Prediction of Binding Pose and Affinity of SARS-CoV-2 Main Protease and Repositioned Drugs by Combining Docking, Molecular Dynamics, and Fragment Molecular Orbital Calculations

Yuma Handa[1,2], Koji Okuwaki[1,3], Yusuke Kawashima[1], Ryo Hatada[3], Yuji Mochizuki[3,4], Yuto Komeiji[2.3.5.7], Shigenori Tanaka[6], Takayuki Furuishi[1], Etsuo Yonemochi[1], Teruki Honma[7], Kaori Fukuzawa[1,2,8*]

1. Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan
2. Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan
3. Department of Chemistry and Research Center for Smart Molecules, Faculty of Science, Rikkyo University, 3-34-1 Nishi-ikebukuro, Toshima-ku, Tokyo 171-8501, Japan
4. Institute of Industrial Science, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
5. Biomedical Research Institute, AIST, Tsukuba Central 6, Tsukuba, Ibaraki 305-8566, Japan
6. Graduate School of System Informatics, Department of Computational Science, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan
7. RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
8. Department of Biomolecular Engineering, Graduate School of Engineering, Tohoku University, 6-6-11 Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan

*Corresponding author: Kaori Fukuzawa (fukuzawa-k@phs.osaka-u.ac.jp)

Abstract

COVID-19 remains a global pandemic, necessitating the urgent development of more effective therapeutics. By combining molecular docking, molecular dynamics (MD), and fragment molecular orbital (FMO) calculations, the binding structure and properties with Mpro were predicted for Nelfinavir (NLF), which was identified as a candidate compound through drug repositioning targeting the Main Protease (Mpro) produced by the causative virus, SARS-CoV-2. For the four docking poses selected by scoring using FMO energy, 100 structures each from the MD trajectory were sampled, and FMO calculations were performed and ranked based on binding energy. Besides the interaction between NLF and each Mpro residue, the desolvation effect of the pocket affected the ranking order. Furthermore, we identified several residues important in ligand recognition, including Glu47, Asp48, Glu166, Asp187, and Gln189, all of which interacted strongly with NLF. Asn142 was mentioned as a residue with hydrogen bonds or CH/π interaction with NLF; however, it was considered a transient interacting residue because of its unstable structure. Moreover, the tert-butyl group of NLF had no interaction with Mpro. Identifying weak interactions provides candidates for substituting ligand functional groups and important suggestions for drug discovery using drug repositioning. Our approach provides a new guideline for structure-based drug design starting from a candidate compound whose complex crystal structure has not been obtained.

1. Introduction

COVID-19 has been prevalent worldwide since 2019, infecting > 767 million people and killing> 6.94 million as of June 2023[1]. The causative virus, severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2), is an enveloped, positive-

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

sense, single-stranded RNA virus belonging to the genus Betacoronavirus. Viral proteins include four structural proteins (sp) that form virus particles and 16 nonstructural proteins (nsp) produced in host cells[2]. During the three-year pandemic, scientists researched potential therapeutic agents, and the Food and Drug Administration approved Actemra® (tocilizumab), which has anti-inflammatory properties[3], Olmient® (baricitinib)[4] and the molecular-targeted drug Veklury® (remdesivir)[5]. Moreover, Emergency Use Authorization approved Bebtelovimab[6], which targets monoclonal antibodies, and Evusheld® (tixagevimab co-packaged with cilgavimab)[7]. Among these approved drugs, Remdesivir was proposed by drug repositioning targeting SARS-CoV-2 RNA-dependent RNA polymerase and confirmed to be effective[8] and was approved for special cases. Moreover, the main protease (Mpro) of nsp5 attracted attention as a drug discovery target, and drug repositioning targeting Mpro was performed in several groups. Consequently, Kaletra® (lopinavir/ritonavir)[9,10] and the HIV-1 Protease inhibitor Nelfinavir (NLF) were reported to have Mpro inhibitory activity[11]. The crystal structure analysis of the NLF-complex has not been published; however, Molecular Mechanics based Molecular Dynamics (MM-MD) calculations (Hereafter, MM-MD is referred to simply as "MD") have proposed the stable structure of the complex and important hydrogen bonds[12]. Furthermore, Xokova® tablets (Ensitrelvir), an oral treatment targeting Mpro, developed in Japan, have been launched in 2022. Ensitrelvir is developed by structure-based drug design using virtual screening based on docking calculations[13]. Notably, repositioning has not yet blazed a path to new drug discovery. In the event of a new pandemic, the concept of drug repositioning is extremely important[#14, 15]. In emergency situations, the first option is to divert existing drugs whose safety and pharmacokinetics have already been confirmed in humans. Secondly, the redevelopment of an existing drug or a compound that has reached an intermediate stage of development can be used as a lead compound, thereby eliminating several drug discovery processes. Therefore, in this study, we focused on NLF, which was found as a candidate compound for drug repositioning targeting Mpro[11] and conducted virtual experiments by computational chemistry on a molecular design using it as a lead compound.

With recent advances in molecular modeling and molecular simulation techniques, the *in silico* approach is often used in the early stages of new drug development, and the computational prediction of binding poses and their interactions is crucial in designing new drugs. Docking and MD calculations, often used in *in silico* drug discovery, are useful methods for rational drug design, as they enable us to know molecular shapes, behaviors, and interaction mechanisms. The docking calculation is simpler than the MD calculation; however, since it is a simulation for a static structure, it is difficult to incorporate structural flexibility in certain aspects. Additionally, classical MD calculations can evaluate the dynamic binding states of molecules. However, since these calculations are based on the empirical force field, they are not quantitatively accurate enough. Notably, quantum chemical calculations can determine the electronic state of a molecule non-empirically based on the first principles. Specifically, the fragment molecular orbital (FMO) method[16,17] can perform high-speed and high-precision quantum chemical calculations for the entire protein and analyze quantitative intramolecular and intermolecular interaction energies[18,19]. The method has been used in drug discovery research, such as protein-ligand binding prediction[20,21] and their quantitative interaction energy analysis[17,22]. Comprehensive FMO calculations have been performed and analyzed for COVID-19-related proteins, and all results have been published in the FMO database (FMODB)[23–25].

In the "MD+FMO" calculation method (also called MM-MD/FMO protocol[17]) used in this study, which combines MD and FMO calculations, multiple structures are sampled from the trajectory of the MD calculation results, and FMO calculations are performed for each structure. Next, quantitative and dynamic analysis can be performed by averaging the obtained interaction energies. Several examples of this approach have been demonstrated. In the case of the SARS-CoV-2 Mpro and N3 inhibitor complex, statistical interaction analyses were performed where 100 structures were sampled from MD calculations, and FMO calculations were performed. MD+FMO calculations for protein-ligand complexes enabled analysis of interactions between ligands and their surrounding residues, considering thermal fluctuations[26]. In another analysis, for Mpro and N3 complex, the number of structure samples was increased to 1000 structures. Notably, applying principal component analysis and singular value decomposition to the interaction evaluation has shown that the relative importance of

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

each residue changes through structural fluctuations[27]. Additionally, when a similar method was used to predict protein-ligand binding affinity for cyclin-dependent kinase-2 and seven of its ligands, the energy obtained using MD+FMO calculation showed a better correlation with the experimentally measured ΔG than the energy for molecular mechanics-optimized X-ray crystal structures[28]. Therefore, studies combining MD and FMO have enabled highly accurate predictions of binding to dynamical structures.

The effect of water is also important in biomolecular simulations. While most MD calculations are based on conducting molecular simulations in water, conventional static FMO calculations have only dealt with adding crystal waters to the system. Since the MD + FMO calculation can use water coordinate information obtained from MD calculation, an explicit solvent water model can be constructed. However, in the examples presented above[26,28], the interactions between water and solutes have not yet been examined in detail. Studies, where water was treated by a continuous solvent model in FMO calculation for protein-ligand binding prediction include the following examples. Two cases have been reported where Molecular Mechanics Poisson–Boltzmann Surface Area (MM-PBSA) was applied to serine/threonine kinase Pim1 and its inhibitor[19], and the FMO-PB method was applied to estrogen receptor and ligand[29]. However, FMO studies that explicitly consider the water model have examined the effect of hydration on the interaction energy and charge distribution of ubiquitin[30] and DNA[31] evaluated the effect of hydration layer thickness on the interaction energy for hydrated DNA/protein complexes and analyzed PIEDA interactions within solutes and between solute and solvent[32,33] but these studies are limited to discussion of solvent effects in their bound state. Conventional quantum mechanical (QM) calculations for explicit water have been used to calculate excitation energies of small molecules[34] and physical quantities such as proton chemical shifts in protein using QM/MM calculations[35], but there have been no discussions of protein-ligand binding interactions or desolvation effects considering explicit water.

This study aimed to explore the MD+FMO approach as a method for drug repositioning. Using the Mpro and NLF complex as an example, we discussed the importance of each functional group in ligand binding from changes in molecular motion and interaction and provided new ideas for drug design. Moreover, we predicted the structure and binding affinity of complexes whose crystal structures have not been solved. Additionally, we predicted the binding energies based on the effect of explicit water.

2. Methods

The calculation flow is shown in Figure 1. First, since the crystal structure of the NLF and Mpro complex has not been solved, we predicted 30 poses by docking calculation using the co-crystal structure of similar compounds as a template. Next, we selected the dominant four poses using FMO energy scoring based on the FMO interaction energy and its binding mode. Afterward, MD simulation was performed for 100 ns each for the four poses selected. Finally, 100 structures each were sampled from the obtained trajectories, and FMO calculations were performed for 400 structures. Furthermore, interaction energy analysis was performed for each pose, and amino acid residues important for NLF recognition of Mpro were identified. Considering the deformation energy of the ligand binding to the protein and the desolvation energy in the explicit solvent, we calculated the ligand binding energy and ranked the four poses of the complex structure from the energy score.

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

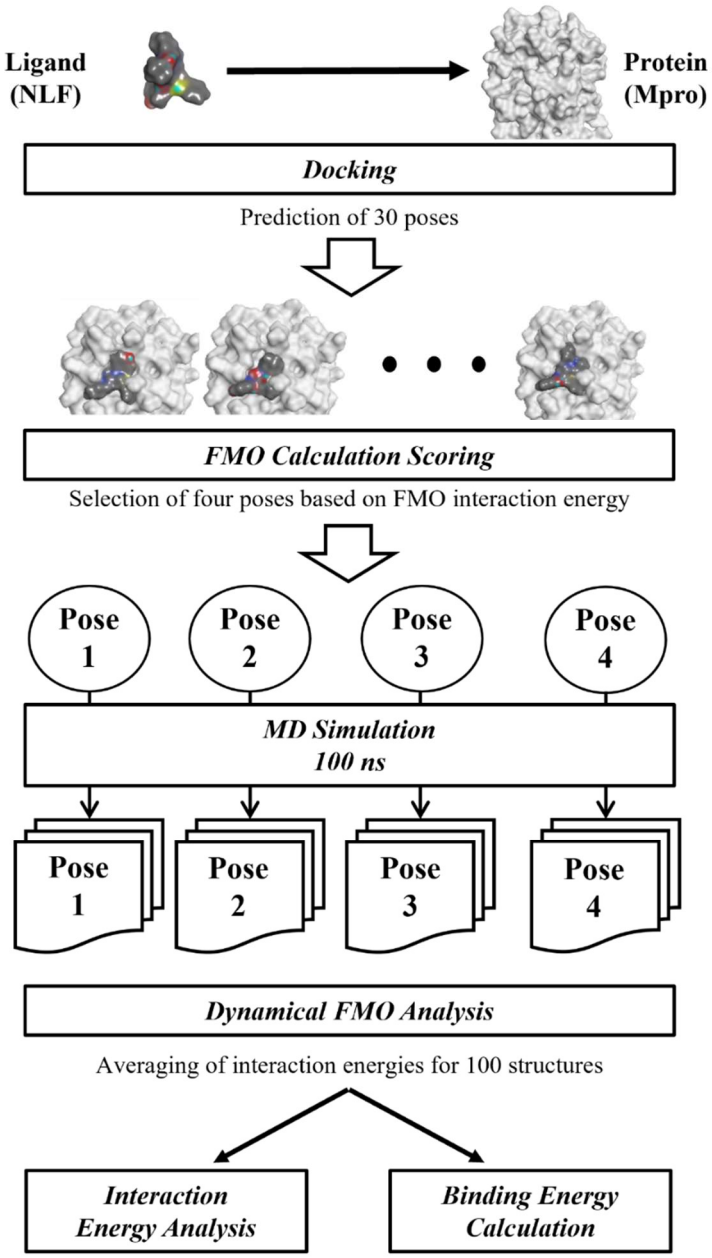

Figure 1. Workflow from docking calculations to MD + FMO calculations.

