp15). Endoribonuclease is a uridylyte-specific endoribonuclease that cleaves the poly-U sequence of RNA,⁶⁹ and a structure with uridine-5'-monophosphate bound is available in the PDB (PDB ID: 6WLC, FMO DB ID: N3G7Q).⁷⁰ The coordinate information of the crystal structure (Figure 13a) suggests only Ser294 as an interacting residue between the whole ligand molecule and its surroundings. On the other hand, the PIEDA interaction energies calculated by FMO (Figure 13b,c) show a strong ES interaction between the uridine portion of the ligand and Lys290 and Ser294, as well as stabilization by DI interactions with Tyr343. A similar analysis on the monophosphate moiety of the ligand showed strong ES interactions with hydrogen-bonded Tyr343, water molecules (HOH540, HOH646), and Lys290. Figure S5 shows the interaction

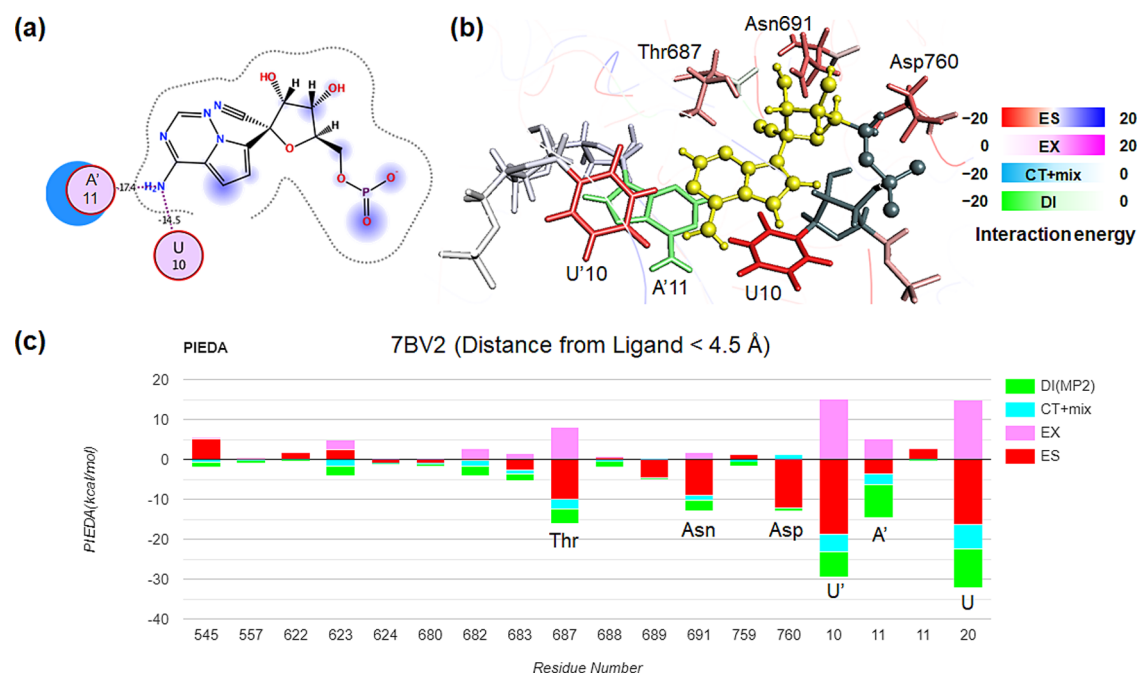


Figure 11. FMO results for the RNA-dependent polymerase, nsp7, nsp8, RNA, and remdesivir complex (PDB ID: 7BV2, FMO DB ID: 1JL3Z). (a) The interaction site between the ligand (remdesivir) and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the ligand (yellow) and a residue fragment within 4.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMO DB web interface. Energy values are in kcal/mol.

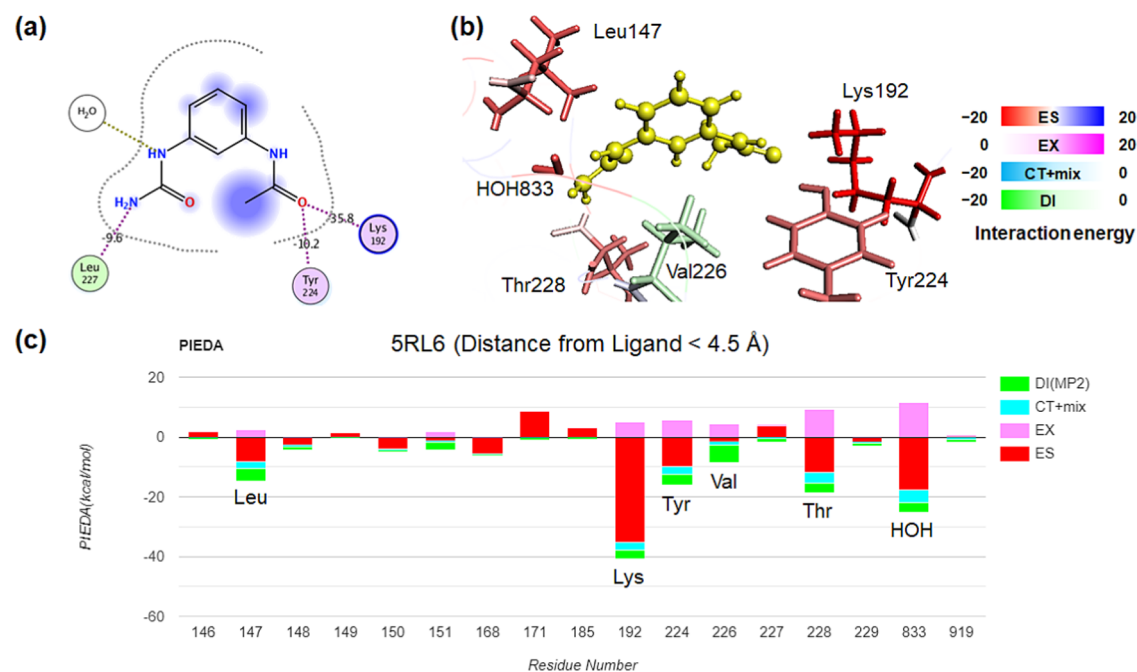


Figure 12. FMO results for the helicase–ligand complex (PDB ID: 5RL6, FMO DB ID: N3YLQ). (a) The interaction site between the entire ligand molecule and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the ligand (yellow) and a residue fragment within 4.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMO DB web interface. Energy values are in kcal/mol.

energies of the entire ligand molecule and its surroundings mapped onto the structure. From the whole ligand molecule, strong ES interactions with Ser294 and Tyr343, which are thought to be important for RNA recognition and degradation, and stabilization by DI interactions with His250 are observed.

3.7. 2'-O-Ribose Methyltransferase (nsp16). Nsp16 is a methyltransferase that forms a complex with nsp10 to transfer a methyl group from a donor to an acceptor, catalyzing methylation at the ribose 2'-O position of nucleotides at the end of viral RNA. The crystal structures of nsp16, nsp10, and their complexes with the methyl group donor, S-adenosylmethionine (SAM), or the demethylation

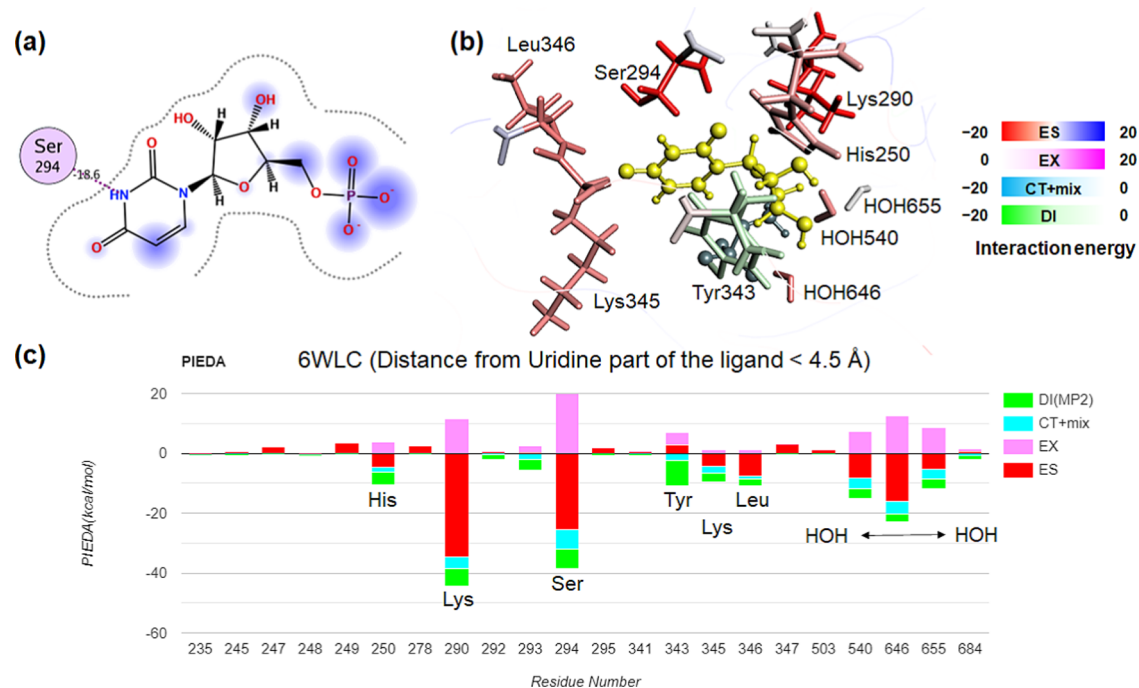


Figure 13. FMO results for the endoribonuclease–ligand complex (PDB ID: 6WLC, FMO DB ID: N3G7Q). (a) The interaction site between the entire ligand molecule and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the uridine portion of the ligand (yellow) and a residue fragment within 4.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMO DB web interface. Energy values are in kcal/mol.

