FMOe: Preprocessing and visualizing package of the fragment molecular orbital method for Molecular Operating Environment and its applications in covalent ligand and metalloprotein analyses

Hirotomo Moriwaki [1,#], Yusuke Kawashima[2,#], Chiduru Watanabe[1,3,], Kikuko Kamisaka[1], Yoshio Okiyama[4], Kaori Fukuzawa[2,5], Teruki Honma[1]*

[1]Center for Biosystems Dynamics Research, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

[2]Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan

[3]JST PRESTO, 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan

[4]Department of Computational Science, Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe, Hyogo 657-8501, Japan

1

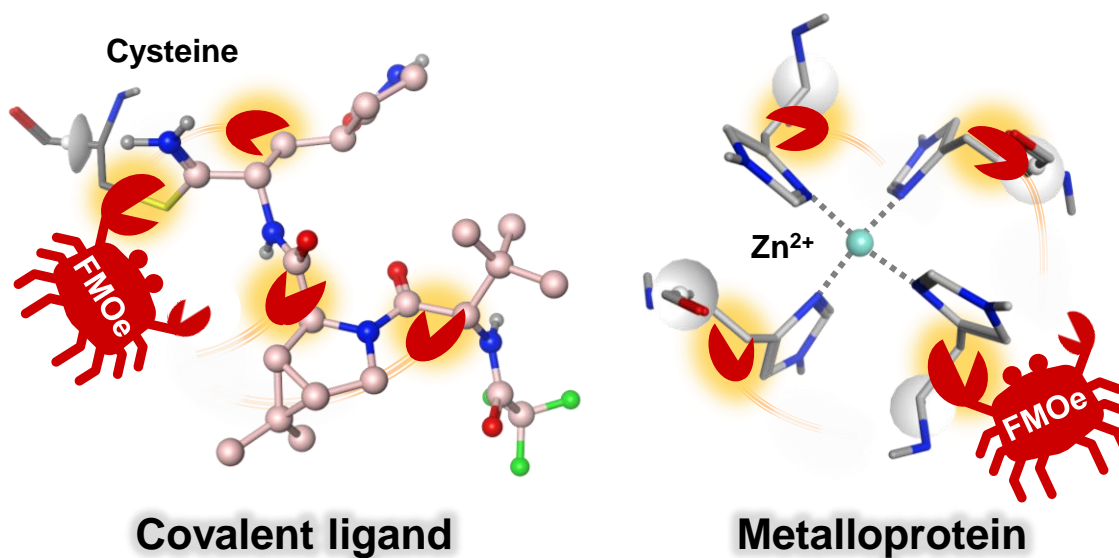[5] Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan

[#] equal contributions

* corresponding author

CW, Email address: chiduru.watanabe@riken.jp

Key Words: FMO, GUI, Pre/post-processing, ABINIT-MP, MOE, covalent ligand, metalloprotein

TOC (1.375 inches high × 3.5 inches)

ABSTRACT

The fragment molecular orbital (FMO) method is an efficient quantum chemical calculation technique for large biomolecules, dividing each into smaller fragments and providing inter-fragment interaction energies (IFIEs) that support our understanding of molecular recognition. The *ab initio* fragment MO method program (ABINIT-MP), an FMO processing software, can automatically divide typical proteins and nucleic acids. In contrast, small molecules such as ligands and hetero systems must be manually divided. Thus, we developed a graphical user interface to easily handle such manual fragmentation as a library for Molecular Operating Environment (MOE) that preprocesses and visualizes FMO calculations. We demonstrated fragmentation with IFIE analyses for the two following cases: 1) covalent cysteine–ligand bonding inside the SARS-CoV-2 main protease (M^{pro}) and nirmatrelvir (Paxlovid) complex, and 2) the metal coordination inside a zinc-bound cyclic peptide. IFIE analysis successfully identified the key amino acid residues for the molecular recognition of nirmatrelrvir with M^{pro} and the details of their interactions (e.g., hydrogen bonds and CH/π interactions) via ligand fragmentation of functional group units. In metalloproteins, we found an efficient and accurate scheme for the fragmentation of Zn^{2+} ions with four histidines coordinated to the ion. FMOe simplifies manual fragmentation, allowing users to experiment with various fragmentation patterns and perform in-depth IFIE analysis with high accuracy. In the future, our findings will provide valuable insight into complicated cases, such as ligand

fragmentation in modality drug discovery, especially for medium-sized molecules and metalloprotein fragmentation around metals.