2.1 Mpro-NLF docking structure

NLF structure was obtained from the Protein Data Bank (PDB) (PDBID:3EL5)[36]. The structural formula of NLF is shown in Figure 2a. The part indicated by the red line in the figure is a peptide-like backbone that mimics a peptide cleaved by a protease. Moreover, the structure of the NLF-like ligand complex was obtained from PDB (PDBID: 6W63, Figure 2b and c). The structural formula shown in Figure 2b has a peptide-like backbone similar to NLF, and the binding mode with Mpro is assumed to be similar to that of NLF; hence, this complex structure was used as a template. Next, hydrogen atoms not determined by X-ray crystallography were added using Protonate 3D function in the Molecular Operating Environment (MOE)[37], considering the protonation state at pH = 7.0. After that, the atomic coordinates were optimized. Subsequently, general docking was performed using MOE for Mpro and NLF, and 30 poses of the complex structure were predicted. The output final score was used for comparison as the docking score. Structural refinement after docking calculation was performed with Tether = 1.0 constraint under induced fit conditions: Tether is the standard deviation in the σ radial direction and assigns a harmonic potential to the specified atom using a force constant of (3/2) kT / σ^2. All modeling in MOE used the AMBER10:EHT force field.

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

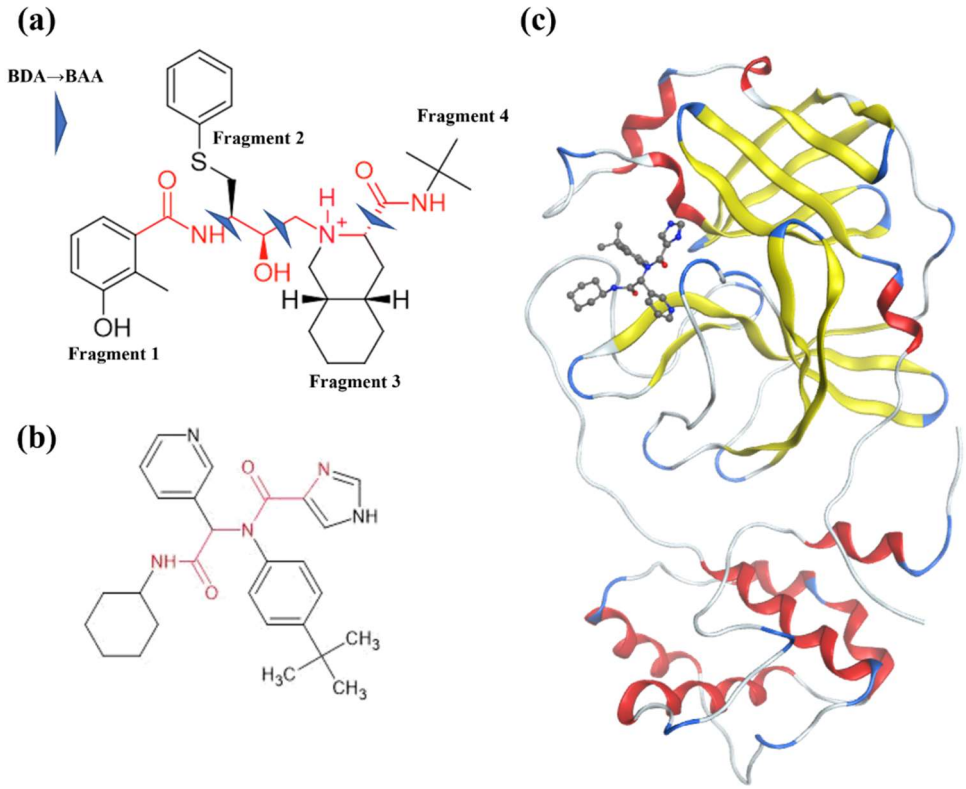

Figure 2. Structural formula and template structure of Nelfinavir (NLF). (a) Structural formula of NLF and fragmentation method. (b) NLF-like ligand. (c) Complex structure of Mpro and NLF-like ligand (PDBID:6w63).

2.2 FMO energy scoring

In the FMO method, the protein complex is divided into fragments, such as amino acid residue units, and the energies of fragment monomers and fragment dimers are determined in the environmental electrostatic potential from surrounding fragments. By integrating these energies, the total energy (E_{total}) and inter-fragment interaction energy (IFIE)[38–40] can be obtained (Equation 1).

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

Where $\tilde{E}_I^{'}$ is the energy of the monomer, excluding the contribution of the environmental electrostatic potential, and $\Delta\tilde{E}_{IJ}$ is IFIE. Using pair interaction energy decomposition analysis (PIEDA)[18,40], IFIE can be further decomposed into the following four components (Equation 2): electrostatic term (ES), exchange-repulsion term (EX), charge transfer with higher-order mixing terms (CT+mix), and dispersion term (DI).

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

The chemical bond can be interpreted by considering the contributions of each term. For example, hydrogen bonds are detected as stabilization energies using the ES and CT+mix terms, whereas dispersion interactions such as CH/π and π-π are detected using the DI term. In this study, Mpro was divided into each amino acid residue, and NLF was divided into four fragments (F1, F2, F3, and F4) (Figure 2a). The ABINIT-MP program [39,42] was used for FMO calculations; electron correlation effects are incorporated by the second-order Møller–Plesset perturbation (MP2) theory, which was efficiently implemented in ABINIT-MP [43-45]. For the basis function[17,39], 6-31G*, standardly used in FMO calculations, was used.

The procedure of FMO scoring is as follows. First, for each of the 30 poses of the docking structure obtained in Section 0, structural optimization was performed using the AMBER10:EHT force field. When hydrogen was not constrained, amino

https://doi.org/10.26434/chemrxiv-2023-hrsvj ORCID: https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

acid side chains within 4.5 Å around the ligand were constrained by Tether = 1.0, and the coordinates of other heavy atoms were fixed. Next, FMO calculation was performed for the optimized structure of 30 poses obtained. Lastly, scoring was performed using the sum of the IFIE of each amino acid residue of the ligand and Mpro (Total IFIE), and the top two poses showing stable values and two poses with different binding modes were selected.

2.3 MD calculation

MD calculation of 100 ns was performed for each of the four poses of the candidate structure obtained in Section 2.2. For MD calculation, a heat process from 0 K to 310 K was performed for 50 ps using the NVT ensemble. Next, an equalization process was performed at 310 K for 50 ps (NPT ensemble). Furthermore, density relaxation was performed for 1 ns (NPT ensemble), and a production run was performed for 100 ns at 310 K (NPT ensemble). Note that the pressure at NPT is 1013 hpa. Notably, Amber10:EHT was used for protein, ligand, and ion parameters. Additionally, TIP3P water was used as the solvent, and Na^+ was used as the counter ion. The bond distances involving hydrogen were not constrained, and the time step was 1 fs. This MD calculation was performed under a periodic boundary condition. Moreover, the MD calculation was performed under the same conditions for the protein apo structure containing no ligand and the ligand structure alone. The MD calculations in this study were performed using the AMBER16 program[46]. The MD calculations in this study were performed with the AMBER16 program[44] using molecular topology files constructed with MOE.

2.4 Dynamical FMO calculation

For each 100 ns trajectory of four complexes obtained using MD calculation, 100 structures were extracted at 50 ps intervals from the latter half of 50 ns. Therefore, 400 structures were extracted, and FMO calculation was performed. Each sample structure was extracted to a droplet form with water molecules within 4 Å of Mpro; this criterion of water layer thickness was determined following Refs[30-33]. Before performing FMO calculations, the geometry was optimized using the AMBER10:EHT force field for each sampling structure. As in the method shown in Section 2.2, hydrogen was not constrained. However, amino acid side chains within 4.5 Å around the ligand were constrained by Tether = 1.0, and the coordinates of other heavy atoms were fixed. Next, FMO calculation was performed at MP2/6-31G* level using ABINIT-MP Program. Additionally, FMO calculation was performed under similar conditions for the protein apo structure containing no ligand. Subsequently, the average value and standard deviation of 100 structures of total IFIE and PIEDA with Mpro for NLF obtained using these FMO calculations were calculated for each pose. Additionally, since NLF is divided into four fragments (Figure 2a), interaction analysis was performed by calculating the average value and standard deviation of 100 structures for IFIE and PIEDA for each fragment.

2.5 Ligand binding energy calculation

The binding energy (ΔE_{bind}) is expressed as the sum of the protein-ligand intermolecular interaction energy (ΔE^{int}), the deformation energy of the ligand (ΔE_{lig}^{def}), and the solvation energies (ΔE^{sol})[19,29] (Equation 3). Additionally, ΔE^{int} is the total energy of the complex (E_{com}) minus the total energy of the protein alone (E_{pro}) and the ligand alone (E_{lig}) (Equation 4). Here, the sum of the IFIEs of the ligand in the complex and each amino acid residue (total IFIE; $\sum\Delta\tilde{E}_{IJ}$) is used for approximation (Equation 5). Additionally, here, I indicate the ligand fragment (NFL) and J indicates each amino acid fragment of the protein (Mpro). Furthermore, $\Delta E_{\mathrm{lig}}^{\mathrm{def}}$ is the deformation energy of the ligand structure, indicated by the difference in total energy between the complexed (lig(com)) and isolated forms in water (lig(sol)) (Equation 6). The average of 100 structures for each pose was used to calculate the complex (lig(com)) of deformation energy ($\Delta E_{\mathrm{lig}}^{\mathrm{def}}$). Moreover, the isolated structure (lig(sol)) was optimized with B97D[47]/6-31G* using Gaussian16, and single-point calculation was performed with MP2/6-31G* using ABINIT-MP ($E_{\mathrm{lig(sol)}}$).

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Furthermore, the solvation energies (ΔE^{sol}) are the complex solvation energies minus the protein and ligand solvation energies (Equation 7).

$$\Delta E_{bind} = \Delta E^{int} + \Delta E_{lig}^{def} + \Delta E^{sol} \quad (3)$$

$$\Delta E^{int} = E_{\mathrm{com}} - (E_{\mathrm{pro}} + E_{\mathrm{lig}}) \quad (4)$$

$$\approx \sum \Delta \tilde{E}_{IJ} \quad (5)$$

$$\Delta E_{\mathrm{lig}}^{\mathrm{def}} = E_{\mathrm{lig(com)}} - E_{\mathrm{lig(sol)}} \quad (6)$$

$$\Delta E^{sol} = E_{com}^{sol} - \left(E_{pro}^{sol} + E_{lig}^{sol}\right) \quad (7)$$

Here, the solvation energy of the complex (E_{com}^{sol}) is half the sum of the interaction energies between ligand and water molecules ($\sum \Delta E_{\mathrm{IK}}^{\mathrm{com}}$) and between amino acids and water molecules ($\sum \Delta E_{\mathrm{JK}}^{\mathrm{com}}$) in the complex form (Equation 8). Additionally, the protein solvation energy (E_{pro}^{sol}) is half the interaction energy between amino acids and water molecules in the apo structure (Equation 9). Similarly, the ligand solvation energy (E_{lig}^{sol}) is half the interaction energy between the ligand and water molecules in the isolated free structure (Equation 10). K indicates a water molecule fragment.

$$E_{com}^{sol} = 1/2(\sum \Delta E_{\mathrm{IK}}^{\mathrm{com}} + \sum \Delta E_{\mathrm{JK}}^{\mathrm{com}}) \quad (8)$$

$$E_{pro}^{sol} = 1/2(\sum \Delta E_{\mathrm{JK}}^{\mathrm{apo}}) \quad (9)$$

$$E_{lig}^{sol} = 1/2(\sum \Delta E_{\mathrm{IK}}^{free}) \quad (10)$$

The method for selecting the amino acid fragment (J) used in determining the protein solvation energy (E_{pro}^{sol}) (Equation 9) is as follows: first, residues whose average distance between nearest neighbor atoms from the ligand is within 4 Å were determined in 100 complex structures sampled for each pose. This was performed for four poses, and amino acid residues targeted in one or more poses were commonly subjected to integration in all poses. Consequently, 25 amino acid residues were used (Thr25, Thr26, Leu27, His41, Ser46, Met49, Leu50, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Leu167, His172, Phe181, Asp187, Arg188, Gln189, Thr190, Ala191, and Gln192) (Figure S4).
In this study, ΔE^{int} is written as $\Delta E^{int(static)}$ for static structure and ΔE^{int} for dynamic structure.

3. Results and Discussion

3.1 Scoring using docking and FMO calculations

The 30-pose composite structure obtained by docking Mpro and NLF is shown in Figure 3. All of the poses were scored using FMO calculation (Table S1), and considering the binding mode and FMO scoring results, the candidate structures were narrowed down to the four poses (Figure 4), which had different binding modes to ligands. As presented in Table S1, the ranking by docking score is 1, 2, 3, and 4 in descending order of scores, with Pose 1 having the best value. However, the ranking by FMO scoring is 3, 1, 4, and 2 in descending order, with Pose 3 showing the best value. Between each pose, the IFIE value and the number of hydrogen and CH/π bonds based on it were different. Here, $\Delta E^{int(static)}$ of Poses 1– 4 is -147– -193 kcal/mol, which was more stable than $\Delta E^{int(static)}$ (Figure S1) (-145 kcal/mol) in the original complex structure with the NLF analog (Figure 2c), used as a template for docking calculations.