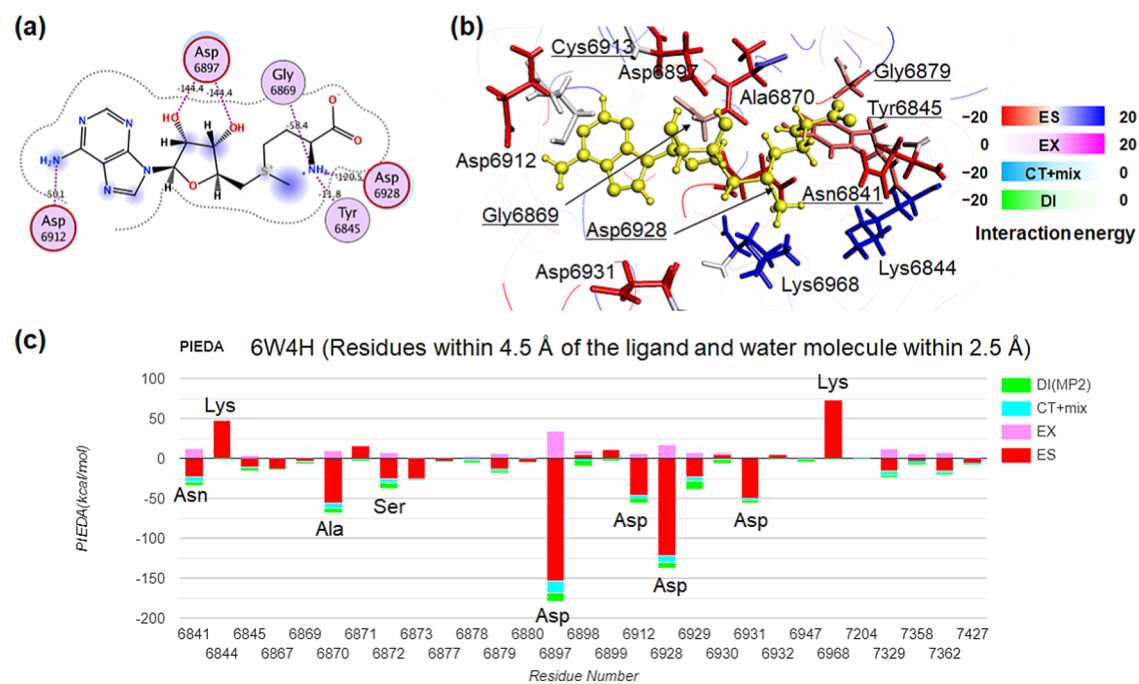


Figure 14. FMO results for the 2'-O-ribose methyltransferase–SAM complex (PDB ID: 6W4H, FMO DB ID: JM9Y9). (a) The interaction site between the entire ligand (SAM) and the surrounding residues, as drawn by MOE. (b) The PIEDA main component between the ligand (yellow) and a residue fragment within 4.5 Å and a water molecule within 2.5 Å, mapped on the molecular structure using a BioStation Viewer. (c) Graphical representation of the PIEDA components as in (b) using the FMO DB web interface. Energy values are in kcal/mol.

product, S-adenosyl-L-homocysteine (SAH), have been published.^{71,72} The results of the analysis of the SAM complex are shown in Figure 14 (PDB ID: 6W4H,⁷² FMO DB ID: JM9Y9). The coordinate information of the crystal structure (Figure 14a) suggests an interaction of three Asp residues (6897, 6912, 6928) with Gly6869 and Tyr6845.

On the other hand, the PIEDA results show that the positively charged SAM is stabilized by strong ES interactions with Asp6897, Asp6912, Asp6928, Asp6931, Asn6841, and Ala6870 (Figure 14b,c). The binding pocket of SAM is composed of amino acids N43 (Asn6841), Y47 (Tyr6845), G71 (Gly6869), G81 (Gly6879), D99 (Asp6897), N101

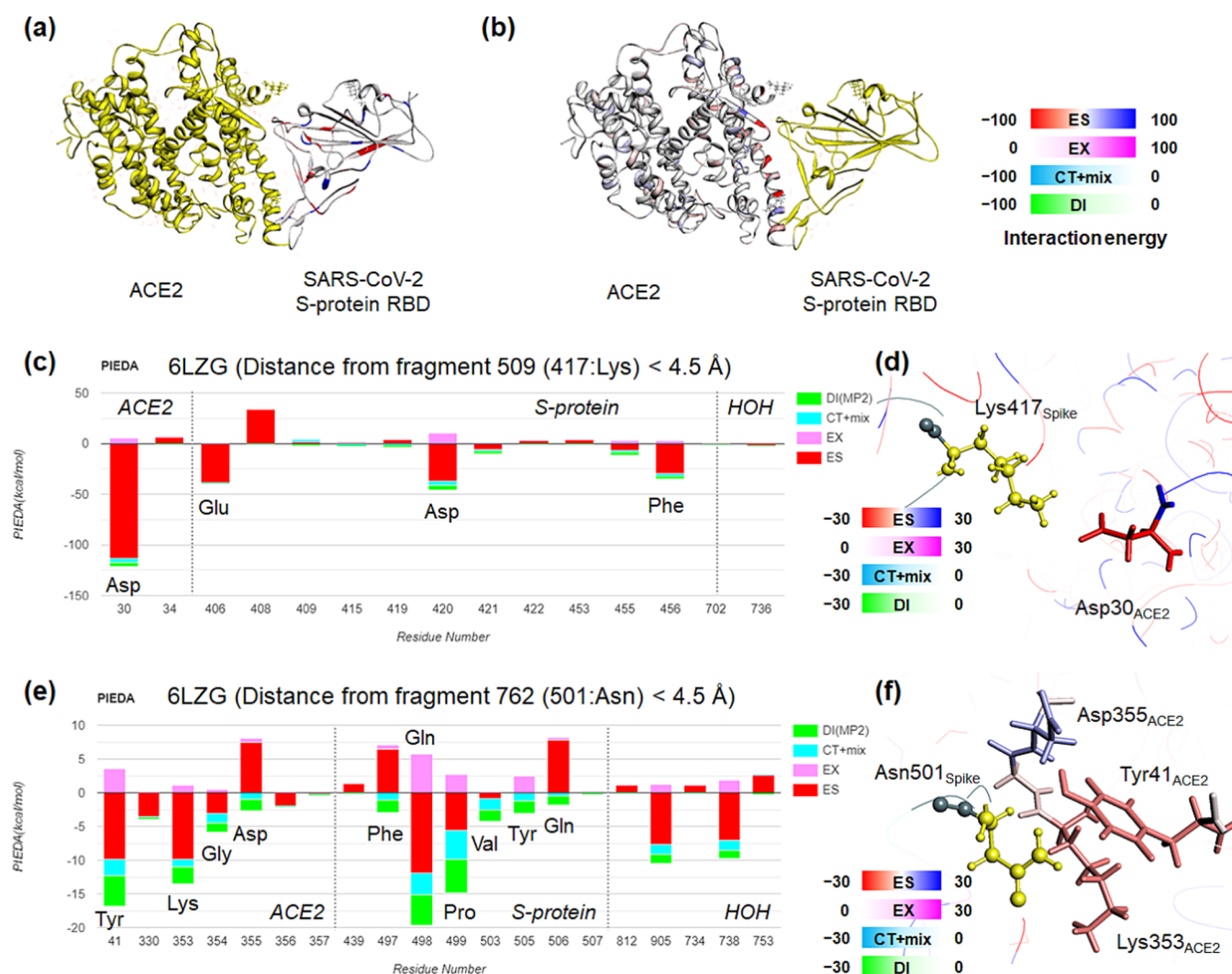


Figure 15. FMO results for PPI between the S protein RBD and ACE2 receptor (PDB ID: 6LZG, FMODB ID: 4NZVN). (a) The PIEDA main component between the ACE2 (yellow) and residue fragments of the RBD, mapped on the molecular structure of the RBD using a BioStation Viewer. (b) The PIEDA main component between the RBD (yellow) and residue fragments of ACE2, mapped on the molecular structure of ACE2. (c) Graphical representation of the PIEDA components between the Lys417 of the S protein and a residue fragment within 4.5 Å using the FMODB web interface. (d) The PIEDA main component as in (c), mapped on the molecular structure. (e) Graphical representation of the PIEDA components between Asn501 of the S protein (yellow) and a residue fragment within 4.5 Å, using the FMODB web interface. (f) The PIEDA main component as in (b), mapped on the molecular structure. Energy values are in kcal/mol.

(Asn6899), C115 (Cys6913), and D130 (Asp6928), all of which are conserved in SARS-CoV and MERS-CoV.⁷¹ The corresponding residues in Figure 14c are underlined. FMO analyses showed that the largest interactions are of the order of D99 (Asp6897), which is hydrogen-bonded to the hydroxyl group of the adenosine sugar backbone of SAM, D130 (Asp6928), and N43 (Asn6841), which are hydrogen-bonded to the amine and carboxyl groups of the methionine portion of SAM, respectively. Recently, a complex structure of the methylated CAP-like RNA and SAH was published and registered to the FMODB (PDB ID: 7LW3, FMODB ID: 2K6MR). In this 2'-O-ribose methyltransferase complex, nsp10 contains two Zn atoms, which requires a special fragmentation process around Zn. The calculation of the Zn-containing structure has been carried out in the example of the complex of nsp10–nsp16 with the candidate inhibitor simefungin (PDB ID: 6YZ1, FMODB ID: 6394Z). As shown in Figure S6, the side-chain portions of the coordinating His

and Cys residues are fragmented to be identical to Zn. We are currently working on automatic fragmentation around Zn.