1. INTRODUCTION

In the fragment molecular orbital (FMO) method developed by Kitaura *et al.*[1], a large biomolecule is divided into small fragments before performing molecular orbital (MO) calculations. After performing MO calculations on a single fragment (monomer) and fragment pairs (dimer), the result is used to obtain the energy of the entire system by two-body approximation. This approach makes it possible to perform simulations of large molecules within a reasonable calculation time. Furthermore, the FMO method facilitates the implementation of pair interaction energy decomposition analysis (PIEDA)[2–4], in which the inter-fragment interaction energy (IFIE) is divided into several components to obtain more detailed information, such as the profiles of compound–protein interactions. Therefore, the application of PIEDA to drug discovery has progressed over the years[5–9]. The FMO method is currently available in several solver programs as general atomic and molecular electronic structure systems (GAMESS)[10], *ab initio* fragment MO method programs (ABINIT-MP)[11,12], and parallelized *ab initio* calculation systems (PAICS)[13]. Fragmentation of a molecule system must be performed prior to the FMO calculation. In particular, when a covalent bond is fragmented, the electron pair in the covalent bond is assigned to one

4

of the fragments, with the atoms from which the electron pairs have been truncated called bond detached atoms (BDAs) and the atoms to which an electron pair is added called bond attached atoms (BAAs). In this instance, it is necessary to obtain information such as the atoms and formal charges of each fragment and the BDAs or BAAs for each divided covalent bond. Several graphical user interface (GUI) programs[14–17] are available for each solver to assist in such a complicated procedure.

The ABINIT-MP program itself has a convenient function that automatically performs fragmentation for a typical system, such as proteins and nucleic acids, according to predefined rules. In contrast, one needs a manual operation to split large-size compounds a system contains or perform a more detailed IFIE analysis in the unit of functional groups[5–9,18]. Currently, such heterogeneous fragmentation is only performed using the GUI software, BioStation Viewer[16]. However, its operability is room for improvement, such as selecting all atoms in the fragment. Molecular Operating Environment (MOE)[19] is a molecular modeling software. It has a built-in scientific vector language (SVL), a programming language that handles molecules, and users can make advanced extensions manually. Previously, we developed a library called "FMOe"[20] using MOE to provide a simple interface for creating input files and visualizing the analysis results for FMO calculations. FMOe has addressed the problems associated with the existing interfaces for manual fragmentation.

5

Recently, there has been a demand for functional group unit interaction analyses via the fragmentation of ligands from the perspective of structure-based drug design in the pharmaceutical industry. Using FMOe will meet these demands and simplify the handling of ligands with covalent bonds and medium-molecular-weight drug candidate compounds (e.g., large compounds and cyclic peptides) for the "diversification of modality." In addition, functionally important transition metal ions must be explicitly treated because they play an essential role in tertiary structure formation and catalysis in metal-containing proteins such as Zn fingers and enzymes. In the past, FMO calculations have often been performed with metals removed, prioritizing the convergency and ease of IFIE analysis. However, when discussing the results of FMO calculations with structural biologists, it has been requested that FMO calculations be performed using the net experimental structure, which still incorporates metals. Therefore, in this study, as an example of analysis using FMOe, we will execute FMO calculations in the following molecular systems: 1) a complex between the SARS-CoV-2 main protease (M^{pro}) and a tripeptide-like inhibitor, nirmatrelvir (brand name: Paxlovid, code name: PF-07321332)[21,22] covalently bonded with Cys145 of M^{pro}; and 2) a cyclic peptide with a Zn^{2+} ion as the structural center.

2. METHODS

2.1. FMO method

6

Herein, we briefly describe the *ab initio* FMO method. In this method, a large molecule or molecular cluster is divided into small fragments, and MO calculations are performed on each monomer and dimer fragment, thus yielding the properties of the entire system. The many-body effects were considered using the environmental electrostatic potentials. The total energies of the FMO calculations are given by Equation (1).

$$E_{\text{total}} = \sum_I E'_I + \sum_{IJ} \Delta\tilde{E}_{IJ}. \qquad (1)$$

Here, E'_I is the monomer energy without the environmental electrostatic potential; $\Delta\tilde{E}_{IJ}$ is the IFIE; and *I* and *J* are fragment indices. In addition, using the Møller–Plesset perturbation method[23], PIEDA was used to analyze the energy components of IFIE, $\Delta\tilde{E}_{IJ}$: the electrostatic interaction (ES), exchange repulsion (EX), charge transfer interaction with mixing terms (CT+mix), and dispersion interaction (DI), as shown in Equation (2).

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

To obtain interaction energy of several fragments for A part $\Delta\tilde{E}_J^A$, such as the inhibitor-binding energy, we summed the IFIEs of all pairs between the fragments for A part and a fragment *J* that called the "IFIE-sum"[24] and is given by the IFIEs shown in Equation (3).

$$\Delta\tilde{E}_J^A = \sum_{I=A} \Delta\tilde{E}_{IJ}. \qquad (3)$$

7

The electron densities of the FMO calculations[25], $\rho(\boldsymbol{r})$, are given by Equation (4).

$$\rho(\boldsymbol{r}) = \sum_I \rho_I(\boldsymbol{r}) + \sum_{IJ} \Delta\rho_{IJ}(\boldsymbol{r}). \qquad (4)$$

Here, ρ_I is the monomer electron density without the environmental electrostatic potential and $\Delta\rho_{IJ}$ is

the dimer electron density.

The details of the FMO calculation, such as structural preprocessing and fragmentation, are described

at the beginning of Section 2.3.