In Pose 1, we confirmed two hydrogen bonds (SH-O bonds) at Thr25 and Cys145 with NLF and four CH/π bonds with Thr25, His41, Asn142, and Glu166. In Pose 2, we confirmed one hydrogen bond with Thr190 (Ala191 C=O) and three CH/π bonds with His41, ARG188, and Gln189. Notably, dispersive interactions with Met165 contributed to stabilization. Pose 3 showed two hydrogen bonds with the main chain of Phe140 and Asn142 and three CH/π bonds with His41, Asn142, and Glu166. Lastly, in Pose 4, we confirmed one hydrogen bond with Thr190 (Ala191 C=O) and three CH/π bonds with Asn142, Glu166, and Gln189 (Figure 4) (Figure S3). Notably, a larger number of hydrogen bonds in each pose resulted in a more stable ES term (Table 1). Similarly, a greater number of CH/π bonds and interactions between hydrophobic functional groups gave a more stable DI term.

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

In a previously reported FMODB COVID-19 special issue[25], the interaction between Mpro and various ligands was comprehensively analyzed using the X-ray co-crystal structures of 110 Mpro-Ligand complexes. Comparing the findings in that study with the present results, the hydrogen bonds with the main chain of Phe140 Asn142 seen in Pose 3 were similar[25,48]. In contrast, the two hydrogen bonds in Pose 1 and the hydrogen bond with Thr190 seen in Poses 2 and 4 were observed in some structures, but not universally.

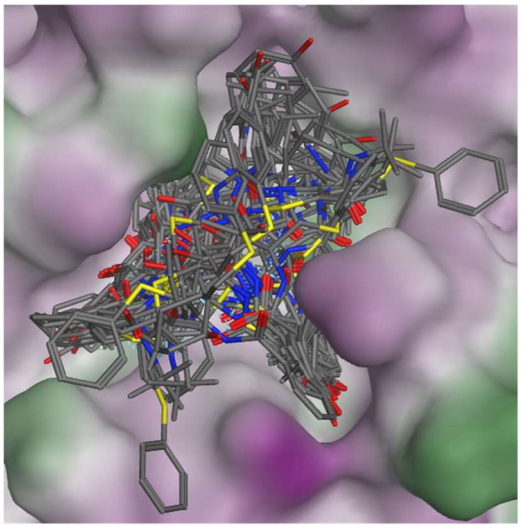

Figure 3. Superposition of 30 structures obtained by docking Mpro and NLF. Among the ligand-binding pockets of Mpro, purple indicates hydrophilic, and green indicates lipophilic.

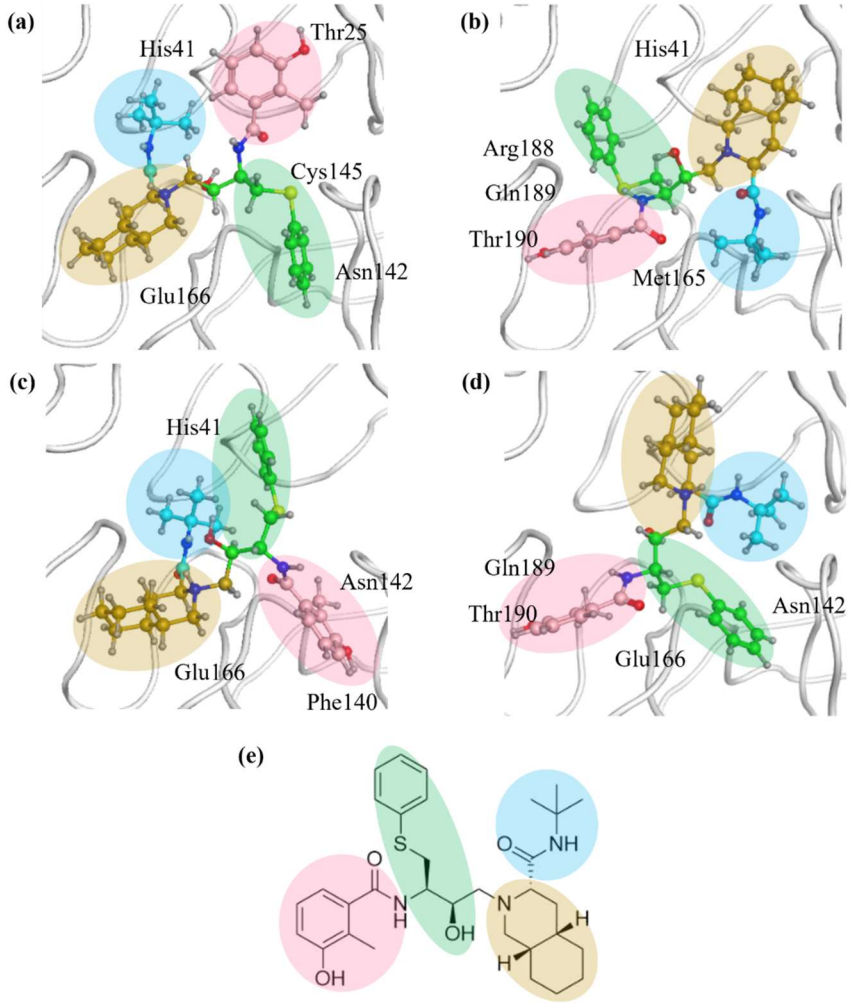

Figure 4. Four structures selected based on docking and FMO scoring. (a) Pose 1, (b) Pose 2, (c) Pose 3, (d) Pose 4.

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Table 1. Interaction and biding energies of each docking pose.

	Pose 1	Pose 2	Pose 3	Pose 4
Docking Structure				
docking score	-9.70	-9.63	-9.07	-9.00
ES	-150.4	-135.1	-139.5	-145.9
EX	105.1	118.1	63.2	89.9
CT + mix	-35.0	-24.4	-33.1	-33.4
DI	-103.5	-105.6	-83.1	-84.9
$\Delta E^{int(static)}$ (Total IFIE)	-183.7	-147.1	-192.6	-174.3
Dynamical FMO analysis				
ES	-116.4 ± 23.3	-110.5 ± 20.4	-199.0 ± 23.7	-138.5 ± 18.6
EX	61.4 ± 15.2	46.7 ± 16.0	92.4 ± 18.6	79.7 ± 13.5
CT + mix	-24.8 ± 5.1	-18.2 ± 5.6	-34.6 ± 5.7	-27.1 ± 3.9
DI	-65.0 ± 10.2	-53.9 ± 9.8	-70.8 ± 6.6	-77.9 ± 6.7
ΔE^{int} (Total IFIE)	-144.8 ± 23.5	-135.9 ± 20.0	-212.0 ± 21.0	-167.2 ± 15.1
Binding Energies				
$\Delta E_{\mathrm{lig}}^{\mathrm{def}}$	73.2 ± 10.9	70.9 ± 9.1	71.2 ± 10.9	80.2 ± 10.3
ΔE^{sol}	61.5	60.2	123.5	107.6
$\Delta E^{int} + \Delta E_{\mathrm{lig}}^{\mathrm{def}}$	-71.6±25.0	-65.0±21.0	-140.7±23.3	-86.9±18.4
$\Delta E^{int} + \Delta E^{sol}$	-83.3	-75.7	-88.5	-59.6
ΔE_{bind}	-10.1	-4.8	-17.2	20.65

Energy values are in kcal/mol.

3.2 MD calculations

Trajectory analysis in MD calculations from 0 ns to 100 ns yielded the root mean square deviation (RMSD) of the complex protein backbone (Figures 5a-5d). The average RMSDs of protein backbone atoms (C, CA, and N) were 2.56, 1.93, 2.22, and 1.80 from Poses 1–4, sequentially and the standard deviations of RMSD were 0.36, 0.38, 0.33, and 0.36, respectively. From these results, although the structure differed from the docking structure, the standard deviation was small, indicating no significant structural change. Notably, the RMSD of the last 10 ns (90 –100 ns) in Pose 2 was slightly > 3.0. However, this is due to structural fluctuations at the C-terminus and did not affect the interaction with the NLF. The average RMSD values of NLF were 2.73, 2.33, 3.30, and 3.55 from Poses 1–4 sequentially, and the standard deviations of RMSD were 0.42, 0.36, 0.98, and 0.34, respectively. Notably, the average value indicates that the structural deformation of NLF is greater than that of Mpro in induced fit. This is particularly noticeable in Poses 3 and 4.

Moreover, when calculating the RMSD only in 50–100 ns, the average values of RMSD were 2.59, 2.52, 3.99, and 3.51, and the standard deviations of RMSD were 0.45, 0.37, 0.15, and 0.31, starting from Pose 1. Compared with the results from 0–100 ns, the standard deviation of RMSD does not change in Poses 1, 2, and 4; however, it is particularly small in Pose 3. This suggests that MD in Pose 3 reached equilibrium after 50 ns. Therefore, we performed structural sampling for FMO calculation focusing on the latter half of 50 ns. Detailed structure analysis for the MD calculation and FMO results is described in Section 3.3.

3.3 Evaluation of ligand binding using dynamical FMO calculation

For each 100 sampled structures extracted from MD trajectories of four poses, interaction, and binding energy were evaluated using FMO calculations. Table 1 and Figure 6 present the average value and standard deviation of the total IFIE for each pose (ΔE^{int}). The order of IFIE stability (ΔE^{int}) was 3, 4, 1, and 2; hence, Pose 3 was the most stable. Although the EX term is small in Poses 1 and 2, the stabilization energies, such as the ES and DI terms, are also small. Moreover, Pose 3 had larger stabilization energy than other poses, and the ES term was particularly stable. Furthermore, Pose 4 had a large

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

stabilization energy similar to Pose 3, and the DI term was the most stable. However, pose 4 was not the most stable because the ES term was smaller than that of Pose 3. The characteristics of each pose are discussed as follows.

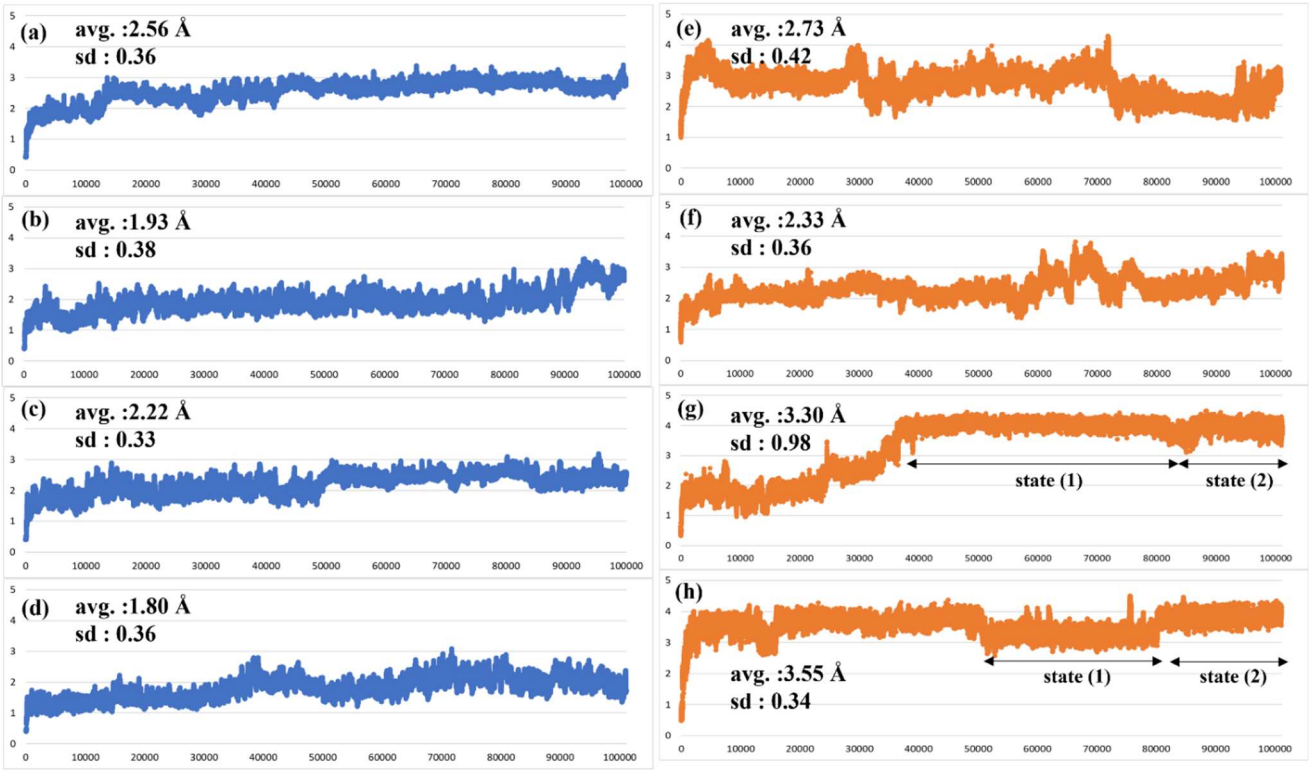

Figure 5. Root mean square deviation (RMSD) of protein main chain and that of ligand in the MD structure of each pose. (a) RMSD of the protein main chain in Pose 1. (b) RMSD of the protein main chain in Pose 2. (c) RMSD of the protein main chain in Pose 3. (d) RMSD of the protein main chain in Pose 4. (e) RMSD of heavy atoms of NLF in Pose 1. (f) RMSD of heavy atoms of NLF in Pose 2. (g) RMSD of heavy atoms of NLF in Pose 3. (h) RMSD of heavy atoms of NLF in Pose 4.