3.8. Spike (S) Protein. One of the structural protein encoded in the viral genome, spike (S) glycoprotein, is a characteristic feature of coronaviruses, protrusions on the surface of the viral particle that binds to the ACE2 receptor on the host cell and allows the virus to enter the cell.^{48,49} The S protein has the highest number of published PDB entries, along with the Mpro. Cryo-EM is the main tool for the structural analysis of the whole trimer, and X-ray data are also available for the receptor-binding domain (RBD)–antibody substructure. Since FMO calculations for the whole trimer are large-scale that cannot be done easily, most of them have been done for the RBD and complexes with antibodies or ACE2 receptors. The champion data of the open/close type structure of the S protein trimer (PDB ID: 6VXX, 6VYB, FMODB ID: MMG1Z, 9YJV2) were obtained by Akisawa et al. using the Fugaku supercomputer.²² For RBD substructures, Watanabe et al. have analyzed complexes

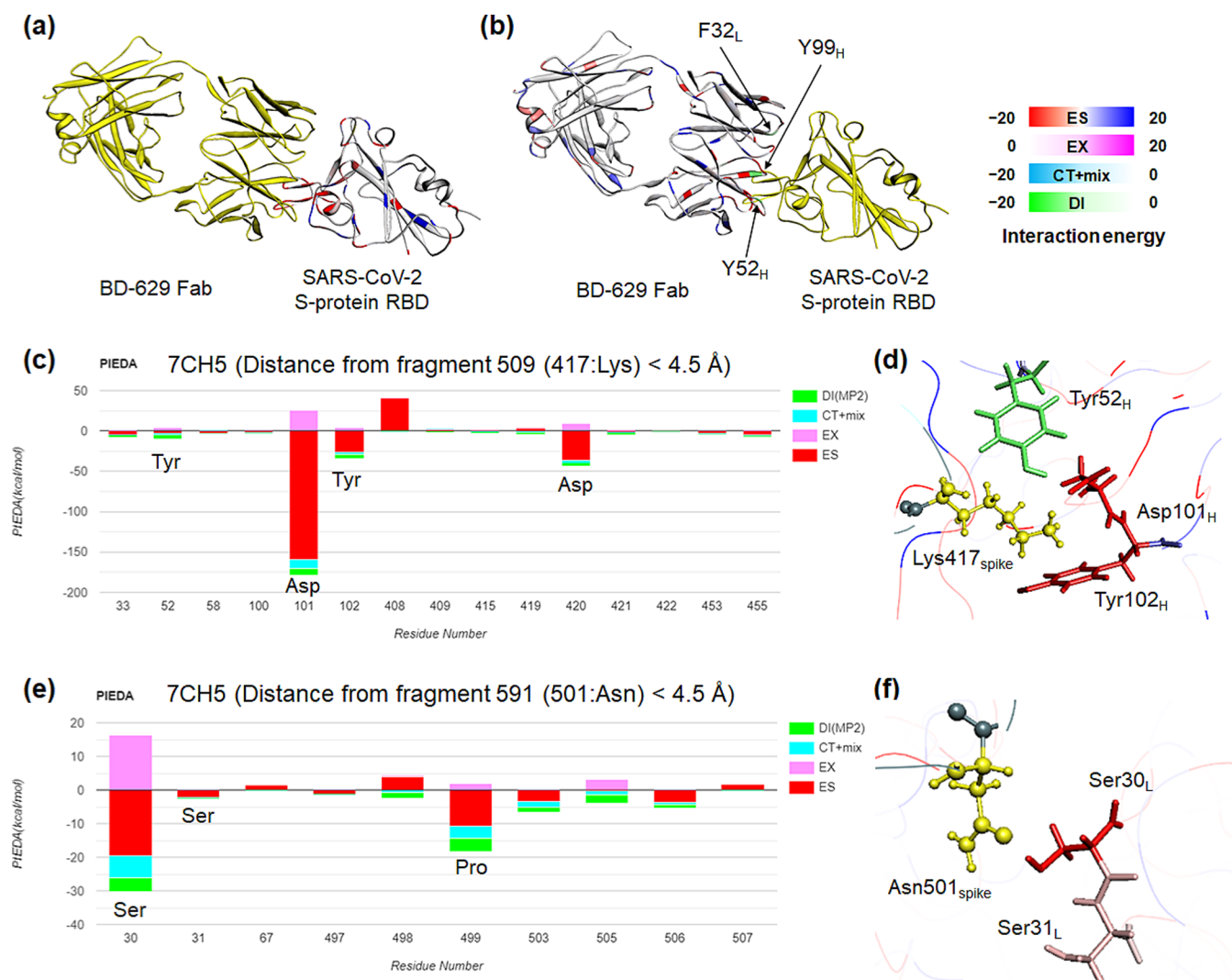


Figure 16. FMO results for PPI between the S protein RBD and neutralizing antibody (PDB ID: 7CH5, FMO DB ID: JMSM9). (a) The PIEDA main component between the BD-629 Fab (yellow) and residue fragments of RBD, mapped on the molecular structure of RBD using a BioStation Viewer. (b) The PIEDA main component between the RBD (yellow) and residue fragments of BD-629 Fab, mapped on the molecular structure of BD-629 Fab. (c) Graphical representation of the PIEDA components between the Lys417 of the S protein and a residue fragment within 4.5 Å using the FMO DB web interface. (d) The PIEDA main component as in (c), mapped on the molecular structure. (e) Graphical representation of the PIEDA components between the Asn501 of the S protein (yellow) and a residue fragment within 4.5 Å, using the FMO DB web interface. (f) PIEDA main component as in (b), mapped on the molecular structure. Energy values are in kcal/mol.

with ACE2 and compared them with SARS,²¹ and Watanabe et al. have studied complexes with antibodies, including epitope analysis²³ and mutant strains,²⁴ and all of these data have been registered in the FMO DB. Here, we quantify the protein–protein interaction (PPI) of RBD–ACE2 and RBD–neutralizing antibodies and analyzed the interaction around the mutant residues Lys417 and Asn501 in South African and Brazilian mutants reported since December 2020^{73–75} as examples of residue analyses.

An example of FMO-based PPI analyses using the RBD of the S protein and ACE2 complex (PDB ID: 6LZG, FMO DB ID: 4NZVN) is shown in Figure 15. To identify the amino acid residues important in molecular recognition, we first mapped the main PIEDA components of the interactions between the whole ACE2 and each residue of the S protein onto the molecular structure of the S protein (Figure 15a) and *vice versa* (Figure 15b). The strength of the interaction energy is indicated by a color gradation for each amino acid

residue. Since ACE2 is a $-26e$ charged protein, the charged amino acid residues throughout the S protein interact electrostatically (Figure 15a). On the other hand, the RBD of the S protein has a charge of $+2e$, indicating that the interacting amino acid residues are more concentrated at the binding interface compared to that in ACE2 (Figure 15b). Due to the large negative charge of ACE2, hydrogen bonds and XH/ π interactions, which are key interactions in molecular recognition of the S protein, are hidden by ES interactions of the charged residues. However, the hydrogen bond can be easily interpreted by analyzing the ES and CT terms of PIEDA, and the XH/ π interaction can be easily interpreted by analyzing the DI term.²¹ In addition, we performed PIEDA analysis of the mutant residues of the S protein, Lys417 and Asn501, in the South African and Brazilian mutants and the surrounding residues (Figure 15b,c). Lys417 forms a salt bridge with Asp30 of ACE2 and shows the strongest attractive interaction (-116.3 kcal/