2.2. Pre-/post-processing interface of FMOe

FMOe was developed using MOE SVL and implemented as a preprocessing/calculation result

visualization function for the FMO calculation using ABINIT-MP software. The results of the FMO

calculation were handled with a checkpoint file (CPF) for visualization and were read by both FMOe and

BioStation Viewer.

2.2.1. Preprocessing function: automatic and manual fragmentation and input file generation

Figure 1 shows the interface for fragmentation. For structures that have undergone hydrogenation and

structural optimization (i.e., structures in which the bond order between atoms is appropriately set on the

8

MOE and the force field and partial charge are assigned), fragmentation can be performed using the following procedure:

When selecting the "*Fragmentation/Merge*" mode from the "*FMOe*" button (Figure 1a) on the right side of the MOE's primary interface window ("*MOE*" window), the "*FMOe Fragmentation*" window (Figure 1b) was launched. Fragmentation of the protein was performed automatically in units of amino acid residues. Each water molecule, ligand, and metal ion was treated as one fragment. In Figure 1a, the fragment point to be automatically divided is displayed on the viewer as a white disk, considering the orientation of BDA and BAA. A small white stud was observed on the side of the BAA. The functions used for fragmenting were also designed to be used in SVL. Therefore, automatic processing can be performed.

The covalent bond serial number to be divided (#), bond split types of automatic or manual fragmentation (T), and information on BDA and BAA are displayed in the "*Bonds*" table in the "*FMOe Fragmentation*" window (Figure 1b). Selecting the division site from the list allows the user to zoom into the site and delete the site between the BDA and BAA.

The serial number of the fragment to be divided (#), residue name (Res.), number of atoms in each fragment (nA), number of atomic orbitals (AO) in the selected basis set (default: 6-31G*), and several types of charge information (D: number of detached electron pairs, A: number of attached electron pairs,

9

and C: fragment formal charge) are displayed in the "*Fragments*" table of the "*FMOe Fragmentation*" window (Figure 1b). By clicking on each fragment in the "*Fragments*" table, all atoms in the fragment can be selected in the "*MOE*" window.

Suppose additional fragmentation, such as ligand fragmentation in the functional group unit, is desired. In that case, the following procedure can be used (Figure 1b and 1c): After clicking the "+" button, meaning the "*Manual Fragmentation*" (Figure 1b), the covalent bond can be divided manually by clicking the BDA and BAA in sequence (Figure 1c). If the sp^3 carbon atom used as the BDA was selected, the atom that was a candidate for BAA is displayed as a ball. A light blue disk represents the fragment point between the BDA and BAA to be manually divided, where a small light blue stud is considered the BAA side. The details of the ligand fragmentation are demonstrated in the SARS-CoV-2 M^{pro} inhibitor, nirmatrelvir, analysis example described in Section 3.1.

Furthermore, merging multiple fragments was possible using the "*Merge*" table in the "*FMOe Fragmentation*" window even for fragments once divided. The details of the fragment merge function are introduced in the metalloprotein analysis example described in Section 3.2.

After performing the fragmentation process correctly, the FMO calculation input file was generated for ABINIT-MP using the following procedure: When clicking the "*Generate*" button (Figure 1b) at the bottom of the "*FMOe Fragmentation*" window, the molecule coordinate file (protein data bank (PDB)),

10

which was required for the FMO calculation, and ABINIT-MP input file (the "ajf file") were generated.

In this process, one can not specify the FMO configuration, including the computational method and basis

set. However, it is possible to reflect the settings in the ajf file in production by loading a user-defined ajf

template containing those settings. Therefore, it is possible to deal flexibly with the incompatibilities of

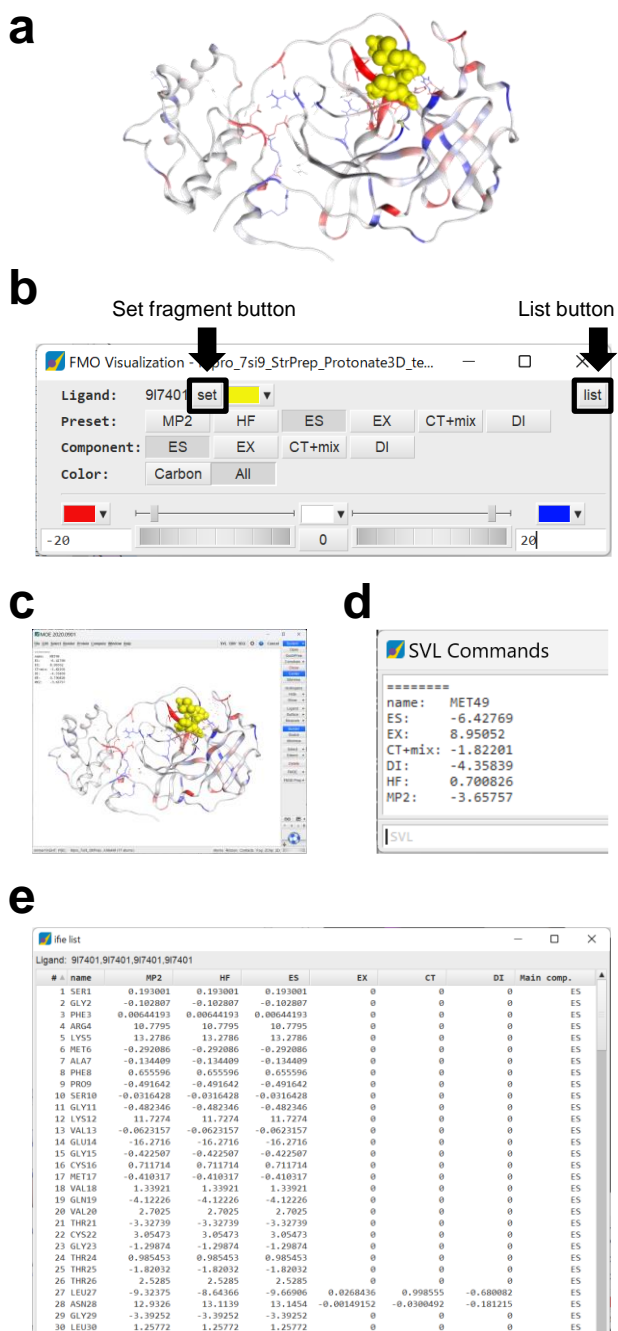configuration files due to differences in the ABINIT-MP versions.