3.3.1 Fluctuations in the Mpro-NLF bond structure and changes in their interactions

(1) Pose 1

Figure 6 depicts the amino acid residues of Mpro interacting with NLF in Pose 1. Here, fragments in which the distance between nearest neighbor atoms between fragments is within 2.0 times the sum of the van der Waals radius of atoms are defined as neighboring fragments. Besides the neighboring fragment, only residues with a total IFIE more stable than -15 kcal/mol or a DI term more stable than -5 kcal/mol are illustrated (similar criteria are used in Figures 7–9). Notably, the positive charge of NLFs overestimates the ES interactions between NLFs and each residue[39]. Notable PIEDA energy values range from several negative tens of kcal/mol for the ES term; however, even -5 kcal/mol for the DI term is significant. In interactions with the whole NLF, Leu 27 had an IFIE of -17.9 ± 8.2 kcal/mol (ES: -18.4 ± 11.0 kcal/mol, CT: -4.5 ± 3.0 kcal/mol, DI: -4.7 ± 1.6 kcal/mol), suggesting hydrogen bonding based on ES and CT energies, and CH/π interactions based on the DI energy. Similarly, consideration was made based on each PIEDA component. Glu 47, Asp 48, and Asp 187 showed strong IFIE peaks; however, they were all due to ES energy. Hence, they are considered electrostatic interactions between fragments not in direct contact. Additionally, Asn142 had an IFIE of -10.7 ± 6.6 kcal/mol (ES: -6.8 ± 6.8 kcal/mol, CT: -4.7 ± 2.7 kcal/mol, DI: -8.9 ± 4.0 kcal/mol), suggesting hydrogen bonding based on stabilization of ES and CT energies and CH/π interaction based on stabilization of DI energy. Moreover, the large standard deviation, particularly in the ES term, suggests that the interaction changes with time in the MD trajectories. Predominantly, Glu 166 had ES energy and large electrostatic interaction (ES: -34.1 ± 5.7 kcal/mol), and the DI energy was slightly stable (DI: -4.7 ± 2.7 kcal/mol). This suggests CH/π interactions. Furthermore, Gln 189 had an IFIE of -14.8 ± 16.3 kcal/mol (ES: -11.5 ± 16.3 kcal/mol, CT: -3.1 ± 3.2 kcal/mol,

https://doi.org/10.26434/chemrxiv-2023-hrsvj ORCID: https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

DI: -6.8 ± 3.8 kcal/mol), suggesting hydrogen bonding based on ES and CT energies and dispersive interactions between hydrophobic functional groups based on the DI energy. The large standard deviation of Gln189 is due to fluctuations in interactions with F3 and F4, which will be described later.

Detailed interactions for each fragment are as follows. For F1, IFIE with Leu 27 was -15.3 ± 7.8 kcal/mol (ES: -15.8 ± 10.6 kcal/mol, CT: -4.5 ± 3.0 kcal/mol, DI: -4.7 ± 1.6 kcal/mol). We observed a hydrogen bond between the OH group of F1 and Leu 27 (the C=O from Leu26 belongs to the Leu 27 fragment due to fragmentation rules) and a CH/π interaction between the benzene ring of F1 and the Leu 27 side chain. For F2, IFIE with Glu 166 was -6.3 ± 3.7 kcal/mol (ES: -4.8 ± 3.5 kcal/mol, CT: -0.9 ± 1.4 kcal/mol, DI: -4.6 ± 2.8 kcal/mol). The CH/π interaction in the benzene ring of F2 with CH in the side chain of Glu 166 was confirmed from the DI energies. The IFIE with Asn142 was -3.2 ± 2.6 kcal/mol (ES: -0.4 ± 2.6 kcal/mol, CT: -3.2 ± 1.5 kcal/mol, DI: -6.7 ± 2.8 kcal/mol), and we confirmed a CH/π interaction between F2 and the Asn142 side chain. From approximately 95 ns, F2 left its binding site and moved to the solvent-exposed site. The interaction energy from 95–100 ns decreased to approximately 70% of the interaction energy from 50–94 ns. In F3 and F4, F3 showed an IFIE with Gln 189 of -12.2 ± 9.6 kcal/mol (ES: -9.9 ± 9.0 kcal/mol, CT: -1.7 ± 1.2 kcal/mol, DI: -3.6 ± 1.4 kcal/mol). F4 showed an IFIE with Gln 189 of -2.1 ± 7.3 kcal/mol (ES: -1.6 ± 8.6 kcal/mol, CT: -1.1 ± 1.8 kcal/mol, DI: -2.8 ± 2.0 kcal/mol). From these results, although Gln189 has timing that forms hydrogen bond bridges with N of F3 and N of F4, it failed to maintain this bridge in several sampling structures, as confirmed by the large standard deviations of IFIE and ES. Notably, the tert-butyl group of F4 did not interact.

Therefore, although Pose 1 showed partially strong binding, it is difficult to say that it is an excellent binding pose because F2 could not maintain the interaction, and F4 was not used well for binding.

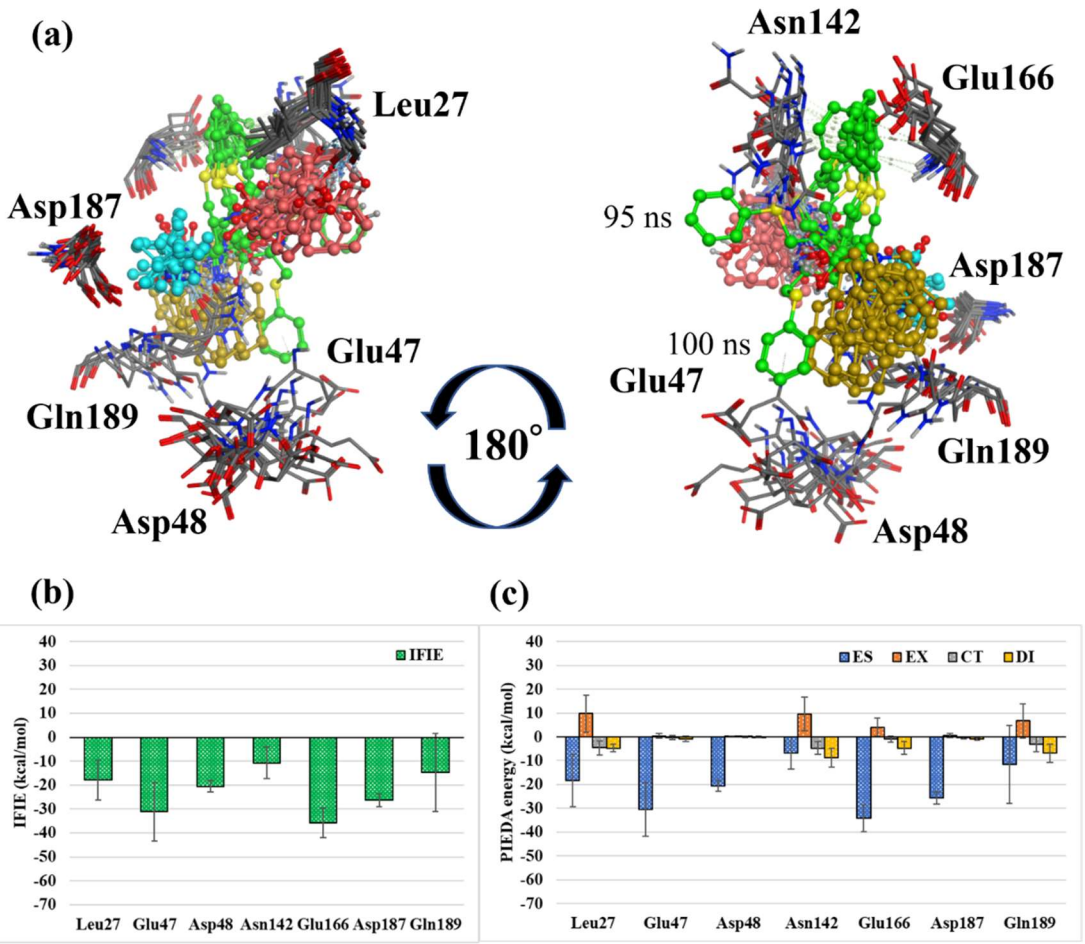

Figure 6. Interactions with surrounding amino acids in Pose 1. (a) Amino acid residues of main protease interacting with Nelfinavir in Pose 1. The four ligand fragments are colored F1: pink, F2: green, F3: yellow, and F4: light blue. (b)(c) Interaction energies between each amino acid residue and NLF; (b) IFIE and (c) PIEDA energies.

https://doi.org/10.26434/chemrxiv-2023-hrsvj ORCID: https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

(2) Pose 2

Figure 7 shows the amino acid residues of the neighboring fragments interacting with the NLF of Pose 2. Regarding the interaction with the whole NLF, IFIE was strong in Glu 47, Asp 48, and Glu 166, but mostly due to the ES term; hence, it is considered a slightly distant electrostatic interaction. Moreover, Met 49 and Met 165 showed a hydrophobic interaction due to the stabilization of the DI energy. However, since it exists near the ligand, the EX energy was large, canceling the stabilizing interaction, thereby weakening the IFIE. Asp 187 had DI and ES terms, suggesting that it acquires CH/π and ES interactions. The contributions of ES, CT, and DI terms suggest that Gln 189 had hydrogen bonds and CH/π interactions. However, considering the large standard deviation, these residues may change their interaction depending on the timing of sampling from MD.

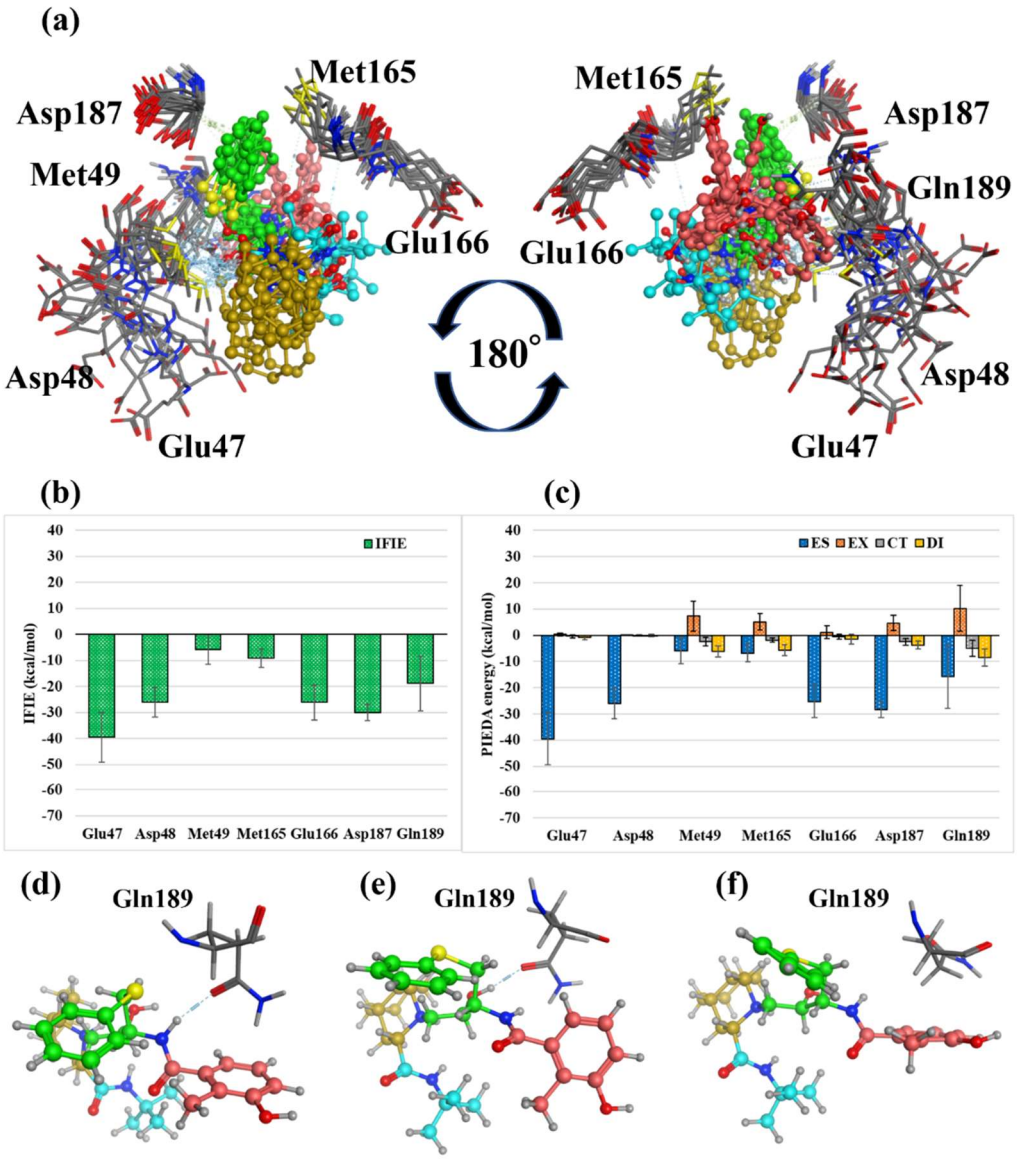

Figure 7. Interactions with surrounding amino acids in Pose 2. (a) Amino acid residues of main protease interacting with Nelfinavir in Pose 2. The four ligand fragments are colored F1: pink, F2: green, F3: yellow, and F4: light blue. **(b)(c)** Interaction energies between each amino acid residue and NLF; (b) IFIE and (c) PIEDA energies. (d) NH...O hydrogen bond between Gln189 and F1 of NLF (at 55 ns of MD calculation). (e) OH...O hydrogen bond between Gln189 and F2 of NLF (at 75 ns of MD calculation). (f) CH/π interaction between Gln189 and F1 of NLF (at 85 ns of MD calculation).