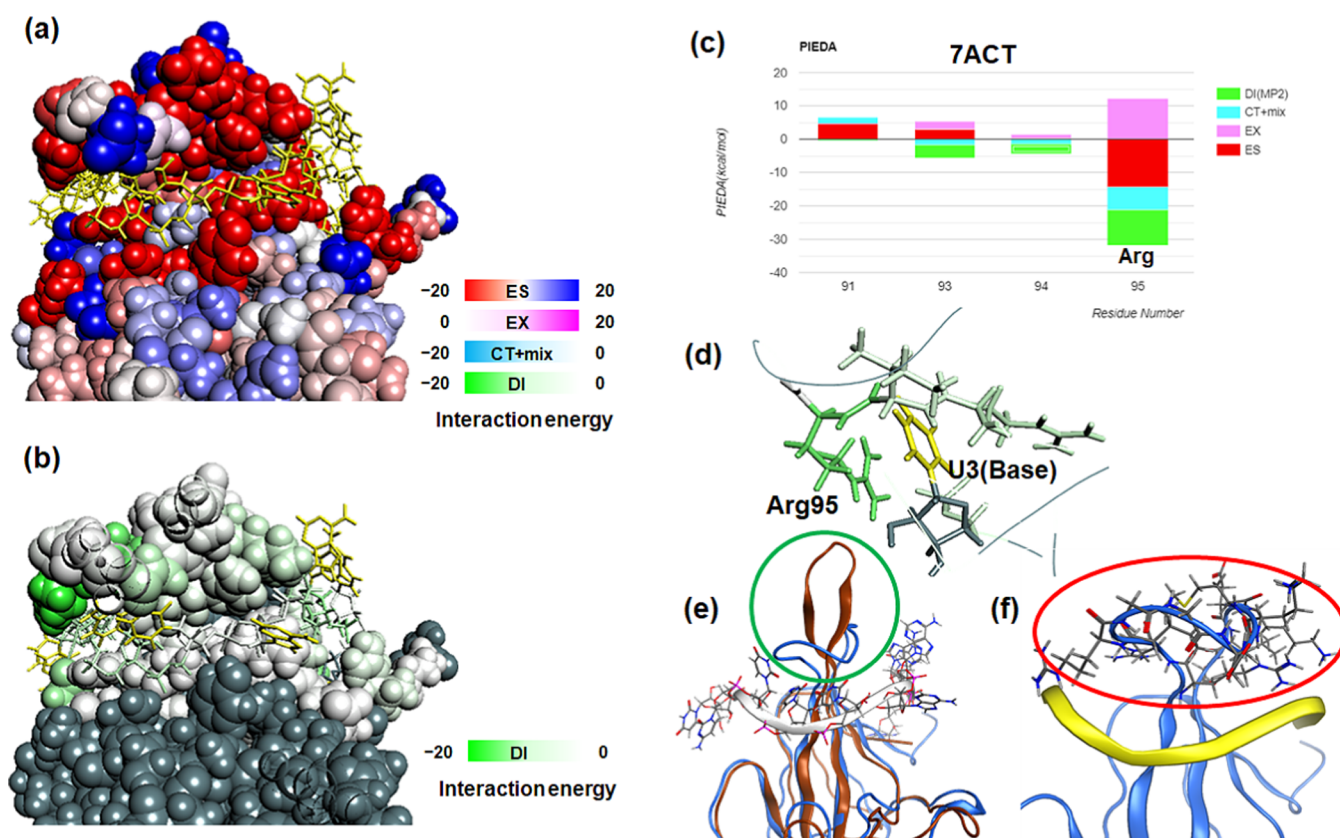


Figure 17. FMO results for the N protein and single-stranded RNA (PDBID: 7ACT, 7CDZ, FMOB ID: N3Y7Q, R12G8). Energies are in kcal/mol. (a) The PIEDA main component between the RNA (yellow) and residue fragments of N protein, mapped on the molecular structure of the N protein using a BioStation Viewer. (b) The PIEDA DI component between the RNA (yellow) and residue fragments of N protein. (c) Graphical representation of the PIEDA components between the U3 base of RNA and a residue fragment within 4.5 Å using the FMOB web interface. (d) The PIEDA main component as in (c), mapped on the molecular structure. (e) The structure of the NTD with (blue) and without (brown) RNA binding. (f) Arg92–Ser105 are framed by the red loop. Energy values are in kcal/mol.

mol) between the two proteins. On the other hand, the main contribution of Asn501 is ES with the surrounding amino acid residues (Tyr41, Lys353, Gly354, etc.). In a previous study, it was reported that the N501Y mutation enhances the ability of the S protein to bind to ACE2 by acquiring more hydrogen bonds and XH/ π interactions than wild type.²¹

As an example of RBD–neutralizing antibodies, we show the FMO results for a structure containing BD-629 Fab, which has many overlapping active sites on the RBD with the ACE2 receptor and a high activity value (IC_{50}) (PDB ID: 7CH5, FMOB ID: JM5M9).²³ First, the main components of PIEDA for the PPI of RBD and BD-629 Fab were mapped on the molecular structure (Figure 16a,b). ES is the main component of many interactions between the RBD and BD-629 Fab, whereas only Tyr52_H, Tyr99_H, and Phe32_L of BD-629 Fab interact with the RBD by DI as the main component (subscripts L and H indicate heavy and light chains of the antibody, respectively). Next, the PIEDA on a single residue of Lys417 (Figure 16c,d) showed that Asp101_H and Tyr102_H exhibit strong ES interactions. In particular, a salt bridge is formed between Lys417 and Asp101_H, and the ES term is about -150 kcal/mol. In addition, Tyr52_H forms a CH/ π interaction with the Lys417 side chain, and the DI term is the major component. On the other hand, considering Asn501 (Figure 16e,f), hydrogen bonds are formed between Asn501 and Ser30_L, indicating a strong ES interaction. These results suggest that Lys417 and Asn501, which are character-

istic of South African and Brazilian mutants, contribute significantly to the interaction with neutralizing antibodies similar to the ACE2 receptor. Details of the epitope analysis by neutralizing antibodies²³ and mutant strains²⁴ are available in the respective references.

3.9. Nucleocapsid (N) Protein. The nucleocapsid (N) protein, an structural protein, binds to viral genomic RNA to form a capsid, which constitutes the viral particle. Both the N-terminal domain (NTD) and C-terminal domain (CTD) are registered in the PDB.⁷⁶ The NTD is a major RNA-binding domain and its complex structure with single- and double-stranded RNA has been determined by solution NMR.⁷⁷ The CTD is a dimerization domain and its dimeric structure has been obtained by X-ray crystallography,^{78,79} suggesting that it also has an RNA-binding ability. The calculation results of these structures are registered in the FMOB, and representative analysis of the complex structure of the NTD and single-stranded RNA (PDB ID: 7ACT, FMOB ID: N3Y7Q) is described below. Figure 17a shows the interaction of each amino acid residue of the N protein from RNA, mapping the main components of the PIEDA on the molecular structure. We confirmed that the molecular recognition of RNA is mainly due to ES interactions. On the other hand, Figure 17b shows the DI interaction between the protein and only the base parts of RNA, indicating that DI plays an important role in recognizing the bases. For example, a large DI is observed in the interaction between the U3 base

and Arg95, indicating π - π stacking between the pyrimidine of U3 and guanidyl group of Arg95 (Figure 17c,d). Figure 17e shows the structure of the NTD with and without RNA binding (PDB ID: 7CDZ, FMO DB ID: R12G8). As shown in Figure 17f, the interaction energy between the red-framed loop (Arg92–Ser105) and RNA is -1191.5 kcal/mol, which accounts for 53.8% of the total interaction. This indicates that the structural change is important in the binding of the N protein to RNA.

3.10. Other SARS-CoV-2. Structures other than the above nine are currently classified as other SARS-CoV-2. Although several cryo-EM structures of nsp1, which suppress the innate immune function of the host, in a huge complex with human 40S ribosomal subunit have been published in the PDB,^{80,81} the FMO DB currently has only the calculated X-ray crystal structure of nsp1 alone (PDB ID 7K7P, FMO DB ID: 5965Z),⁸² and the complex structure is under calculation. The structures of nsp2 (PDB ID: 7MSW) and nsp14 (PDB ID: 7DIY), which were published in the PDB in May 2021, are currently under investigation. For structural protein, a five-helix bundle structure of the envelope (E) protein⁸³ (PDB ID: 7K3G, FMO DB ID: 4NZYN) and a complex structure of the E protein C-terminal DRLV motif and human cell junction protein PALS1⁸⁴ (PDB ID: 7M4R, FMO DB ID: Z2JJN) have been registered in the FMO DB. For open reading form proteins, the FMO calculation results for ORF3a (PDB ID: 6XDC), ORF7a (PDB ID: 7C13), ORF8 (PDB ID: 7JX6), OLF9b (PDB ID: 6Z4U), and the ORF9b-human Tom70 (Mitochondrial import receptor subunit) complex (PDB ID: 7KDT)⁸⁵ have been registered in the FMO DB.

4. CONCLUSIONS

In this study, we present an overview of the COVID-19 special feature in FMO DB, describing the contents of the registered data and analysis methods for each type of SARS-CoV-2-related protein. We have collected data mainly on the structures classified as representative entries in the COVID-19 special page of the PDBj but it is difficult to completely cover them due to some technical issues. First, regarding the size of the calculated protein, complex structures of up to 1000 residues are mainly registered in the FMO DB. Among the FMO calculations performed by ABINIT-MP so far, the largest-sized protein is the trimer of S protein (3300 residues),²² of both open and closed structures. In the near future, we shall perform FMO calculations for large complex systems, such as RTCs of RdRp, including helicase, nsp1–ribosomal subunit, and spike trimer–antibody, and elucidate the molecular recognition mechanism in these complex systems. Second, for S glycoproteins containing carbohydrate chains, we have calculated structural models with and without carbohydrate modification and registered both of them in the FMO DB (e.g., PDB ID: 6LZG, FMO DB ID: 4NZVN and YQG92). Furthermore, for the NMR structure, the PDB contains dozens of MODEL structures due to the fluctuating structure in solution. We would like to register all of these MODELS as a standard of FMO DB in the future. The fourth issue is the handling of Zn atoms: for the Zn-containing proteins PLpro, RdRp, helicase, nsp10, and ACE2, only a few of the complex structures have been currently calculated in the presence of Zn because of the large amount of manual work involved in fragmentation (see Section 3.7 for details). Since Zn is located far from the active site in these proteins,

its presence or absence would not affect important interactions. Automatic fragmentation around Zn is currently under development and will be recalculated and registered in the FMO DB as soon as it is available within this year. We shall address these technical issues in the activities of the FMO drug design consortium (FMO DB).^{86,87} The last issue is the expansion of the FMO programs that are eligible for FMO DB registration. Although the current FMO DB includes only results calculated using the ABINIT-MP program,^{32,33} some groups have performed FMO calculations for COVID-19-related proteins, such as Mpro^{29,30} and S proteins,²⁵ using the GAMESS program.^{34–36} We currently improve FMO DB so that the results of FMO calculations can be registered in GAMESS.

Experimental structural studies of SARS-CoV-2 by structural biologists worldwide and their publications in the PDB are a valuable source of information for the fight against COVID-19. The interaction data published in the FMO DB enhance our understanding of SARS-CoV-2 by adding quantitative interaction to structural information. Clarifying structural features of therapeutic agents important to protein interactions would not only help in the rational design of therapeutic agents but also contribute to the epitope analysis of mutant strains, which is currently the focus of global attention, and the identification of interactions leading to vaccine development. We shall continuously update the FMO DB by performing FMO calculations on newly released important PDB structures, which we hope will help overcome COVID-19.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.1c00694>.