11
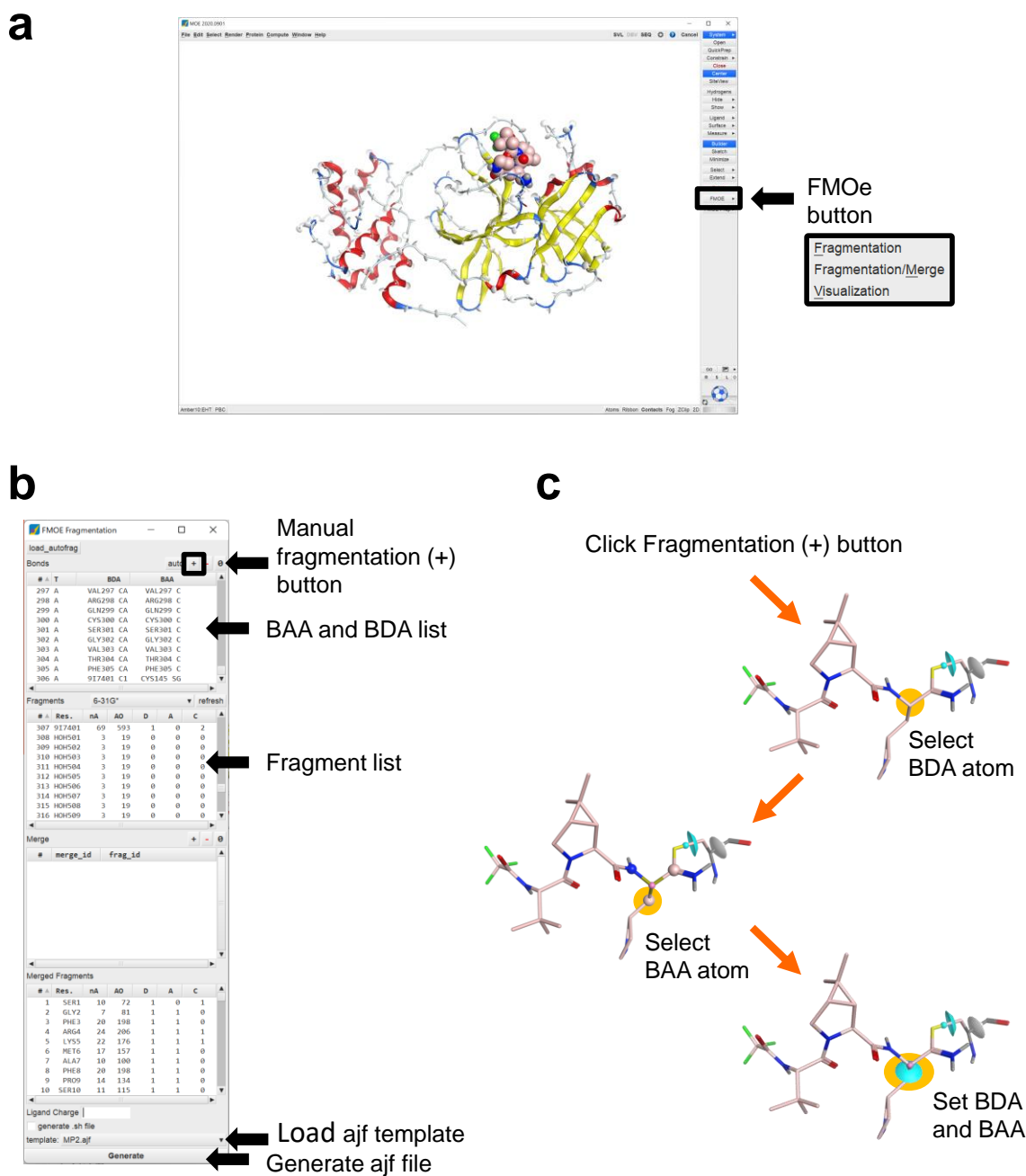
Figure 1. Manual fragmentation interface by the "*Fragmentation/Merge*" mode of FMOe. The

"*FMOe*" start button and the "*MOE*" window after automatic fragmentation processing (a), "*FMOe*

Fragmentation" window (b), and manual fragmentation processing procedure (c) are shown. The structure used was that of a SARS-CoV-2 M^{pro} and nirmatrelvir complex (PDB ID: 7SI9).