Detailed interactions for each fragment are as follows. In F1, IFIE with Gln189 was -5.2 ± 5.0 kcal/mol (ES: -4.3 ± 6.4 kcal/mol, CT: -2.4 ± 2.3 kcal/mol, DI: -3.9 ± 3.0 kcal/mol). We observed a hydrogen bond between the N of F1 and the

https://doi.org/10.26434/chemrxiv-2023-hrsvj ORCID: https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

carbonyl of the Gln 189 side chain and a CH/π bond between the benzene ring of F1 and the CH of the Gln 189 side chain. (Figures 7d and f) However, the standard deviation of each component was large; hence, their interactions were not necessarily retained. In F2, IFIE with Gln 189 was -9.0 ± 4.1 kcal/mol (ES: -6.8 ± 5.0 kcal/mol, CT: -2.6 ± 1.5 kcal/mol, DI: -4.5 ± 1.7 kcal/mol). We observed a hydrogen bond between the OH group of F2 and the carbonyl of Gln 189; however, the interaction was not preserved because of large standard deviations. Moreover, we suggested the formation of hydrophobic interactions between CHs of F2 and Gln 189 side chains. Additionally, the DI energy with Met 165 was -4.0 ± 1.5 kcal/mol, and the DI energy with Asp 187 was -3.7 ± 1.5 kcal/mol, suggesting the formation of a CH/π bond. (Figure 7e) F3 and Met49 had a DI energy of -2.6 ± 1.5 kcal/mol, suggesting a hydrophobic interaction; however, we observed no other interacting residues with IFIE > -5 kcal/mol. Similarly, in F4, no interacting residues with IFIE >-5 kcal/mol were identified. Therefore, although Pose 2 formed partially strong interactions, F3 and F4 were not used well for binding, suggesting that it is not a good binding pose.

(3) Pose 3

Figure 8 depicts the amino acid residues of the neighboring fragments interacting with the NLF of Pose 3. Since this pose had a larger ES term than other poses, the energy for each amino acid had a larger ES term. Regarding the interaction with the whole NLF, Glu47, Asp48, and Asp187 had mostly ES terms; however, Glu166 and Gln189 had CT and ES terms, suggesting that they form hydrogen bonds. Although Met49 had a smaller total IFIE than other residues, it had a DI term of -10.0 ± 3.3 kcal/mol, suggesting a strong hydrophobic interaction. Lastly, although Asn142 did not show a large interaction overall, it acquired a strong interaction depending on the time of MD sampling.

Detailed interactions for each fragment are as follows. In F1, IFIE with Glu166 was -24.5 ± 9.7 kcal/mol (ES: -29.9 ± 13.7 kcal/mol, CT: -7.1 ± 4.0 kcal/mol, DI: -5.1 ± 2.0 kcal/mol). We observed a hydrogen bond between the hydroxy group of F1 and the side chain carbonyl group of Glu166 and a CH/π bond between the benzene ring of F1 and CH of the side chain of Glu166. Moreover, F2 had a DI energy of -6.2 ± 2.1 kcal/mol with Met49, and the benzene ring of F2 had a CH/π bond with the side chain of Met49. F3 interacted with Gln189 at -29.2 ± 3.1 kcal/mol (ES: -28.8 ± 5.0 kcal/mol, CT: -5.1 ± 1.2 kcal/mol, DI: -6.1 ± 0.8 kcal/mol). Besides the hydrophobic interactions with the hydrocarbon moiety of F3, the N of F3 and the side-chain carbonyl of Gln189 formed hydrogen bonds. Furthermore, F4 interacted with Gln189 of -12.5 ± 1.4 kcal/mol (ES: -15.5 ± 2.7 kcal/mol, CT: -2.5 ± 1.0 kcal/mol). Therefore, the N of F4 is hydrogen-bonded to the carbonyl group of the Gln189 side chain, similar to F3. Notably, the tert-butyl group of F4 did not interact.

In the above-described Figure 5, we considered that the state of ligands in MD was divided into two types, 50–85 ns (State 1) and 85–100 ns (State 2). However, this was mainly due to structural changes in F2. Here, Met49 maintained the binding by cooperating with F2; however, Asn142 changed its structure in each state; therefore, interaction differed (Figure 8d-f). Moreover, the total IFIE (ΔE^{int}) was more than thrice higher in State 2 than in State 1, and the interaction was enhanced in State 2 and ES, CT, and DI. This is because in State 2, the carbonyl of F1 and N of Asn142 formed a new hydrogen bond, and the benzene ring of F1 and the main chain carbonyl of Leu141 (Asn142 C=O) formed a new π/π interaction.

Therefore, Pose 3 could use all fragments for binding, and its binding energy (ΔE^{int}) was the most stable among the four poses. Moreover, ΔE^{int} increased from -203.2 ± 17.2 kcal/mol to -232.5 ± 13.2 kcal/mol with the change from State 1 to State 2. This confirms that State 2 is stable in terms of binding energy. Consequently, functional groups that have fluctuating interactions with specific residues, such as Asn142, are candidates for molecular optimization, and it is expected that converting them to stable interactions will improve binding affinity.

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

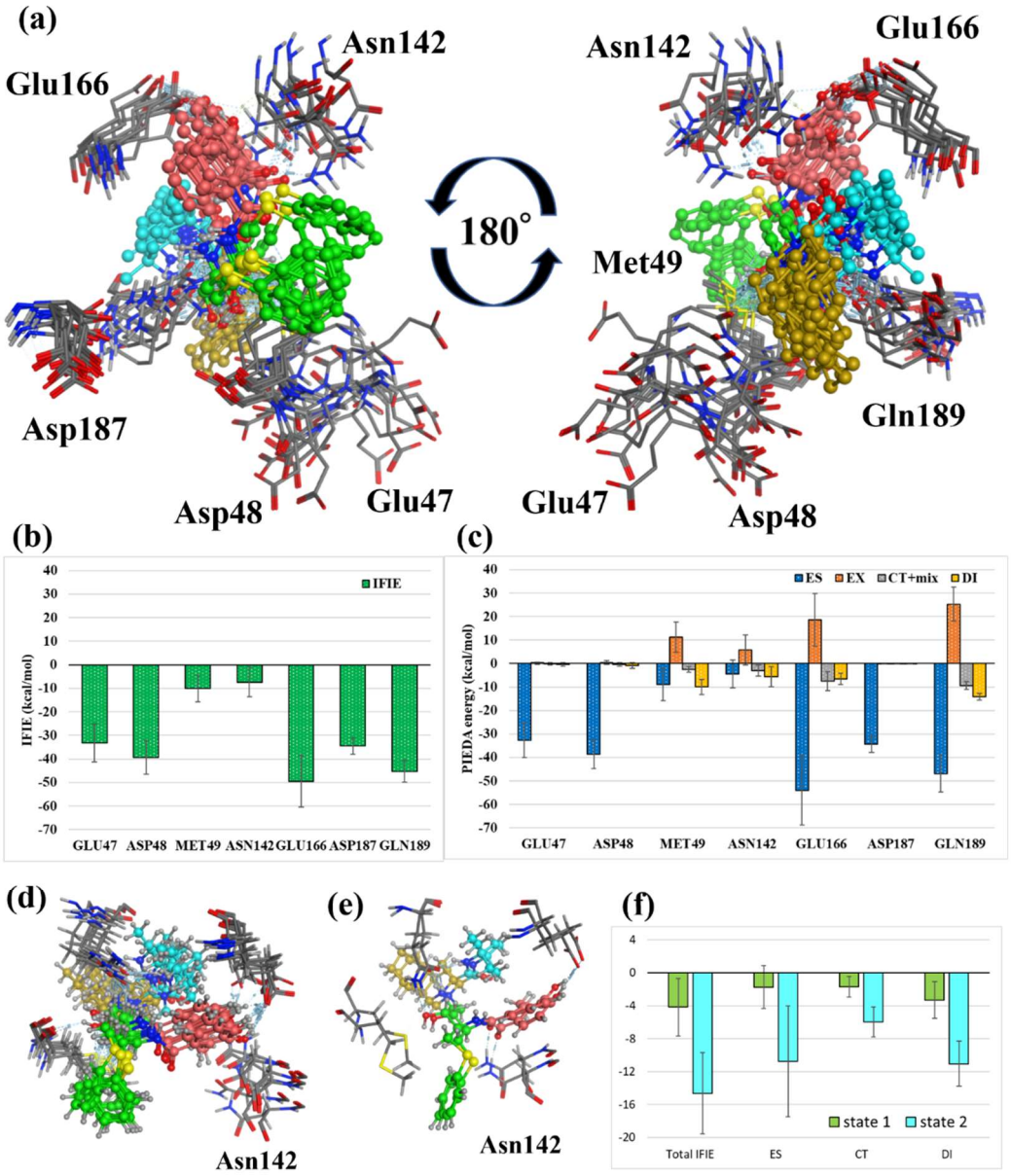

Figure 8. Interactions with surrounding amino acids in Pose 3. (a) Amino acid residues of main protease interacting with Nelfinavir in Pose 3. The four ligand fragments are colored F1: pink, F2: green, F3: yellow, and F4: light blue. (b)(c) Interaction energies between each amino acid residue and NLF; (b) IFIE and (c) PIEDA energies. (d) Structure of Asn142 and NLF from 50 ns to 85 ns. (State 1). (d) Structure of Asn142 and NLF from 85 ns to 100 ns. (State 2). (f) IFIE and PIEDA in State 1 and State 2.

(4) Pose 4

Figure 9 shows the amino acid residues of the neighboring fragments interacting with the NLF of Pose 4. Regarding the interaction with the whole NLF, Glu47, Asp48, and Asp187 had strong IFIEs, which were mostly electrostatic interactions due to the ES term. Moreover, Met165, Glu166, and Gln189 had DI and ES terms attributed to CH/π interaction with the ligand. Asn142 had ES, CT, and DI terms, suggesting hydrogen bonding and CH/π interactions. However, since standard deviations were large, the structural changes due to MD were significant, as with Pose 3.

Detailed interactions for each fragment are as follows. F1 interacted with Gln189 at -14.4 ± 4.8 kcal/mol (ES: -12.6 ± 6.3 kcal/mol, CT: -5.2 ± 1.4 kcal/mol, DI: -8.7 ± 2.0 kcal/mol), suggesting hydrogen and CH/π bonding constituted by NH/O hydrogen bonds between the carbonyl group of F1 and the Gln189 side chain and CH/π bonds between the benzene ring of F1 and the Gln189 side chain. Additionally, FI interacted with Met165 at -6.4 ± 1.4 kcal/mol in the DI term, suggesting the formation of a CH/π bond between the benzene ring of F1 and the side chain of Met165. F2 formed a CH/π bond between its benzene ring and the side chain of Asn142 since the DI energy with Asn142 was -6.9 ± 2.3 kcal/mol. Additionally, the DI

https://doi.org/10.26434/chemrxiv-2023-hrsvj ORCID: https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

energy of F2 and Glu166 was -5.9 ± 1.5 kcal/mol, suggesting that the benzene ring of F2 formed a CH/π bond with the side chain of Glu166. F2 underwent structural changes at 80 ns (Figures 9d and e) but retained the CH/π bond with Asn142. F3 had electrostatic interactions with negatively charged residues, such as Glu47 and Asp48, due to the positive charge of N; however, no hydrogen bonding or CH/π interactions were observed. Lastly, F4 formed a hydrogen bond with Asn142 in the structure from 50–75 ns (State 1) (IFIE: -11.9 ± 4.8 kcal/mol, ES: -12.6 ± 6.0 kcal/mol, CT: -2.6 ± 1.2 kcal/mol). However, the structure after 80 ns lost hydrogen bonding (IFIE: -0.1 ± 1.3 kcal/mol, ES: 0.2 ± 1.3 kcal/mol, CT: -0.0 ± 0.1 kcal/mol). Notably, the tert-butyl group did not interact in Pose 4.

Therefore, Pose 4 could use all fragments for binding, and its binding energy (ΔE^{int}) was more stable than those of Pose 1 and 2, which only partially interacted. Regarding Asn142, confirming the interaction observed from the whole ligand (Figure 9f), State 1 acquired more than twice the interaction compared with State 2. Additionally, the interaction was enhanced in ES and CT due to structural changes in F2, F4, and Asn142 at 80 ns.