Structures of SARS-CoV-2-related proteins are displayed in Figure S1. Interaction energies between the whole ligand are divided into multiple fragments and the surrounding residues (Figures S2 and S5). The substrate-binding pocket of the main protease and the structure of the 110 compounds shown in Figure 7b are displayed for each ligand cluster classified by SOM (Figure S3). The full IFIE map between all fragments that consist of the complex structure of RdRp–nsp7–nsp8–RNA–remdesivir (Figure S4). The fragmentation method of the four-coordinated Zn atom and the surrounding Cys and His residues (Figure S6) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Kaori Fukuzawa – Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, Tokyo 142-8501, Japan; Department of Biomolecular Engineering, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan; orcid.org/0000-0001-5357-8250; Email: k-fukuzawa@hoshi.ac.jp

Koichiro Kato – Department of Applied Chemistry, Graduate School of Engineering, Kyushu University, Fukuoka 819-0395, Japan; Center for Molecular Systems (CMS), Kyushu University, Fukuoka 819-0395, Japan; orcid.org/0000-0003-4392-8741; Email: kato.koichiro.957@m.kyushu-u.ac.jp

Chiduru Watanabe – RIKEN Center for Biosystems Dynamics Research, Yokohama, Kanagawa 230-0045, Japan; JST PRESTO, Kawaguchi, Saitama 332-0012, Japan; orcid.org/0000-0002-0742-3896; Email: chiduru.watanabe@riken.jp

Teruki Honma – RIKEN Center for Biosystems Dynamics Research, Yokohama, Kanagawa 230-0045, Japan; orcid.org/0000-0003-3761-9504; Email: honma.teruki@riken.jp

Authors

Yusuke Kawashima – Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, Tokyo 142-8501, Japan; orcid.org/0000-0001-8465-6848

Yuma Handa – Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, Tokyo 142-8501, Japan; orcid.org/0000-0001-8660-8487

Ami Yamamoto – Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, Tokyo 142-8501, Japan

Kazuki Watanabe – Graduate School of Pharmaceutical Sciences, Osaka University, Suita, Osaka 565-0871, Japan; Graduate School of Pharmaceutical Sciences, Chiba University, Chiba 260-8675, Japan; orcid.org/0000-0002-8119-4446

Tatsuya Ohyama – RIKEN Center for Biosystems Dynamics Research, Yokohama, Kanagawa 230-0045, Japan; Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, Kobe 650-0047, Japan; orcid.org/0000-0002-8375-2608

Kikuko Kamisaka – RIKEN Center for Biosystems Dynamics Research, Yokohama, Kanagawa 230-0045, Japan; orcid.org/0000-0002-1215-2525

Daisuke Takaya – RIKEN Center for Biosystems Dynamics Research, Yokohama, Kanagawa 230-0045, Japan; orcid.org/0000-0001-6020-8151

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jcim.1c00694>

Notes

The authors declare no competing financial interest. All structure files and a set of input/output files used for FMO calculations are available at the FMO DB (<https://drugdesign.riken.jp/FMODB/>). Simple data analysis can be performed using the FMO DB web interface, and detailed analysis can be performed using the BioStation Viewer software (<https://fmodd.jp/biostationviewer-dl/>). FMO Software ABINIT-MP has been pre-installed on the high-performance computing infrastructure (HPCI) machines (https://www.hpci-office.jp/pages/e_appli_abinit-mp).

ACKNOWLEDGMENTS

The authors thank Prof. Yuji Mochizuki and Dr. Koji Okuwaki of Rikkyo University, Prof. Shigenori Tanaka of Kobe University, Dr. Yuto Komeiji of the National Institute of Advanced Industrial Science and Technology, and Dr. Tatsuya Nakano and Dr. Yoshio Okiyama of the National Institute of Health Sciences (NIHS) for general discussions related to FMO calculations. In particular, Prof. Mochizuki provided us the Oakforest-PACS version of ABINIT-MP and the calculation data of the spike protein trimer. The authors

are grateful to Prof. Tatsuya Takagi of the Osaka University for discussing the PPI of S protein, Dr. Kazumi Tsuda and Dr. Shu Koyama of Science & Technology Systems, Inc. for their help in the development of the FMO DB, and Dr. Junichi Ishida of Mizuho Research & Technologies, Ltd. for his help in some of the calculations. The authors would like to thank Enago for the English language review. 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 Research and Development (AMED) (Grant Number JP21am0101113). C.W. acknowledges JST, PRESTO grant (JPMJPR18GD). C.W. and D.T. acknowledge JSPS KAKENHI (Grant Number 18K06619). This research was supported by the RIKEN programs for Drug Discovery and Medical Technology Platforms (DMP) and COVID-19 project. This research was performed under the activities of the FMO drug design consortium (FMODD). The FMO calculations were partially performed using the Oakforest-PACS supercomputer (project ID: hp200101), Fugaku supercomputer (project ID: hp210130), the HOKUSAI supercomputer (RIKEN Advanced Center for Computing and Communications, Saitama, Japan, project ID: Q20306), and the TSUBAME3.0 supercomputer (Tokyo Institute of Technology, Tokyo, Japan). A part of PIEDA calculations was performed with the MIZUHO/BioStation software package (version 4.0).

REFERENCES

- (1) WHO Coronavirus Disease (COVID-19) Situation Reports. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/> (accessed June 8, 2021).
- (2) Burley, S. K.; Berman, H. M.; Bhikadiya, C.; Bi, C.; Chen, L.; Costanzo, L. D.; Christie, C.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ghosh, S.; Goodsell, D. S.; Green, R. K.; Guranovic, V.; Guzenko, D.; Hudson, B. P.; Liang, Y.; Lowe, R.; Peisach, E.; Periskova, I.; Randle, C.; Rose, A.; Sekharan, M.; Shao, C.; Tao, Y.-P.; Valasatava, Y.; Voigt, M.; Westbrook, J.; Young, J.; Zardecki, C.; Zhuravleva, M.; Kurisu, G.; Nakamura, H.; Kengaku, Y.; Cho, H.; Sato, J.; Kim, J. Y.; Ikegawa, Y.; Nakagawa, A.; Yamashita, R.; Kudou, T.; Bekker, G.-J.; Suzuki, H.; Iwata, T.; Yokochi, M.; Kobayashi, N.; Fujiwara, T.; Velankar, S.; Kleywegt, G. J.; Anyango, S.; Armstrong, D. R.; Berrisford, J. M.; Conroy, M. J.; Dana, J. M.; Deshpande, M.; Gane, P.; Gáborová, R.; Gupta, D.; Gutmanas, A.; Koča, J.; Mak, L.; Mir, S.; Mukhopadhyay, A.; Nadzirin, N.; Nair, S.; Patwardhan, A.; Paysan-Lafosse, T.; Pravda, L.; Salih, O.; Sehna, D.; Varadi, M.; Vařeková, R.; Markley, J. L.; Hoch, J. C.; Romero, P. R.; Baskaran, K.; Maziuk, D.; Ulrich, E. L.; Wedell, J. R.; Yao, H.; Livny, M.; Ioannidis, Y. E. Protein Data Bank: The Single Global Archive for 3D Macromolecular Structure Data. *Nucleic Acids Res.* **2019**, *47*, D520–D528.
- (3) Waman, V. P.; Sen, N.; Varadi, M.; Daina, A.; Wodak, S. J.; Zoete, V.; Velankar, S.; Orengo, C. The Impact of Structural Bioinformatics Tools and Resources on SARS-Cov-2 Research and Therapeutic Strategies. *Briefings Bioinf.* **2021**, *22*, 742–768.
- (4) Kinjo, A. R.; Bekker, G.-J.; Suzuki, H.; Tsuchiya, Y.; Kawabata, T.; Ikegawa, Y.; Nakamura, H. Protein Data Bank Japan (PDBj): Updated User Interfaces, Resource Description Framework, Analysis Tools for Large Structures. *Nucleic Acids Res.* **2017**, *45*, D282–D288.
- (5) Kinjo, A. R.; Bekker, G.-J.; Wako, H.; Endo, S.; Tsuchiya, Y.; Sato, H.; Nishi, H.; Kinoshita, K.; Suzuki, H.; Kawabata, T.; Yokochi, M.; Iwata, T.; Kobayashi, N.; Fujiwara, T.; Kurisu, G.; Nakamura, H. New Tools and Functions in Data-Only Activities at Protein Data Bank Japan (PDBj). *Protein Sci.* **2018**, *27*, 95–102.