2.2.2. Post-processing function: IFIE/PIEDA analysis

Figure 2 shows the visualization interface for the results of FMO calculations. Post-processing for the IFIE/PIEDA analysis was performed as follows: The "*Visualization*" mode (Figure 1a) was selected using the "*FMOe*" button on the right side of the "*MOE*" window to read the CPF file. The CPF file describes the FMO calculation results executed using ABINIT-MP. IFIE/PIEDA analysis was performed automatically (Figure 2a). The fragment to be analyzed was set to the ligand molecule by default. The fragment to be analyzed; components of interaction energies, MP2, HF, ES, EX, CT+mix, and DI to be visualized; and the color scale range of the interaction energy was specified in the "*FMO Visualization*" window (Figure 2b). These operations immediately reflected the coloring of the interaction energy value of the selected component on the molecule in the viewer (Figure 2a). To change the fragment to be analyzed, at least one atom of the target fragment was selected in the "*MOE*" window. Then, the "*Set*" button, meaning the "*Set Fragment,*" was selected in the "*FMO Visualization*" window to reflect the results of the IFIE/PIEDA analysis of the target fragment. The IFIE or IFIE-sum of the target fragment with the fragment(s) selected on the screen is displayed in the upper left of the "*MOE*" window (Figure

13

2c) and is also shown in the "*SVL Commands*" window (Figure 2d). The "*List*" button (Figure 2b) can be

pressed at the top right of the "*FMO Visualization*" window to display a list of IFIE/PIEDA for the target

fragment on the screen (Figure 2e).

14

a

b

Set fragment button   List button

FMO Visualization - ...pro_7si9_StrPrep_Protonate3D_te...

Ligand: 9l7401 set
Preset: MP2   HF   ES   EX   CT+mix   DI
Component: ES   EX   CT+mix   DI
Color: Carbon   All

-20   0   20

c

d

SVL Commands

========
name: MET49
ES: -6.42769
EX: 8.95052
CT+mix: -1.82201
DI: -4.35839
HF: 0.700826
MP2:   -3.65757

SVL

e

ifie list

Ligand: 9l7401,9l7401,9l7401,9l7401

#	name	MP2	HF	ES	EX	CT	DI	Main comp.
1	SER1	0.193001	0.193001	0.193001	0	0	0	ES
2	GLY2	-0.102807	-0.102807	-0.102807	0	0	0	ES
3	PHE3	0.00644193	0.00644193	0.00644193	0	0	0	ES
4	ARG4	10.7795	10.7795	10.7795	0	0	0	ES
5	LYS5	13.2786	13.2786	13.2786	0	0	0	ES
6	MET6	-0.292086	-0.292086	-0.292086	0	0	0	ES
7	ALA7	-0.134409	-0.134409	-0.134409	0	0	0	ES
8	PHE8	0.655596	0.655596	0.655596	0	0	0	ES
9	PRO9	-0.491642	-0.491642	-0.491642	0	0	0	ES
10	SER10	-0.0316428	-0.0316428	-0.0316428	0	0	0	ES
11	GLY11	-0.482346	-0.482346	-0.482346	0	0	0	ES
12	LYS12	11.7274	11.7274	11.7274	0	0	0	ES
13	VAL13	-0.0623157	-0.0623157	-0.0623157	0	0	0	ES
14	GLU14	-16.2716	-16.2716	-16.2716	0	0	0	ES
15	GLY15	-0.422507	-0.422507	-0.422507	0	0	0	ES
16	CYS16	0.711714	0.711714	0.711714	0	0	0	ES
17	MET17	-0.410317	-0.410317	-0.410317	0	0	0	ES
18	VAL18	1.33921	1.33921	1.33921	0	0	0	ES
19	GLN19	-4.12226	-4.12226	-4.12226	0	0	0	ES
20	VAL20	2.7025	2.7025	2.7025	0	0	0	ES
21	THR21	-3.32739	-3.32739	-3.32739	0	0	0	ES
22	CYS22	3.05473	3.05473	3.05473	0	0	0	ES
23	GLY23	-1.29874	-1.29874	-1.29874	0	0	0	ES
24	THR24	0.985453	0.985453	0.985453	0	0	0	ES
25	THR25	-1.82032	-1.82032	-1.82032	0	0	0	ES
26	THR26	2.5285	2.5285	2.5285	0	0	0	ES
27	LEU27	-9.32375	-8.64366	-9.66906	0.0268436	0.998555	-0.680082	ES
28	ASN28	12.9326	13.1139	13.1454	-0.00149152	-0.0300492	-0.181215	ES
29	GLY29	-3.39252	-3.39252	-3.39252	0	0	0	ES
30	LEU30	1.25772	1.25772	1.25772	0	0	0	ES