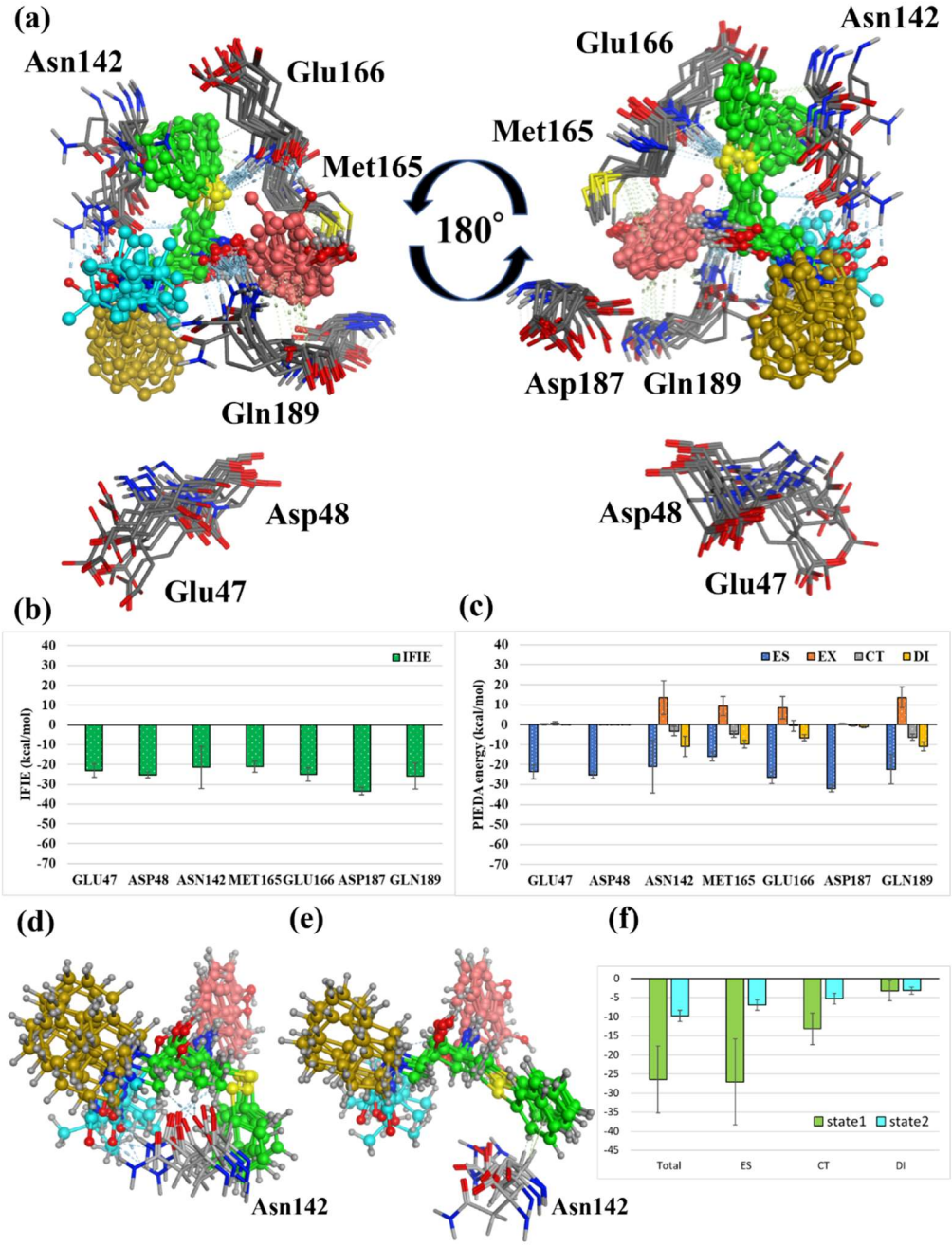

Figure 9. Interactions with surrounding amino acids in Pose 4. (a) Amino acid residues of main protease interacting with Nelfinavir in Pose 4. The four ligand fragments are colored F1: pink, F2: green, F3: yellow, and F4: light blue. (b)(c)

https://doi.org/10.26434/chemrxiv-2023-hrsvj ORCID: https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

The tert-butyl group of NLF, which was not found to interact with Mpro through the four poses, was compared with HIV-1 protease, the original target of NLF. In the complex structure of HIV-1 protease and NLF (PDBID:3el5), the binding interaction energy $\Delta E^{int(static)}$ was -155 kcal/mol, slightly more stable than Pose2. However, the tert-butyl group, which did not interact with Mpro, interacted strongly in the DI term with the surrounding hydrophobic amino acid residues (Ala28, Asp29, Asp30, Val32, Ile47, etc.) in the HIV-1 protease (Figure S2). In other words, all functional groups were utilized in the interaction with the original target protein. Thus, identifying such functional groups with less robust interactions is a candidate for functional group optimization and is an important suggestion for drug discovery using drug repositioning.

3.3.2 Ligand binding energy

Table 1 presents the results (ΔE^{int}) of calculating the average value and standard deviation of total IFIE from the FMO calculation results of each of the 100 structures for Poses 1–4. Comparing the dynamically averaged interaction energy ΔE^{int} with the static interaction energy $\Delta E^{int(static)}$ confirmed that Pose 3 was consistently the most stable. Moreover, ΔE^{int} showed that Pose 4 was the second most stable structure, followed by Pose 1, and the stability was reversed from the static state. This is because although Pose 1 had two hydrogen bonds and Pose 4 had one hydrogen bond in the static structure, and both poses had three bonds in the dynamic structure, the hydrogen bond dependence on the time was not maintained in Pose 1 and Pose 4 was in a stronger binding state. In the DI term in $\Delta E^{int(static)}$, Poses 1 and 2 were more stable than Poses 3 and 4 by approximately -20 kcal/mol. Conversely, in the DI term in ΔE^{int}, Poses 3 and 4 showed stable values ranging from -5–10 kcal/mol. Furthermore, in the static structure, there are four CH/π interactions in Pose 1, three in Pose 2, three in Pose 3, and three in Pose 4. However, in the dynamical structure, there are three CH/π interactions in Pose 1, two in Pose 2, two in Pose 3, and five in Pose 4. The number of interactions decreased in Poses 1 and 2 and increased in Pose 4, explaining why the DI term became more stable than Pose 1 and Pose 2. Moreover, Pose 3 had few CH/π interactions; however, its hydrophobic interactions with Gln189 and others stabilized the DI term.

Next, the deformation energy ($\Delta E_{\mathrm{lig}}^{\mathrm{def}}$) of ligands of each pose and the binding energy, $\Delta E^{int} + \Delta E_{\mathrm{lig}}^{\mathrm{def}}$, are presented in Table 1. Poses 1–3 had similar binding energy and Pose 4 had larger deformation energy than other poses. This is because Pose 4 has a wider shape than the other poses and is connected distortedly. Additionally, $\Delta E^{int} + \Delta E_{\mathrm{lig}}^{\mathrm{def}}$ were 3, 4, 1, and 2 in descending order. Since the deformation energy is the same as the magnitude relationship of ΔEint, it has little effect on the difference in ΔEint.

Furthermore, the desolvation energy (ΔE^{sol}) was determined. Using the method defined in Section 2.5, 25 amino acids within 4 Å average distance from ligands were selected. The solvation energies of Poses 3 and 4 were approximately 120 kcal/mol, twice more than those of Poses 1 and 2 (approximately 60 kcal/mol) (Table 1). Figures 10 (a) and (b) show the actual water molecules around ligands and binding sites in Poses 2 and 3. In Pose 2, the water molecules were trapped between the ligand and the protein in the pocket (Figure 10a). However, in Pose 3, the ligand was tightly bound to the protein, and almost no water molecules entered the binding site (Figure 10b). Notably, water molecules in the ligand-binding pocket reduce direct interactions between the ligand and the amino acid residues at the binding site. Therefore, ΔE^{int} was less stabilized and the desolvation energy (ΔE^{sol}) was small because less water was displaced by binding. Specifically, in Poses 1 and 2, many water molecules remained at the binding sites; hence, ΔE^{int} was less stabilized and ΔE^{sol} was small. However, in Poses 3 and 4, ΔE^{int} was greatly stabilized and ΔE^{sol} was large, indicating that the water molecules at the binding sites were excluded. Such direct interactions with water molecules cannot be handled using the FMO-PB calculation [29] or the continuum model PCM[49]. The result of this time involved manually arranging water molecules. ΔE_{bind} that has ultimately been obtained is presented in Table 1. In ΔE_{bind}, we suggested that 3, 1, 2, and 4 were the stable binding poses, in order of decreasing

https://doi.org/10.26434/chemrxiv-2023-hrsvj ORCID: https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

stability. Additionally, ΔE_{bind} in Poses 3, 1 and 2 was negative, but positive in Pose 4 and could not be stabilized by binding. This is attributed to the large energy loss due to desolvation and the large deformation energy.

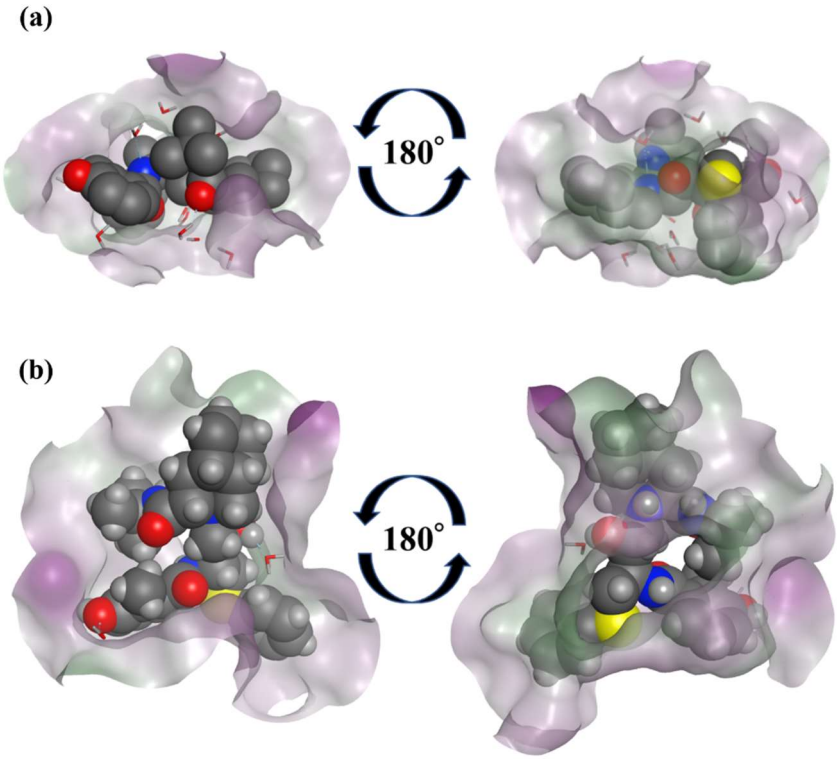

Figure 10. Water molecules in the ligand binding pocket. (a) Structure of Pose 2 at 100 ns and water molecules can be seen between the ligand and the protein. (b) Structure of Pose 3 at 100 ns, with almost no water molecules between the ligand and the protein.

Although the binding energy values differed depending on the pose, some amino acids interacting with the ligand were common in the four binding poses. Glu47, Asp48, and Asp187 acquired electrostatic interactions in all binding poses. Similarly, Glu166 acquired electrostatic interactions in all binding poses and hydrogen bonds or CH/π interactions depending on the pose. Moreover, Gln189 had a hydrogen bond and CH/π interaction in all poses; however, it fluctuated greatly depending on the pose. Asn142 was identified as a residue with hydrogen bond or CH/π interaction in all poses excluding Pose 2; however, it was a fluctuating interacting residue in all binding poses of Poses 1, 3, and 4. Furthermore, Met49 and Met165 showed a strong interaction in two poses, and Leu27 showed strong interaction in one pose. Gln189 maintained hydrogen bonds with the peptide-like backbone of F3 and F4 in Pose 3 and F1 in Pose 4. In contrast, in other poses, there were times when the nitrogen of the peptide-like backbone of F3 and F4 in Pose 1 and the OH group of the peptide-like backbone of F1 and F2 in Pose 2 made hydrogen bond bridge with Gln189, however, several of the sampling structures failed to retain the bond. Similarly, Asn142 was hydrogen-bonded with the peptide-like backbone of F1 in Pose 3 and F4 in Pose 4. The stability of ΔE^{int} in Poses 3 and 4 suggest that strong hydrogen bonding with the peptide-like backbone is one of the factors responsible for the stability of binding to Mpro.