- (6) Komatsu, T. S.; Okimoto, N.; Koyama, Y. M.; Hirano, Y.; Morimoto, G.; Ohno, Y.; Taiji, M. Drug Binding Dynamics of the Dimeric SARS-Cov-2 Main Protease, Determined by Molecular Dynamics Simulation. *Sci. Rep.* **2020**, *10*, No. 16986.
- (7) Molecular Dynamics Simulations Related to SARS-CoV-2. http://www.deshawresearch.com/resources_sarscov2.html (accessed June 8, 2021).
- (8) Mori, T.; Jung, J.; Kobayashi, C.; Dokainish, H. M.; Re, S.; Sugita, Y. Elucidation of Interactions Regulating Conformational Stability and Dynamics of SARS-CoV-2 S-protein. *Biophys. J.* **2021**, *120*, 1060–1071.
- (9) Shaw, D. E.; Grossman, J. P.; Bank, J. A.; Batson, B.; Butts, J. A.; Chao, J. C.; Deneroff, M. M.; Dror, R. O.; Even, A.; Fenton, C. H.; Forte, A.; Gagliardo, J.; Gill, G.; Greskamp, B.; Ho, C. R.; Ierardi, D. J.; Iserovich, L.; Kuskin, J. S.; Larson, R. H.; Layman, T.; Lee, L.-S.; Lerer, A. K.; Li, C.; Killebrew, D.; Mackenzie, K. M.; Mok, S. Y.-H.; Moraes, M. A.; Mueller, R.; Nociolo, L. J.; Peticolas, J. L.; Quan, T.; Ramot, D.; Salmon, J. K.; Scarpazza, D. P.; Schafer, U. B.; Siddique, N.; Snyder, C. W.; Spengler, J.; Tang, P. T. P.; Theobald, M.; Toma, H.; Towles, B.; Vitale, B.; Wang, S. C.; Young, C. In *Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer*, SC'14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2014; pp 41–53.
- (10) Komatsu, T. S.; Ohno, Y.; Morimoto, G.; Zhang, H.; Koyama, Y. M.; Taiji, M. MDGRAPE-4A. Japanese Patent JP2020-64560A.
- (11) GENESIS-2.0. <https://github.com/genesis-release-r-ccs/genesis-2.0> (accessed June 8, 2021).
- (12) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (13) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism From Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (14) 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*, 1722–1734.
- (15) 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*, 4785–4790.
- (16) Nguyen, H. L.; Thai, N. Q.; Truong, D. T.; Li, M. S. Remdesivir Strongly Binds to Both RNA-Dependent RNA Polymerase and Main Protease of SARS-CoV-2: Evidence from Molecular Simulations. *J. Phys. Chem. B* **2020**, *124*, 11337–11348.
- (17) Koulgi, S.; Jani, V.; Uppuladinne, M. V. N.; Sonavane, U.; Joshi, R. Remdesivir-Bound and Ligand-Free Simulations Reveal the Probable Mechanism of Inhibiting the RNA Dependent RNA Polymerase of Severe Acute Respiratory Syndrome Coronavirus 2. *RSC Adv.* **2020**, *10*, 26792–26803.
- (18) Zhang, L.; Zhou, R. Structural Basis of the Potential Binding Mechanism of Remdesivir to SARS-CoV-2 RNA-Dependent RNA Polymerase. *J. Phys. Chem. B* **2020**, *124*, 6955–6962.
- (19) Kitaura; et al. Fragment Molecular Orbital Method: an Approximate Computational Method for Large Molecules. *Chem. Phys. Lett.* **1999**, *313*, 701–706.
- (20) *Recent Advances of the Fragment Molecular Orbital Method*; In Mochizuki, Y.; Tanaka, S.; Fukuzawa, K., Eds.; Springer: Singapore, 2021.
- (21) Watanabe, C.; Okiyama, Y.; Tanaka, S.; Fukuzawa, K.; Honma, T. Molecular Recognition of SARS-Cov-2 Spike Glycoprotein: Quantum Chemical Hot Spot and Epitope Analyses. *Chem. Sci.* **2021**, *12*, 4722–4739.
- (22) Akisawa, K.; Hatada, R.; Okuwaki, K.; Mochizuki, Y.; Fukuzawa, K.; Komeiji, Y.; Tanaka, S. Interaction Analyses Of SARS-Cov-2 Spike Protein Based on Fragment Molecular Orbital Calculations. *RSC Adv.* **2021**, *11*, 3272–3279.
- (23) Watanabe, K.; Watanabe, C.; Honma, T.; Tian, Y.-S.; Kawashima, Y.; Kawashita, N.; Takagi, T.; Fukuzawa, K. Intermolecular Interaction Analyses on SARS-CoV-2 Spike Protein Receptor Binding Domain and Human Angiotensin-Converting Enzyme 2 Receptor-Blocking Antibody/Peptide Using Fragment Molecular Orbital Calculation. *J. Phys. Chem. Lett.* **2021**, *12*, 4059–4066.
- (24) Watanabe, K.; Honma, T.; Tian, Y.-S.; Kawashima, Y.; et al. Computational Ab Initio Interaction Analyses Between Neutralizing Antibody and SARS-Cov-2 Variant Spike Proteins Using the Fragment Molecular Orbital Method. *Bull. Chem. Soc. Jpn.* **2021**, *94*, 1794–1798.
- (25) Lim, H.; Baek, A.; Kim, J.; Kim, M. S.; Liu, J.; Nam, K. Y.; Yoon, J.; No, K. T. Hot Spot Profiles of SARS-Cov-2 and Human ACE2 Receptor Protein Interaction Obtained by Density Functional Tight Binding Fragment Molecular Orbital Method. *Sci. Rep.* **2020**, *10*, No. 16862.
- (26) Hatada, R.; Okuwaki, K.; Mochizuki, Y.; Handa, Y.; Fukuzawa, K.; Komeiji, Y.; Okiyama, Y.; Tanaka, S. Fragment Molecular Orbital Based Interaction Analyses on COVID-19 Main Protease - Inhibitor N3 Complex (PDB ID: 6LU7). *J. Chem. Inf. Model.* **2020**, *60*, 3593–3602.
- (27) Hatada, R.; Okuwaki, K.; Akisawa, K.; Mochizuki, Y.; Handa, Y.; Fukuzawa, K.; Komeiji, Y.; Okiyama, Y.; Tanaka, S. Statistical Interaction Analyses Between SARS-Cov-2 Main Protease and Inhibitor N3 By Combining Molecular Dynamics Simulation and Fragment Molecular Orbital Calculation. *Appl. Phys. Express* **2021**, *14*, No. 027003.
- (28) Tanaka, S.; Tokutomi, S.; Hatada, R.; Okuwaki, K.; Akisawa, K.; Fukuzawa, K.; Komeiji, Y.; Okiyama, Y.; Mochizuki, Y. Dynamic Cooperativity of Ligand-Residue Interactions Evaluated with the Fragment Molecular Orbital Method. *J. Phys. Chem. B* **2021**, *125*, 6501–6512.
- (29) Nutho, B.; Mahalapbutr, P.; Hengphasatporn, K.; Pattarangoon, N. C.; Simanon, N.; Shigeta, Y.; Hannongbua, S.; Rungrotmongkol, T. Why Are Lopinavir and Ritonavir Effective against the Newly Emerged Coronavirus 2019? Atomistic Insights into the Inhibitory Mechanisms. *Biochemistry* **2020**, *59*, 1769–1779.
- (30) Deetanya, P.; Hengphasatporn, K.; Wilasluck, P.; Shigeta, Y.; Rungrotmongkol, T.; Wangkanont, K. Interaction of 8-Anilino-naphthalene-1-Sulfonate with SARS-CoV-2 Main Protease and Its Application as a Fluorescent Probe for Inhibitor Identification. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 3364–3371.
- (31) 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. *J. Mol. Graphics Model.* **2020**, *100*, No. 107695.
- (32) Mochizuki, Y.; Nakano, T.; Sakakura, K.; Okiyama, Y.; Watanabe, H.; Kato, K.; Akinaga, Y.; Sato, S.; Yamamoto, J.-i.; Yamashita, K.; Murase, T.; Ishikawa, T.; Komeiji, Y.; Kato, Y.; Watanabe, N.; Tsukamoto, T.; Mori, H.; Okuwaki, K.; Tanaka, S.; Kato, A.; Watanabe, C.; Fukuzawa, K. The ABINIT-MP Program. In *Recent Advances of the Fragment Molecular Orbital Method*; Mochizuki, Y.; Tanaka, S.; Fukuzawa, K., Eds.; Springer: Singapore, 2021; pp 53–67.
- (33) 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*, 10310–10344.
- (34) 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.; Humana Press: New York, NY, 2020; Vol. 2114, pp 37–48.
- (35) Fedorov, D. G. The Fragment Molecular Orbital Method: Theoretical Development, Implementation in GAMESS, and Applications. *WIREs Comput. Mol. Sci.* **2017**, *7*, No. e1322.