Figure 2. Inter-fragment interaction energy (IFIE) analysis and pair interaction energy decomposition analysis (PIEDA) interface by the "*Visualization*" mode of FMOe. The IFIE (MP2) values (attractive interaction: red, repulsive interaction: blue) of the fragment (yellow) selected for

15

interaction analysis and the fragment of the exception are mapped on the molecular structure (a); the

"*FMO Visualization*" window allows for the specification of IFIE/PIEDA components and color

scales (b); the IFIE or IFIE-sum of the target fragment (yellow) with the fragment(s) selected on the

screen is displayed in the upper left of the "*MOE*" window (c) and is also shown in the "*SVL*

Commands" window (d); and the IFIE/PIEDA list can be obtained from the "*List*" button in the

"*FMO Visualization*" window (e). For the FMO data used, the FMO calculation results of the SARS-

CoV-2 M^{pro} and nirmatrelvir complex (PDB ID: 7SI9; FMODB ID: 4LQRN) were used.

3. RESULTS AND DISCUSSION

3.1 Detailed interaction analysis of functional units of covalent ligand using manual fragmentation

function

As an example of a relatively sizeable covalent inhibitor, FMO calculations were performed for the

complex between SARS-CoV-2 M^{pro} and nirmatrelvir (PDB ID: 7SI9)[21,22]. Nirmatrelvir is an orally active

SARS-CoV-2 M^{pro} inhibitor developed by Pfizer (Pty) Ltd. (USA). This covalent inhibitor binds directly

to the enzyme catalyst (Cys145) (Figures 3a and 3b). After FMO calculation, IFIE/PIEDA analysis was

performed to clarify the critical interaction between nirmatrelvir and each amino acid residue of M^{pro}.

First, the complex structure was pretreated via the following MOE procedure: the structure preparation

16

function was applied to structurally complement the missing heavy atoms on the amino acid residues; consequently, hydrogen atoms on the complex were added using the Protonate3D function. Subsequently, structural optimization was performed with the Amber10:EHT force field using the energy minimization function of MOE. The atom constraints during structural optimization were as follows: all heavy atoms registered in the PDB were fixed; the complemented heavy atoms on amino acid residues and all hydrogen atoms were unconstrained. After structural optimization, M^{pro} and nirmatrelvir (3-letter code of ligand name in PDB: 9I7) were divided into each amino acid residue and four fragments (9I7(1)–9I7(4)), respectively, according to the manual fragmentation procedure of FMOe (described in Section 2.1; Figure 3). The covalent bond site of Cys145 between M^{pro} and nirmatrelvir was fragmented into the main and side chains of cysteine. The FMO-MP2/6-31G* calculation was performed based on the input file obtained using FMOe. This FMO calculation result was registered in FMODB (https://drugdesign.riken.jp/FMODB/)[26,27] with the code (FMODB ID) 4LQRN.

Figures 3d–3g, S2a, S2c, S2e, and S2g show the IFIE analyses between each amino acid residue of SARS-CoV-2 M^{pro} and nirmatrelvir, and Figures S1a–S1p, S2b, S2d, S2f, and S2h show their PIEDA results, where the interaction energies with four ligand fragments of nirmatrelvir are represented (Figure

17

3c). For every ligand fragment (9I7(1)–9I7(4)), the tendency of the interaction energies with neighboring amino acid residues was different.

Figures 3d and S2a of the IFIE analysis show that the 9I7(1) fragment has the strongest attractive interaction with the Leu167 fragment (−22.1 kcal/mol). The other major attractive interacting fragments were Gln192 (−4.9 kcal/mol), Glu166 (−4.7 kcal/mol), Thr190 (−4.3 kcal/mol), Arg188 (−4.1 kcal/mol), and Gln189 (−3.8 kcal/mol). The reason for the strong interaction between the 9I7(1) and Leu167 fragment pair was suspected to be the NH–O hydrogen bond formed by the oxygen atom in the carbonyl group on the main chain of Glu166 (Leu167 fragment) and the hydrogen atom attached to a nitrogen atom in the amide group of 9I7(1). When verifying the PIEDA results (Figures S1a, S1i, and S2b), the ES component between the 9I7(1) fragment and the Leu167 fragment had a strong attractive interaction (−22.4 kcal/mol). In addition, the CT+mix component of its fragment pair, characteristic of hydrogen bond formation, exhibited a strong attractive interaction (−7.0 kcal/mol). In Gln189, the hydrogen atoms on the CA and CG carbon atoms in the main and side chains made CH–O hydrogen bonds with the oxygen atom of the amide of 9I7(1); a CH/π interaction was formed between the hydrogen atom on the trimethyl group of 9I7(1) and a π-orbital on the amide in the side chain of Gln189. Therefore, the CT+mix and DI components had weak attractive interactions of −2.0 and −3.9 kcal/mol, respectively (Figures S1a, S1i, and S2b). Here, 9I7(1) contains hydrophobic functional groups, such as trimethyl and trifluoromethane. Hence, the

interaction between some amino acid residues in the vicinity and the DI component was confirmed (Figures S1m and S2b).