For compounds that inhibit the Mpro, described in a study [25] that comprehensively analyzed the interaction of 110 complex ligand structures registered in FMODB, CH/π interactions with Asn142, hydrogen bonds with the main chain of Glu166, and CH/π interactions with the β-carbon are important. Furthermore, a study[13] on the crystal structure of Ensitrelvir (PDBID: 7VTH), which targets the Mpro, reported that Ensitrelvir hydrogen-bonds with the main chain NH of Glu166. Moreover, Asn142, identified as a residue with a fluctuating interaction in this study, had a missing atom in the side chain in the crystal structure of Ensitrelvir. Therefore, in the future, better binding properties can be acquired by designing compounds with stable

https://doi.org/10.26434/chemrxiv-2023-hrsvj ORCID: https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

interactions with Asn142. Although the results obtained using our method do not perfectly predict the binding pose itself, we were able to identify interacting residues important for ligand recognition.

4. Conclusions

We performed docking, MD, and FMO (MD + FMO) calculations to estimate the binding poses and clarify the binding properties of the Mpro and NLF complex, a drug repositioning candidate, whose crystal structure has not been solved. Furthermore, we identified the intermolecular interactions between the NLF functional groups and Mpro residues, which are important for ligand recognition. We ranked each pose by evaluating the binding energy. In addition to intermolecular interaction, the desolvation effect was important in this procedure. The exclusion of water molecules from the ligand binding pocket is believed to be important for Mpro binding to NLF. The calculation of desolvation energy by explicit solvent consumes enormous computational resources. However, the use of the Fugaku supercomputer has made it possible. In the analysis of the intermolecular interaction focusing on the NLF side, following interactions were crucial in all poses : the CH/π and π-π interactions between the π electrons of the benzene rings of F1 and F2 and surrounding residues and hydrogen bonding interactins with the OH group of F1 and surrounding residues. Moreover, the tert-butyl group of F4 was not involved in the binding; hence, it was considered unimportant for binding to the Mpro. However, F4 was an important functional group in binding to HIV-1 protease—the original target of NLF. Therefore, functional groups that became less important due to drug repositioning could be identified using our method. Moreover, the hydrogen bond between the peptide-like backbone and Mpro was considered important. This backbone portion formed hydrogen bonds with Gln189 in Poses 3 and 4. However, in Poses 1 and 2, hydrogen bonds were rarely formed at this site. Therefore, we inferred that the strong hydrogen bond at the peptide-like site is a factor in the binding stability of Mpro and NLF.

Additionally, on the Mpro side, amino acid residues important for binding (Glu47, Asp48, Glu166, Asp187, Gln189), common to all binding poses, could be identified. These amino acid residues were considered important in drug design. Furthermore, we suggested that it is important to have hydrogen bonds with peptide-like backbones and multiple π-rich functional groups, such as benzene rings, in interactions with these residues, particularly for NLF inhibition.

Our “MD+FMO” approach can predict the local stable structure of protein-ligand complexes considering the conformational fluctuations in the solvent and identifying important intermolecular interactions in the complexes. Quantitative interaction energies based on quantum chemical calculations generate ideas for subsequent molecular design. Our approach provides a new guideline for structure-based drug design starting from a candidate compound whose complex crystal structure has not been obtained.

5. Data and Software Availability

All structure files and a set of input/output files used for FMO calculations are available at the FMODB (https://drugdesign.riken.jp/FMODB/); FODBIDs are listed in Table S2. Simple data analysis can be performed using the FMODB 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 high-performance computing infrastructure machines (https://www.hpci-office.jp/pages/e_appli_abinit-mp).

Acknowledgments

The authors thank Dr. Tatsuya Nakano and Dr. Yoshio Okiyama of the Graduate School of System Informatics, Department of Computational Science, Kobe University, and Dr. Chiduru Watanabe of RIKEN Center for Biosystems Dynamics Research for general discussions related to FMO calculations. This study was partly supported by the Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research)

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

(BINDS) from the Japan Agency for Medical Research and Development (AMED) (Grant Number JP23ama121030). YM would also like to thank Rikkyo SFR for its financial support. This research was performed as an activity of the FMO drug design consortium (FMODD). The FMO calculations were performed using the Oakforest-PACS supercomputer (project ID: hp200146), Fugaku supercomputer (project ID: g9330001 and hp220143), and the TSUBAME3.0 supercomputer (Tokyo Institute of Technology, Tokyo, Japan).

References

(1) *WHO Coronavirus (COVID-19) Dashboard*. https://covid19.who.int (accessed 2023-06-15).

(2) Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y.; Yuan, M.-L.; Zhang, Y.-L.; Dai, F.-H.; Liu, Y.; Wang, Q.-M.; Zheng, J.-J.; Xu, L.; Holmes, E. C.; Zhang, Y.-Z. A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* **2020**, *579* (7798), 265–269. https://doi.org/10.1038/s41586-020-2008-3.

(3) Xu, X.; Han, M.; Li, T.; Sun, W.; Wang, D.; Fu, B.; Zhou, Y.; Zheng, X.; Yang, Y.; Li, X.; Zhang, X.; Pan, A.; Wei, H. Effective Treatment of Severe COVID-19 Patients with Tocilizumab. *Proceedings of the National Academy of Sciences* **2020**, *117* (20), 10970–10975. https://doi.org/10.1073/pnas.2005615117.

(4) Marconi, V. C.; Ramanan, A. V.; de Bono, S.; Kartman, C. E.; Krishnan, V.; Liao, R.; Piruzeli, M. L. B.; Goldman, J. D.; Alatorre-Alexander, J.; de Cassia Pellegrini, R.; Estrada, V.; Som, M.; Cardoso, A.; Chakladar, S.; Crowe, B.; Reis, P.; Zhang, X.; Adams, D. H.; Ely, E. W.; Ahn, M.-Y.; Akasbi, M.; Alatorre-Alexander, J.; Altclas, J. D.; Ariel, F.; Ariza, H. A.; Atkar, C.; Bertetti, A.; Bhattacharya, M.; Briones, M. L.; Budhraja, A.; Burza, A.; Camacho Ortiz, A.; Caricchio, R.; Casas, M.; Cevoli Recio, V.; Choi, W. S.; Cohen, E.; Comulada-Rivera, A.; Cook, P.; Cornejo Juarez, D. P.; Daniel, C.; Degrecci Relvas, L. F.; Dominguez Cherit, J. G.; Ellerin, T.; Enikeev, D.; Erico Tanni Minamoto, S.; Estrada, V.; Fiss, E.; Furuichi, M.; Giovanni Luz, K.; Goldman, J. D.; Gonzalez, O.; Gordeev, I.; Gruenewald, T.; Hamamoto Sato, V. A.; Heo, E. Y.; Heo, J. Y.; Hermida, M.; Hirai, Y.; Hutchinson, D.; Iastrebner, C.; Ioachimescu, O.; Jain, M.; Juliani Souza Lima, M. P.; Khan, A.; Kremer, A. E.; Lawrie, T.; MacElwee, M.; Madhani-Lovely, F.; Malhotra, V.; Martínez Resendez, M. F.; McKinnell, J.; Milligan, P.; Minelli, C.; Moran Rodriguez, M. A.; Parody, M. L.; Paulin, P.; Pellegrini, R. de C.; Pemu, P.; Procopio Carvalho, A. C.; Puoti, M.; Purow, J.; Ramesh, M.; Rea Neto, A.; Rea Neto, A.; Robinson, P.; Rodrigues, C.; Rojas Velasco, G.; Saraiva, J. F. K.; Scheinberg, M.; Schreiber, S.; Scublinsky, D.; Sevciovic Grumach, A.; Shawa, I.; Simon Campos, J.; Sofat, N.; Som, M.; Spinner, C. D.; Sprinz, E.; Stienecker, R.; Suarez, J.; Tachikawa, N.; Tahir, H.; Tiffany, B.; Vishnevsky, A.; Westheimer Cavalcante, A.; Zirpe, K. Efficacy and Safety of Baricitinib for the Treatment of Hospitalised Adults with COVID-19 (COV-BARRIER): A Randomised, Double-Blind, Parallel-Group, Placebo-Controlled Phase 3 Trial. *The Lancet Respiratory Medicine* **2021**, *9* (12), 1407–1418. https://doi.org/10.1016/S2213-2600(21)00331-3.

(5) Spinner, C. D.; Gottlieb, R. L.; Criner, G. J.; Arribas López, J. R.; Cattelan, A. M.; Soriano Viladomiu, A.; Ogbuagu, O.; Malhotra, P.; Mullane, K. M.; Castagna, A.; Chai, L. Y. A.; Roestenberg, M.; Tsang, O. T. Y.; Bernasconi, E.; Le Turnier, P.; Chang, S.-C.; SenGupta, D.; Hyland, R. H.; Osinusi, A. O.; Cao, H.; Blair, C.; Wang, H.; Gaggar, A.; Brainard, D. M.; McPhail, M. J.; Bhagani, S.; Ahn, M. Y.; Sanyal, A. J.; Huhn, G.; Marty, F. M.; GS-US-540-5774 Investigators. Effect of Remdesivir vs Standard Care on Clinical Status at 11 Days in Patients With Moderate COVID-19: A Randomized Clinical Trial. *JAMA* **2020**, *324* (11), 1048–1057. https://doi.org/10.1001/jama.2020.16349.

(6) Westendorf, K.; Žentelis, S.; Wang, L.; Foster, D.; Vaillancourt, P.; Wiggin, M.; Lovett, E.; van der Lee, R.; Hendle, J.; Pustilnik, A.; Sauder, J. M.; Kraft, L.; Hwang, Y.; Siegel, R. W.; Chen, J.; Heinz, B. A.; Higgs, R. E.; Kallewaard, N. L.; Jepson, K.; Goya, R.; Smith, M. A.; Collins, D. W.; Pellacani, D.; Xiang, P.; de Puyraimond, V.; Ricicova, M.; Devorkin, L.;

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Pritchard, C.; O'Neill, A.; Dalal, K.; Panwar, P.; Dhupar, H.; Garces, F. A.; Cohen, C. A.; Dye, J. M.; Huie, K. E.; Badger, C. V.; Kobasa, D.; Audet, J.; Freitas, J. J.; Hassanali, S.; Hughes, I.; Munoz, L.; Palma, H. C.; Ramamurthy, B.; Cross, R. W.; Geisbert, T. W.; Menachery, V.; Lokugamage, K.; Borisevich, V.; Lanz, I.; Anderson, L.; Sipahimalani, P.; Corbett, K. S.; Yang, E. S.; Zhang, Y.; Shi, W.; Zhou, T.; Choe, M.; Misasi, J.; Kwong, P. D.; Sullivan, N. J.; Graham, B. S.; Fernandez, T. L.; Hansen, C. L.; Falconer, E.; Mascola, J. R.; Jones, B. E.; Barnhart, B. C. LY-CoV1404 (Bebtelovimab) Potently Neutralizes SARS-CoV-2 Variants. *Cell Reports* **2022**, *39* (7), 110812. https://doi.org/10.1016/j.celrep.2022.110812.

(7) Akinosoglou, K.; Rigopoulos, E.-A.; Kaiafa, G.; Daios, S.; Karlafti, E.; Ztriva, E.; Polychronopoulos, G.; Gogos, C.; Savopoulos, C. Tixagevimab/Cilgavimab in SARS-CoV-2 Prophylaxis and Therapy: A Comprehensive Review of Clinical Experience. *Viruses* **2022**, *15* (1), 118. https://doi.org/10.3390/v15010118.

(8) Beigel, J. H.; Tomashek, K. M.; Dodd, L. E.; Mehta, A. K.; Zingman, B. S.; Kalil, A. C.; Hohmann, E.; Chu, H. Y.; Luetkemeyer, A.; Kline, S.; Lopez de Castilla, D.; Finberg, R. W.; Dierberg, K.; Tapson, V.; Hsieh, L.; Patterson, T. F.; Paredes, R.; Sweeney, D. A.; Short, W. R.; Touloumi, G.; Lye, D. C.; Ohmagari, N.; Oh, M.; Ruiz-Palacios, G. M.; Benfield, T.; Fätkenheuer, G.; Kortepeter, M. G.; Atmar, R. L.; Creech, C. B.; Lundgren, J.; Babiker, A. G.; Pett, S.; Neaton, J. D.; Burgess, T. H.; Bonnett, T.; Green, M.; Makowski, M.; Osinusi, A.; Nayak, S.; Lane, H. C. Remdesivir for the Treatment of Covid-19 — Final Report. *New England Journal of Medicine* **2020**, NEJMoa2007764. https://doi.org/10.1056/NEJMoa2007764.

(9) Nutho, B.; Mahalapbutr, P.; Hengphasatporn, K.; Pattaranggoon, 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* (18), 1769–1779. https://doi.org/10.1021/acs.biochem.0c00160.

(10) Cao, B.; Wang, Y.; Wen, D.; Liu, W.; Wang, J.; Fan, G.; Ruan, L.; Song, B.; Cai, Y.; Wei, M.; Li, X.; Xia, J.; Chen, N.; Xiang, J.; Yu, T.; Bai, T.; Xie, X.; Zhang, L.; Li, C.; Yuan, Y.; Chen, H.; Li, H.; Huang, H.; Tu, S.; Gong, F.; Liu, Y.; Wei, Y.; Dong, C.; Zhou, F.; Gu, X.; Xu, J.; Liu, Z.; Zhang, Y.; Li, H.; Shang, L.; Wang, K.; Li, K.; Zhou, X.; Dong, X.; Qu, Z.; Lu, S.; Hu, X.; Ruan, S.; Luo, S.; Wu, J.; Peng, L.; Cheng, F.; Pan, L.; Zou, J.; Jia, C.; Wang, J.; Liu, X.; Wang, S.; Wu, X.; Ge, Q.; He, J.; Zhan, H.; Qiu, F.; Guo, L.; Huang, C.; Jaki, T.; Hayden, F. G.; Horby, P. W.; Zhang, D.; Wang, C. A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19. *New England Journal of Medicine* **2020**, *382* (19), 1787–1799. https://doi.org/10.1056/NEJMoa2001282.