- (36) Fedorov, D. G. Recent Development of the Fragment Molecular Orbital Method in GAMESS. In *Recent Advances of the Fragment Molecular Orbital Method*; Mochizuki, Y.; Tanaka, S.; Fukuzawa, K., Eds.; Springer: Singapore, 2021; pp 31–51.
- (37) Morao, I.; Heifetz, A.; Fedorov, D. G. Accurate Scoring in Seconds with the Fragment Molecular Orbital and Density-Functional Tight-Binding Methods. In *Quantum Mechanics in Drug Discovery*; Heifetz, A., Ed.; Humana Press: New York, NY, 2020; Vol. 2114, pp 143–148.
- (38) Acharya, A.; Agarwal, R.; Baker, M. B.; Baudry, J.; Bhowmik, D.; Boehm, S.; Byler, K. G.; Chen, S. Y.; Coates, L.; Cooper, C. J.; Demerdash, O.; Daidone, I.; Eblen, J. D.; Ellingson, S.; Forli, S.; Glaser, J.; Gumbart, J. C.; Gunnels, J.; Hernandez, O.; Irlle, S.; Kneller, D. W.; Kovalevsky, A.; Larkin, J.; Lawrence, T. J.; LeGrand, S.; Liu, S. H.; Mitchell, J. C.; Park, G.; Parks, J. M.; Pavlova, A.; Petridis, L.; Poole, D.; Pouchard, L.; Ramanathan, A.; Rogers, D. M.; Santos-Martins, D.; Scheinberg, A.; Sedova, A.; Shen, Y.; Smith, J. C.; Smith, M. D.; Soto, C.; Tsaris, A.; Thavappiragasam, M.; Tillack, A. F.; Vermaas, J. V.; Vuong, V. Q.; Yin, J.; Yoo, S.; Zahran, M.; Zanetti-Polzi, L. Supercomputer-Based Ensemble Docking Drug Discovery Pipeline with Application to Covid-19. *J. Chem. Inf. Model.* **2020**, *60*, 5832–5852.
- (39) Takaya, D.; Watanabe, C.; Nagase, S.; Kamisaka, K.; Okiyama, Y.; Moriwaki, H.; Yuki, H.; Sato, T.; Kurita, N.; Yagi, Y.; Takagi, T.; Kawashita, N.; Takaba, K.; Ozawa, T.; Takimoto-Kamimura, M.; Tanaka, S.; Fukuzawa, K.; Honma, T. FMO: The World's First Database of Quantum Mechanical Calculations for Biomacromolecules Based on the Fragment Molecular Orbital Method. *J. Chem. Inf. Model.* **2021**, *61*, 777–794.
- (40) Watanabe, C.; Watanabe, H.; Okiyama, Y.; Takaya, D.; Fukuzawa, K.; Tanaka, S.; Honma, T. Development of an Automated Fragment Molecular Orbital (FMO) Calculation Protocol Toward Construction of Quantum Mechanical Calculation Database for Large Biomolecules. *Chem-Bio Inf. J.* **2019**, *19*, 5–18.
- (41) BioStation Viewer, Open 1.0 rev. 2. <https://fmod.jp/biostationviewer-dl/> (accessed June 8, 2021).
- (42) *Molecular Operating Environment (MOE), 2019.01*; Chemical Computing Group ULC: Montreal, QC, Canada, 2021.
- (43) Fukuzawa, K.; Watanabe, C.; Okiyama, Y.; Nakano, T. How to Perform FMO Calculation in Drug Discovery. In *Recent Advances of the Fragment Molecular Orbital Method*; Mochizuki, Y.; Tanaka, S.; Fukuzawa, K., Eds.; Springer: Singapore, 2021; pp 93–125.
- (44) Mochizuki, Y.; Nakano, T.; Koikegami, S.; Tanimori, S.; Abe, Y.; Nagashima, U.; Kitaura, K. A Parallelized Integral-Direct Second-Order Moller-Plesset Perturbation Theory Method with a Fragment Molecular Orbital Scheme. *Theor. Chem. Acc.* **2004**, *112*, 442–452.
- (45) Mochizuki, Y.; Koikegami, S.; Nakano, T.; Amari, S.; Kitaura, K. Large Scale MP2 Calculations with Fragment Molecular Orbital Scheme. *Chem. Phys. Lett.* **2004**, *396*, 473–479.
- (46) Fedorov, D. G.; Kitaura, K. Pair Interaction Energy Decomposition Analysis. *J. Comput. Chem.* **2007**, *28*, 222–237.
- (47) 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*, 1–9.
- (48) Barnes, C. O.; West, A. P.; Huey-Tubman, K. E.; Hoffmann, M. A. G.; Sharaf, N. G.; Hoffman, P. R.; Koranda, N.; Gristick, H. B.; Gaebler, C.; Muecksch, F.; Lorenzi, J. C. C.; Finkin, S.; Häggelöf, T.; Hurley, A.; Millard, K. G.; Weisblum, Y.; Schmidt, F.; Hatziioannou, T.; Bieniasz, P. D.; Caskey, M.; Robbiani, D. F.; Nussenzweig, M. C.; Bjorkman, P. J. Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **2020**, *182*, 828–842.
- (49) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281–292.
- (50) Hilgenfeld, R. From SARS to MERS: Crystallographic Studies on Coronaviral Proteases Enable Antiviral Drug Design. *FEBS J.* **2014**, *281*, 4085–4096.
- (51) Yan, L.; Ge, J.; Zheng, L.; Zhang, Y.; Gao, Y.; Wang, T.; Huang, Y.; Yang, Y.; Gao, S.; Li, M.; Liu, Z.; Wang, H.; Li, Y.; Chen, Y.; Guddat, L. W.; Wang, Q.; Rao, Z.; Lou, Z. Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis. *Cell* **2021**, *184*, 184–193.
- (52) Lim, K. P.; Ng, L. F. P.; Liu, D. X. Identification of a Novel Cleavage Activity of the First Papain-Like Proteinase Domain Encoded by Open Reading Frame 1a of the Coronavirus Avian Infectious Bronchitis Virus and Characterization of the Cleavage Products. *J. Virol.* **2000**, *74*, 1674–1685.
- (53) Harcourt, B. H.; Jukneliene, D.; Kanjanahaluethai, A.; Bechill, J.; Severson, K. M.; Smith, C. M.; Rota, P. A.; Baker, S. C. Identification of Severe Acute Respiratory Syndrome Coronavirus Replicase Products and Characterization of Papain-Like Protease Activity. *J. Virol.* **2004**, *78*, 13600–13612.
- (54) Gao, X.; Qin, B.; Chen, P.; Zhu, K.; Hou, P.; Wojdyla, J. A.; Wang, M.; Cui, S. Crystal Structure of SARS-CoV-2 Papain-like Protease. *Acta Pharm. Sin. B* **2021**, *11*, 237–245.
- (55) Frieman, M.; Ratia, K.; Johnston, R. E.; Mesecar, A. D.; Baric, R. S. Severe Acute Respiratory Syndrome Coronavirus Papain-Like Protease Ubiquitin-Like Domain and Catalytic Domain Regulate Antagonism of IRF3 and NF- κ B Signaling. *J. Virol.* **2009**, *83*, 6689–6705.
- (56) Shin, D.; Mukherjee, R.; Grewe, D.; Bojkova, D.; Baek, K.; Bhattacharya, A.; Schulz, L.; Widera, M.; Mehdipour, A. R.; Tascher, G.; Geurink, P. P.; Wilhelm, A.; van der Heden van Noort, G. J.; Ova, H.; Muller, S.; Knobeloch, K. P.; Rajalingam, K.; Schulman, B. A.; Cinatl, J.; Hummer, G.; Ciesek, S.; Dikic, I. Papain-Like Protease Regulates SARS-Cov-2 Viral Spread and Innate Immunity. *Nature* **2020**, *587*, 657–662.
- (57) Rut, W.; Lv, Z.; Zmudzinski, M.; Patchett, S.; Nayak, D.; Snipas, S. J.; El Oualid, F.; Huang, T. T.; Bekes, M.; Drag, M.; Olsen, S. K. Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework for anti-COVID-19 drug design. *Sci. Adv.* **2020**, *6*, No. eabd4596.
- (58) Michalska, K.; Kim, Y.; Jdrzejczak, R.; Maltseva, N. I.; Stols, L.; Endres, M.; Joachimiak, A. Crystal Structures of SARS-Cov-2 ADP-Ribose Phosphatase: from the Apo Form to Ligand Complexes. *IUCrJ* **2020**, *7*, 814–824.
- (59) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of M(pro) from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* **2020**, *582*, 289–293.
- (60) Dai, W.; Zhang, B.; Jiang, X.-M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L.-K.; Xu, Y.; Yang, H.; Liu, H. Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-Cov-2 Main Protease. *Science* **2020**, *368*, 1331–1335.
- (61) Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J.; Hilgenfeld, R.; Yuen, K. Y.; Wong, L.; Gao, G.; Chen, S.; Chen, Z.; Ma, D.; Bartlam, M.; Rao, Z. Design of Wide-Spectrum Inhibitors Targeting Coronavirus Main Proteases. *PLoS Biol.* **2005**, *3*, No. e324.
- (62) Kurauchi, R.; Watanabe, C.; Fukuzawa, K.; Tanaka, S. Novel Type of Virtual Ligand Screening on the Basis of Quantum-Chemical Calculations for Protein-Ligand Complexes and Extended Clustering Techniques. *Comput. Theor. Chem.* **2015**, *1061*, 12–22.
- (63) Ma, C.; Sacco, M. D.; Hurst, B.; Townsend, J. A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M. T.; Chen, Y.; Wang, J. Boceprevir, GC-376, and Calpain Inhibitors II, XII Inhibit SARS-CoV-2 Viral Replication by Targeting the Viral Main Protease. *Cell Res.* **2020**, *30*, 678–692.
- (64) Hillen, H. S.; Kokic, G.; Farnung, L.; Dienemann, C.; Tegunov, D.; Cramer, P. Structure of Replicating SARS-CoV-2 Polymerase. *Nature* **2020**, *584*, 154–156.