From Figures 3e and S2c of the IFIE analysis, the 9I7(2) fragment was attracted to Glu166 (−7.0 kcal/mol), Glu189 (−4.1 kcal/mol), Met165 (−3.5 kcal/mol), Leu167 (−3.2 kcal/mol), and His41 (−3.0 kcal/mol). Therefore, it can be confirmed that the carbonyl oxygen atom of 9I7(2) and the hydrogen atom attached to a nitrogen atom in the main chain of Glu166 were hydrogen-bonded based on the ES and CT+mix components (−7.5 and −1.6 kcal/mol, respectively; Fig. S1b, S1j). The 9I7(2) fragment contained hydrophobic functional groups like methane. Hence, the interaction between some amino acid residues in the vicinity (e.g., His41 and Met165) and the DI component can be confirmed (Figures S1n and S2d). Furthermore, it was confirmed from the analysis of the DI component that a CH/π interaction was formed between the hydrogen atom on the methyl group of 9I7(2) and a π-orbital on the indole ring of His41 (DI component: −3.8 kcal/mol), as a characteristic interaction.

From Figures 3f and S2e of the IFIE analysis, the 9I7(3) fragment was attracted to the Asp187 (−32.4 kcal/mol), Glu166 (−25.8 kcal/mol), Met165 (−23.2 kcal/mol (including an oxygen atom in the main chain of His164)), HOH540 (−21.2 kcal/mol), His164 (−14.7 kcal/mol), Asn142 (−12.4 kcal/mol), Leu27 (−9.1 kcal/mol), and His41 (−8.8 kcal/mol). The ES component of PIEDA (Figures S1c and S2f) shows the presence of strong electrostatic interactions of the 9I7(3) fragment containing a charged imine with

19

charged amino acid residues, polar amino acid residues near 9I7(3), and water (HOH540). However, when the CT+mix component of PIEDA was confirmed (Figures S1g and S2f), only the Met165 fragment and water (HOH540) exhibited attractive interactions. This result indicates that the CT+mix component can ensure the presence of hydrogen bond-forming residues; that is, the hydrogen atom attached to a nitrogen atom of the amide group of 9I7(3) and the oxygen atom in the main chain of His164 (Met165 fragment) were NH–O hydrogen-bonded. Therefore, the Met165 fragment exhibited attractive interactions. In addition, analysis of the CT+mix components revealed that the NH of the 9I7(3) charged imine and the oxygen atom of the water molecule (HOH540) were NH–O hydrogen-bonded.

Figures 3g and S2g of the IFIE analysis show that the 9I7(4) fragment was attracted to His163 (−16.4 kcal/mol), Glu166 (−13.4 kcal/mol), Leu141 (−7.0 kcal/mol), Met165 (−3.9 kcal/mol), and Leu167 (−3.4 kcal/mol). This result can be confirmed from the ES and CT+mix components of PIEDA (Figures S1d, S1l, and S2h); the carbonyl oxygen atom on the 2-pyrrolidone of 9I7(4) and the hydrogen atom attached to the nitrogen atom on the imidazole ring in the side chain of His163 were NH–O hydrogen-bonded. In addition, the hydrogen atom attached to the nitrogen atom on the 2-pyrrolidone of 9I7(4) and the oxygen atom on the carboxylic acid in the side chain of Glu166 were also NH–O hydrogen-bonded. From the DI component of PIEDA (Figures S1p and S2h), it was inferred that the hydrogen atom attached to the carbon atoms on the alkyl group of 2-pyrrolidone and the π-orbitals on the amide groups of the amino acid

20

residues around the inhibitor (e.g., the main chain between Met165 and Glu166 (DI: −6.2 kcal/mol) and the side chain of Asn142 (DI: −2.6 kcal/mol)) exhibited CH/π interactions.