(11) Ohashi, H.; Watashi, K.; Saso, W.; Shionoya, K.; Iwanami, S.; Hirokawa, T.; Shirai, T.; Kanaya, S.; Ito, Y.; Kim, K. S.; Nomura, T.; Suzuki, T.; Nishioka, K.; Ando, S.; Ejima, K.; Koizumi, Y.; Tanaka, T.; Aoki, S.; Kuramochi, K.; Suzuki, T.; Hashiguchi, T.; Maenaka, K.; Matano, T.; Muramatsu, M.; Saijo, M.; Aihara, K.; Iwami, S.; Takeda, M.; McKeating, J. A.; Wakita, T. Potential Anti-COVID-19 Agents, Cepharanthine and NLF, and Their Usage for Combination Treatment. *iScience* **2021**, *24* (4), 102367. https://doi.org/10.1016/j.isci.2021.102367.

(12) Sargolzaei, M. Effect of NLF Stereoisomers on Coronavirus Main Protease: Molecular Docking, Molecular Dynamics Simulation and MM/GBSA Study. *Journal of Molecular Graphics and Modelling* **2021**, *103*, 107803. https://doi.org/10.1016/j.jmgm.2020.107803.

(13) Unoh, Y.; Uehara, S.; Nakahara, K.; Nobori, H.; Yamatsu, Y.; Yamamoto, S.; Maruyama, Y.; Taoda, Y.; Kasamatsu, K.; Suto, T.; Kouki, K.; Nakahashi, A.; Kawashima, S.; Sanaki, T.; Toba, S.; Uemura, K.; Mizutare, T.; Ando, S.; Sasaki, M.; Orba, Y.; Sawa, H.; Sato, A.; Sato, T.; Kato, T.; Tachibana, Y. Discovery of S-217622, a Noncovalent Oral SARS-CoV-2

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

3CL Protease Inhibitor Clinical Candidate for Treating COVID-19. *Journal of Medicinal Chemistry* **2022**, *65* (9), 6499–6512. https://doi.org/10.1021/acs.jmedchem.2c00117.

(14) Jourdan, J.-P.; Bureau, R.; Rochais, C.; Dallemagne, P. Drug Repositioning: A Brief Overview. Journal of Pharmacy and Pharmacology 2020, 72 (9), 1145–1151. https://doi.org/10.1111/jphp.13273.

(15) Hurle, M. R.; Yang, L.; Xie, Q.; Rajpal, D. K.; Sanseau, P.; Agarwal, P. Computational Drug Repositioning: From Data to Therapeutics. Clinical Pharmacology & Therapeutics 2013, 93 (4), 335–341. https://doi.org/10.1038/clpt.2013.1.

(16) 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.

(17) Mochizuki, Y.; Tanaka, S.; Fukuzawa, K. eds.; *Recent Advances of the Fragment Molecular Orbital Method*. **2021**, Springer Singapore, https://doi.org/10.1007/978-981-15-9235-5

(18) 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. *Journal of Computer Chemistry, Japan* **2015**, *advpub*, 2014–0039. https://doi.org/10.2477/jccj.2014-0039.

(19) Fedorov, D. G.; Kitaura, K. Subsystem Analysis for the Fragment Molecular Orbital Method and Its Application to Protein–Ligand Binding in Solution. *The Journal of Physical Chemistry A* **2016**, *120* (14), 2218–2231. https://doi.org/10.1021/acs.jpca.6b00163.

(20) Chiduru, W.; Hirofumi, W.; Kaori, F.; Lorien J. P.; Yoshio, O.; Hitomi Y.; Shigeyuki Y.; Hirofumi, N.; Shigenori T.; Teruki, H. Theoretical Analysis of Activity Cliffs among Benzofuranone-Class Pim1 Inhibitors Using the Fragment Molecular Orbital Method with Molecular Mechanics Poisson–Boltzmann Surface Area (FMO+MM-PBSA) Approach. *Journal of Chemical Information and Modeling* **2017** *57* (12), 2996–3010. https://pubs.acs.org/doi/10.1021/acs.jcim.7b00110

(21) Fukuzawa, K.; Kitaura, K.; Uebayasi, M.; Nakata, K.; Kaminuma, T.; Nakano, T. Ab Initio Quantum Mechanical Study of the Binding Energies of Human Estrogen Receptor α with Its Ligands: An Application of Fragment Molecular Orbital Method. *Journal of Computational Chemistry* **2005**, *26* (1), 1–10. https://doi.org/10.1002/jcc.20130.

(22) Fukuzawa, K.; Tanaka, S. Fragment Molecular Orbital Calculations for Biomolecules. *Current Opinion in Structural Biology* **2022**, *72*, 127–134. https://doi.org/10.1016/j.sbi.2021.08.010.

(23) 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. FMODB: The World's First Database of Quantum Mechanical Calculations for Biomacromolecules Based on the Fragment Molecular Orbital Method. *Journal of Chemical Information and Modeling* **2021**, *61* (2), 777–794. https://doi.org/10.1021/acs.jcim.0c01062.

(24) 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 Informatics Journal* **2019**, *19*, 5–18. https://doi.org/10.1273/cbij.19.5.

(25) Fukuzawa, K.; Kato, K.; Watanabe, C.; Kawashima, Y.; Handa, Y.; Yamamoto, A.; Watanabe, K.; Ohyama, T.; Kamisaka, K.; Takaya, D.; Honma, T. Special Features of COVID-19 in the FMODB: Fragment Molecular Orbital

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Calculations and Interaction Energy Analysis of SARS-CoV-2-Related Proteins. *Journal of Chemical Information and Modeling* **2021**, *61* (9), 4594–4612. https://doi.org/10.1021/acs.jcim.1c00694.

(26) 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. *Applied Physics Express* **2021**, *14* (2), 027003. https://doi.org/10.35848/1882-0786/abdac6.

(27) 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. *Journal of Physical Chemistry B* **2021**, *125* (24), 6501–6512. https://doi.org/10.1021/acs.jpcb.1c03043.

(28) Takaba, K.; Watanabe, C.; Tokuhisa, A.; Akinaga, Y.; Ma, B.; Kanada, R.; Araki, M.; Okuno, Y.; Kawashima, Y.; Moriwaki, H.; Kawashita, N.; Honma, T.; Fukuzawa, K.; Tanaka, S. Protein–Ligand Binding Affinity Prediction of Cyclin-Dependent Kinase-2 Inhibitors by Dynamically Averaged Fragment Molecular Orbital-Based Interaction Energy. *Journal of Computational Chemistry* **2022**, *43* (20), 1362–1371. https://doi.org/10.1002/jcc.26940.

(29) 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. *Journal of Physical Chemistry B* **2019**, *123* (5), 957–973. https://doi.org/10.1021/acs.jpcb.8b09326.

(30) Komeiji, Y.; Ishida, T.; Fedorov, D. G.; Kitaura, K. Change in a Protein's Electronic Structure Induced by an Explicit Solvent: An Ab Initio Fragment Molecular Orbital Study of Ubiquitin. *Journal of Computational Chemistry* **2007**, *28* (10), 1750–1762. https://doi.org/10.1002/jcc.20686.

(31) Fukuzawa, K.; Kurisaki, I.; Watanabe, C.; Okiyama, Y.; Mochizuki, Y.; Tanaka, S.; Komeiji, Y. Explicit Solvation Modulates Intra- and Inter-Molecular Interactions within DNA: Electronic Aspects Revealed by the Ab Initio Fragment Molecular Orbital (FMO) Method. *Computational and Theoretical Chemistry* **2015**, *1054*, 29–37. https://doi.org/10.1016/j.comptc.2014.11.020.

(32) Komeij, Y.; Okiyama, Y.; Mochizuki, Y.; Fukuzawa, K. Explicit Solvation of a Single-Stranded DNA, a Binding Protein, and Their Complex: A Suitable Protocol for Fragment Molecular Orbital Calculation. Chem-Bio Informatics Journal 2017, 17, 72–84. https://doi.org/10.1273/cbij.17.72.

(33) Komeiji, Y.; Okiyama, Y.; Mochizuki, Y.; Fukuzawa, K. Interaction between a Single-Stranded DNA and a Binding Protein Viewed by the Fragment Molecular Orbital Method. BCSJ 2018, 91 (11), 1596–1605. https://doi.org/10.1246/bcsj.20180150.

(34) Provorse Long, M. R.; Isborn, C. M. Combining Explicit Quantum Solvent with a Polarizable Continuum Model. J. Phys. Chem. B 2017, 121 (43), 10105–10117. https://doi.org/10.1021/acs.jpcb.7b06693.

(35) He, X.; Zhu, T.; Wang, X.; Liu, J.; Zhang, J. Z. H. Fragment Quantum Mechanical Calculation of Proteins and Its Applications. Acc. Chem. Res. 2014, 47 (9), 2748–2757. https://doi.org/10.1021/ar500077t.

(36) King, N. M.; Prabu-Jeyabalan, M.; Bandaranayake, R. M.; Nalam, M. N. L.; Nalivaika, E. A.; Özen, A.; Haliloğlu, T.; Yılmaz, N. K.; Schiffer, C. A. Extreme Entropy–Enthalpy Compensation in a Drug-Resistant Variant of HIV-1 Protease. *ACS Chemical Biology* **2012**, *7* (9), 1536–1546. https://doi.org/10.1021/cb300191k.

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

(37) *Molecular Operating Environment (MOE) | MOEsaic | PSILO*. https://www.chemcomp.com/Products.htm (accessed 2023-02-07).

(38) Fedorov, D. G.; Nagata, T.; Kitaura, K. Exploring Chemistry with the Fragment Molecular Orbital Method. *Physical Chemistry Chemical Physics* **2012**, *14* (21), 7562–7577. https://doi.org/10.1039/C2CP23784A.

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

(40) 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. *Journal of Chemical Information and Modeling* **2006**, *46* (1), 221–230. https://doi.org/10.1021/ci050262q.

(41) Fedorov, D. G.; Kitaura, K. Pair Interaction Energy Decomposition Analysis. *Journal of Computational Chemistry* **2007**, *28* (1), 222–237. https://doi.org/10.1002/jcc.20496.

(42) Mochizuki, Y.; Nakano, T.; Sakakura, K.; Okiyama, Y.; Watanabe, H.; Kato, K.; Akinaga, Y.; Sato, S.; Yamamoto, J.; 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: Enhanced Performance and Applicability; Mochizuki, Y., Tanaka, S., Fukuzawa, K., Eds.; Springer: Singapore, 2021; pp 53–67. https://doi.org/10.1007/978-981-15-9235-5_4.

(43) Mochizuki, Y.; Nakano, T.; Koikegami, S.; Tanimori, S.; Abe, Y.; Nagashima, U.; Kitaura, K. A Parallelized Integral-Direct Second-Order Møller–Plesset Perturbation Theory Method with a Fragment Molecular Orbital Scheme. *Theoretical Chemistry Accounts* **2004**, *112* (5), 442–452. https://doi.org/10.1007/s00214-004-0602-3.

(44) Mochizuki, Y.; Koikegami, S.; Nakano, T.; Amari, S.; Kitaura, K. Large Scale MP2 Calculations with Fragment Molecular Orbital Scheme. *Chemical Physics Letters* **2004**, *396* (4), 473–479. https://doi.org/10.1016/j.cplett.2004.08.082.

(45) Mochizuki, Y.; Yamashita, K.; Murase, T.; Nakano, T.; Fukuzawa, K.; Takematsu, K.; Watanabe, H.; Tanaka, S. Large Scale FMO-MP2 Calculations on a Massively Parallel-Vector Computer. Chemical Physics Letters 2008, 457 (4), 396–403. https://doi.org/10.1016/j.cplett.2008.03.090.

(46) Case D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A.W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, T. S. Lee, A. W.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B; Mermelstein, D.; Merz, K.M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman P. A. AMBER 2016. **2016,** University of California, San Francisco.

(47) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. Journal of Computational Chemistry 2006, 27 (15), 1787–1799. https://doi.org/10.1002/jcc.20495.

(48) 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* (6497), 1331–1335. https://doi.org/10.1126/science.abb4489.

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

(49) Fedorov, D. G.; Kitaura, K.; Li, H.; Jensen, J. H.; Gordon, M. S. The Polarizable Continuum Model (PCM) Interfaced with the Fragment Molecular Orbital Method (FMO). *Journal of Computational Chemistry* **2006**, *27* (8), 976–985. https://doi.org/10.1002/jcc.204006.

https://doi.org/10.26434/chemrxiv-2023-hrsvj **ORCID:** https://orcid.org/0000-0001-5357-8250 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0