- (65) Yin, W.; Mao, C.; Luan, X.; Shen, D. D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; Chang, S.; Xie, Y. C.; Tian, G.; Jiang, H. W.; Tao, S. C.; Shen, J.; Jiang, Y.; Jiang, H.; Xu, Y.; Zhang, S.; Zhang, Y.; Xu, H. E. Structural Basis for Inhibition of the RNA-Dependent RNA Polymerase from SARS-CoV-2 by Remdesivir. *Science* **2020**, *368*, 1499–1504.
- (66) Chen, J.; Malone, B.; Llewellyn, E.; Grasso, M.; Shelton, P. M. M.; Olinares, P. D. B.; Maruthi, K.; Eng, E. T.; Vatandaslar, H.; Chait, B. T.; Kapoor, T. M.; Darst, S. A.; Campbell, E. A. Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell* **2020**, *182*, 1560–1573.
- (67) Nudler, E. RNA Polymerase Backtracking in Gene Regulation and Genome Instability. *Cell* **2012**, *149*, 1438–1445.
- (68) Kaplan, D. L.; O'Donnell, M. DnaB Drives DNA Branch Migration and Dislodges Proteins While Encircling Two DNA Strands. *Mol. Cell* **2002**, *10*, 647–657.
- (69) Hackbart, M.; Deng, X.; Baker, S. C. Coronavirus Endoribonuclease Targets Viral Polyuridine Sequences to Evade Activating Host Sensors. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 8094–8103.
- (70) Kim, Y.; Wower, J.; Maltseva, N.; Chang, C.; Jedrzejczak, R.; Wilamowski, M.; Kang, S.; Nicolaescu, V.; Randall, G.; Michalska, K.; Joachimiak, A. Tipiracil Binds to Uridine Site and Inhibits Nsp15 Endoribonuclease NendoU from SARS-CoV-2. *Commun. Biol.* **2021**, *4*, No. 193.
- (71) Lin, S.; Chen, H.; Ye, F.; Chen, Z.; Yang, F.; Zheng, Y.; Cao, Y.; Qiao, J.; Yang, S.; Lu, G. Crystal Structure of SARS-CoV-2 Nsp10/Nsp16 2'-O-Methylase and Its Implication on Antiviral Drug Design. *Signal Transduction Targeted Ther.* **2020**, *5*, No. 131.
- (72) Rosas-Lemus, M.; Minasov, G.; Shuvalova, L.; Inniss, N. L.; Kiryukhina, O.; Brunzelle, J.; Satchell, K. J. F. High-Resolution Structures of the SARS-CoV-2 2'-O-Methyltransferase Reveal Strategies for Structure-Based Inhibitor Design. *Sci. Signal.* **2020**, *13*, No. eabe1202.
- (73) Kirby, T. New Variant of SARS-CoV-2 in UK Causes Surge of COVID-19. *Lancet Respir. Med.* **2021**, *9*, e20–e21.
- (74) Tegally, H.; Wilkinson, E.; Giovanetti, M.; Iranzadeh, A.; Fonseca, V.; Giandhari, J.; Doolabh, D.; Pillay, S.; San, E. J.; Msomi, N.; Mlisana, K.; Von Gottberg, A.; Walaza, S.; Allam, M.; Ismail, A.; Mohale, T.; Glass, A. J.; Engelbrecht, S.; Van Zyl, G.; Preiser, W.; Petruccione, F.; Sigal, A.; Hardie, D.; Marais, G.; Hsiao, M.; Korsman, S.; Davies, M.-A.; Tyers, L.; Mudau, I.; York, D.; Maslo, C.; Goedhals, D.; Abrahams, S.; Laguda-Akingba, O.; Alisoltani-Dehkordi, A.; Godzik, A.; Wibmer, C. K.; Sewell, B. T.; Lourenço, J.; Alcantara, L. C. J.; Pond, S. L. K.; Weaver, S.; Martin, D.; Lessells, R. J.; Bhiman, J. N.; Williamson, C.; De Oliveira, T. *Emergence and Rapid Spread of a New Severe Acute Respiratory Syndrome-Related Coronavirus 2 (SARS-CoV-2) Lineage with Multiple Spike Mutations in South Africa*; Cold Spring Harbor Laboratory, 2020.
- (75) Voloch, C. M.; Da Silva Francisco, R.; De Almeida, L. G. P.; Cardoso, C. C.; Brustolini, O. J.; Gerber, A. L.; Guimarães, A. P. D. C.; Mariani, D.; Da Costa, R. M.; Ferreira, O. C.; Frauches, T. S.; De Mello, C. M. B.; Leitão, I. D. C.; Galliez, R. M.; Faffe, D. S.; Castifeiras, T. M. P. P.; Tanuri, A.; De Vasconcelos, A. T. R.; et al. Genomic Characterization of a Novel SARS-CoV-2 Lineage from Rio de Janeiro, Brazil. *J. Virol.* **2021**, *95*, No. e00119-21.
- (76) Peng, Y.; Du, N.; Lei, Y.; Dorje, S.; Qi, J.; Luo, T.; Gao, G. F.; Song, H. Structures of the SARS-CoV-2 Nucleocapsid and Their Perspectives for Drug Design. *EMBO J.* **2020**, *39*, No. e105938.
- (77) Dinesh, D. C.; Chalupska, D.; Silhan, J.; Koutna, E.; Nencka, R.; Veverka, V.; Boura, E. Structural Basis of RNA Recognition by the SARS-CoV-2 Nucleocapsid Phosphoprotein. *PLoS Pathog.* **2020**, *16*, No. e1009100.
- (78) Zinzula, L.; Basquin, J.; Bohn, S.; Beck, F.; Klumpe, S.; Pfeifer, G.; Nagy, I.; Bracher, A.; Hartl, F. U.; Baumeister, W. High-Resolution Structure and Biophysical Characterization of the Nucleocapsid Phosphoprotein Dimerization Domain from the Covid-19 Severe Acute Respiratory Syndrome Coronavirus 2. *Biochem. Biophys. Res. Commun.* **2021**, *538*, 54–62.
- (79) Yang, M.; He, S.; Chen, X.; Huang, Z.; Zhou, Z.; Zhou, Z.; Chen, Q.; Chen, S.; Kang, S. Structural Insight Into the SARS-CoV-2 Nucleocapsid Protein C-Terminal Domain Reveals a Novel Recognition Mechanism for Viral Transcriptional Regulatory Sequences. *Front. Chem.* **2021**, *12*, No. 624765.
- (80) Thoms, M.; Buschauer, R.; Ameisemeier, M.; Koepke, L.; Denk, T.; Hirschenberger, M.; Kratzat, H.; Hayn, M.; Mackens-Kiani, T.; Cheng, J.; Straub, J. H.; Sturzel, C. M.; Frohlich, T.; Berninghausen, O.; Becker, T.; Kirchhoff, F.; Sparrer, K. M. J.; Beckmann, R. Structural Basis for Translational Shutdown and Immune Evasion by the Nsp1 Protein of SARS-CoV-2. *Science* **2020**, *369*, 1249–1255.
- (81) Schubert, K.; Karousis, E. D.; Jomaa, A.; Scaiola, A.; Echeverria, B.; Gurzeler, L. A.; Leibundgut, M.; Thiel, V.; Muhlemann, O.; Ban, N. SARS-CoV-2 Nsp1 Binds the Ribosomal Mrna Channel to Inhibit Translation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 959–966.
- (82) Clark, L. K.; Green, T. J.; Petit, C. M. Structure of Nonstructural Protein 1 from SARS-CoV-2. *J. Virol.* **2021**, *95*, 1–12.
- (83) Mandala, V. S.; McKay, M. J.; Shcherbakov, A. A.; Dregni, A. J.; Kolocouris, A.; Hong, M. Structure and Drug Binding of The SARS-CoV-2 Envelope Protein Transmembrane Domain in Lipid Bilayers. *Nat. Struct. Mol. Biol.* **2020**, *27*, 1202–1208.
- (84) Chai, J.; Cai, Y.; Pang, C.; Wang, L.; McSweeney, S.; Shanklin, J.; Liu, Q. Structural Basis for SARS-CoV-2 Envelope Protein Recognition of Human Cell Junction Protein PALS1. *Nat. Commun.*, **2021**, *143*, 3433–3438.
- (85) Gordon, D. E.; Hiatt, J.; Bouhaddou, M.; Rezelj, V. V.; Ulferts, S.; Braberg, H.; Jureka, A. S.; Obernier, K.; Guo, J. Z.; Batra, J.; Kaake, R. M.; Weckstein, A. R.; Owens, T. W.; Gupta, M.; Pourmal, S.; Titus, E. W.; Cakir, M.; Souchey, M.; McGregor, M.; Cakir, Z.; Jang, G.; O'Meara, M. J.; Tummino, T. A.; Zhang, Z.; Foussard, H.; Rojic, A.; Zhou, Y.; Kuchenov, D.; Huttenhain, R.; Xu, J.; Eckhardt, M.; Swaney, D. L.; Fabius, J. M.; Ummadi, M.; Tutuncuoglu, B.; Rathore, U.; Modak, M.; Haas, P.; Haas, K. M.; Naing, Z. Z. C.; Pulido, E. H.; Shi, Y.; Barrio-Hernandez, I.; Memon, D.; Petsalaki, E.; Dunham, A.; Marrero, M. C.; Burke, D.; Koh, C.; Vallet, T.; Silvas, J. A.; Azumaya, C. M.; Billesbolle, C.; Brilot, A. F.; Campbell, M. G.; Diallo, A.; Dickinson, M. S.; Diwanji, D.; Herrera, N.; Hoppe, N.; Kratochvil, H. T.; Liu, Y.; Merz, G. E.; Moritz, M.; Nguyen, H. C.; Nowotny, C.; Puchades, C.; Rizo, A. N.; Schulze-Gahmen, U.; Smith, A. M.; Sun, M.; Young, I. D.; Zhao, J.; Asanow, D.; Biel, J.; Bowen, A.; Braxton, J. R.; Chen, J.; Chio, C. M.; Chio, U. S.; Deshpande, I.; Doan, L.; Faust, B.; Flores, S.; Jin, M.; Kim, K.; Lam, V. L.; Li, F.; Li, J.; Li, Y. L.; Li, Y.; Liu, X.; Lo, M.; Lopez, K. E.; Melo, A. A.; Moss, F. R., 3rd; Nguyen, P.; Paulino, J.; Pawar, K. I.; Peters, J. K.; Pospiech, T. H., Jr.; Safari, M.; Sangwan, S.; Schaefer, K.; Thomas, P. V.; Thwin, A. C.; Trenker, R.; Tse, E.; Tsui, T. K. M.; Wang, F.; Whitis, N.; Yu, Z.; Zhang, K.; Zhang, Y.; Zhou, F.; Saltzberg, D.; Consortium, Q. S. B.; Hodder, A. J.; Shun-Shion, A. S.; Williams, D. M.; White, K. M.; Rosales, R.; Kehrer, T.; Miorin, L.; Moreno, E.; Patel, A. H.; Rihn, S.; Khalid, M. M.; Vallejo-Gracia, A.; Fozouni, P.; Simoneau, C. R.; Roth, T. L.; Wu, D.; Karim, M. A.; Ghousaini, M.; Dunham, I.; Berardi, F.; Weigang, S.; Chazal, M.; Park, J.; Logue, J.; McGrath, M.; Weston, S.; Haupt, R.; Hastie, C. J.; Elliott, M.; Brown, F.; Burness, K. A.; Reid, E.; Dorward, M.; Johnson, C.; Wilkinson, S. G.; Geyer, A.; Giesel, D. M.; Baillie, C.; Raggett, S.; Leech, H.; Toth, R.; Goodman, N.; Keough, K. C.; Lind, A. L.; Zoonomia, C.; Klesh, R. J.; Hemphill, K. R.; Carlson-Stevermer, J.; Oki, J.; Holden, K.; Maures, T.; Pollard, K. S.; Salii, A.; Agard, D. A.; Cheng, Y.; Fraser, J. S.; Frost, A.; Jura, N.; Kortemme, T.; Manglik, A.; Southworth, D. R.; Stroud, R. M.; Alessi, D. R.; Davies, P.; Frieman, M. B.; Ideker, T.; Abate, C.; Jouvenet, N.; Kochs, G.; Shoichet, B.; Ott, M.; Palmarini, M.; Shokat, K. M.; Garcia-Sastre, A.; Rassen, J. A.; Grosse, R.; Rosenberg, O. S.; Verba, K. A.; Basler, C. F.; Vignuzzi, M.; Peden, A. A.; Beltrao, P.; Krogan, N. J.; et al. Comparative

Host-Coronavirus Protein Interaction Networks Reveal Pan-Viral Disease Mechanisms. *Science* **2020**, *370*, No. eabe9403.

(86) Fukuzawa, K.; Tanaka, S.; Yagi, Y.; Kurita, N.; Kawashita, N.; Takaba, K.; Honma, T. In *Recent Advances of the Fragment Molecular Orbital Method*; Mochizuki, Y.; Tanaka, S.; Fukuzawa, K., Eds.; Springer: Singapore, 2021; pp 127–181.

(87) FMO Drug Design Consortium. <https://fmodd.jp/top-en/> (accessed June 8, 2021).