In summary, IFIE/PIEDA analysis with ligand fragmentation enabled us to identify the critical amino acid residues for the binding of nirmatrelvir to SARS-CoV-2 M^{pro} and their types of interactions on a functional group of the ligand.
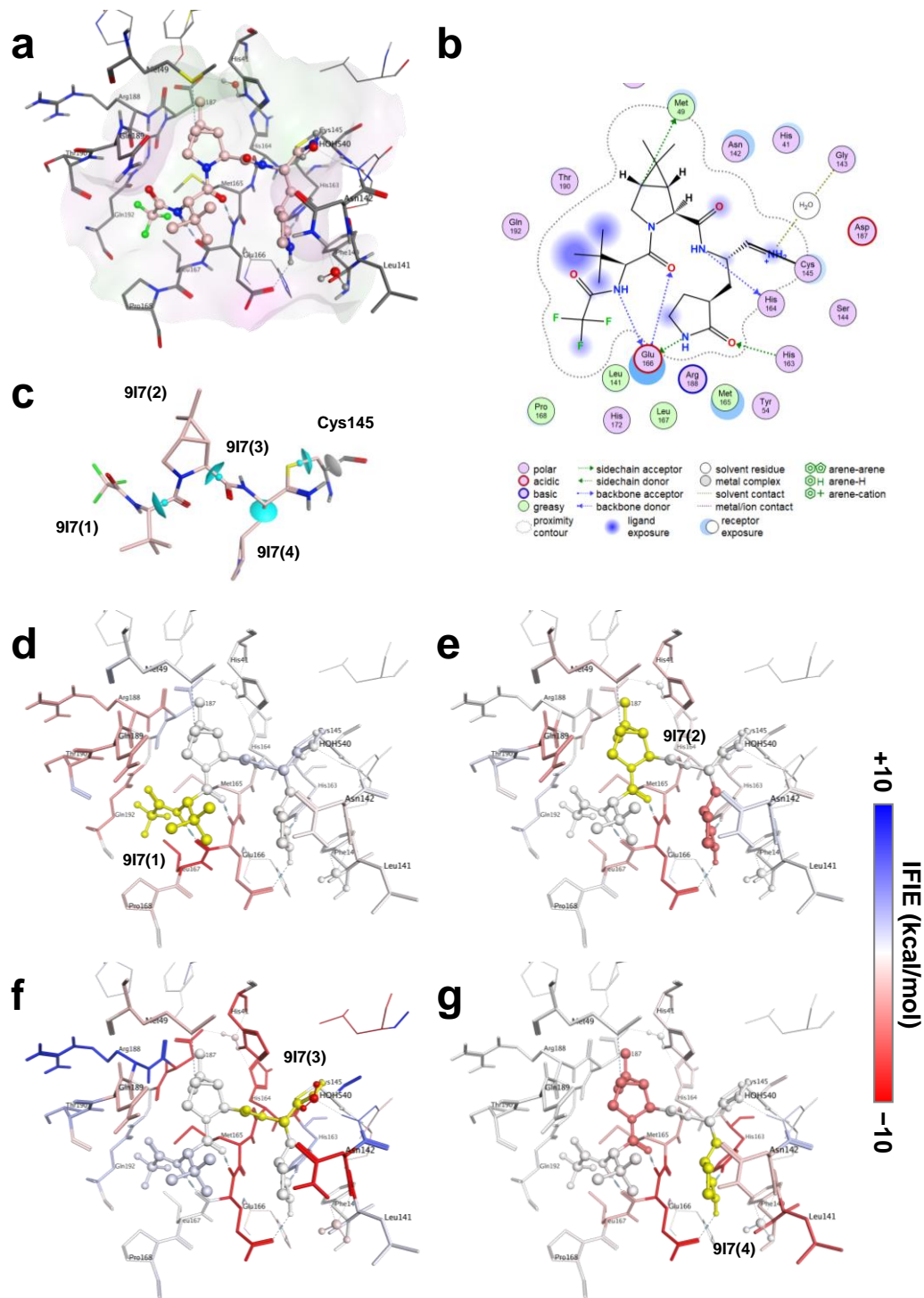
Figure 3. Interaction analysis between SARS-CoV-2 M^pro and nirmatrelvir. The ligand-binding

pocket of nirmatrelvir of SARS-CoV-2 M^pro is represented by the gray stick model for amino acid

22

residues and the pink ball and stick model for nirmatrelvir (a). Coordinate-based ligand interaction (2D) diagram (b). Fragment points of nirmatrelvir and Cys145 (c). The light blue disk indicates the bond detached atom (BDA)–bond attached atom (BAA), and the atoms on the side with the small light blue studs are BAAs (c). IFIE analysis of the ligand fragments (9I7(1)–9I7(4)) and the amino acid residues of SARS-CoV-2 M$^{\text{pro}}$ is shown (d)–(g).

3.2 Accuracy verification of FMO calculation for Zn-containing proteins using the fragment merge function

An example in which fragmentation processing is complicated is a molecular system containing a transition metal ion (e.g., Zn^{2+} and Mg^{2+}) or an ion (e.g., Ca^+ and Na^+). For example, the structure of the 6UFA entry registered in the PDB contains a Zn^{2+} ion at the center of the cyclic peptide, where the Zn^{2+} ion forms a coordination bond with the four surrounding histidines: His6, D-His12, His18, and D-His24 (Figure 4a). The cyclic peptide has the sequence KLqeXHklQEXhKLqeXHklQEXh, where X represents α-aminoisobutyric acid (AIB), and upper- and lower-case letters indicate L- and D-amino acids, respectively. Treating a Zn^{2+} ion as a single fragment worsens the self-consistent field convergence of monomers of the ion, the amino acid residue fragments coordinated with it, and dimers containing the ion[28]. This is because the proper electronic state cannot be solved, as the coordination bond is broken.

23

Thus, treating the Zn^{2+} ion and the histidines in the vicinity coordinated with the ion as one fragment is necessary to avoid breaking the coordination bond. However, a large fragment will be formed if this merging process is performed on a fragment of a typical amino acid residue unit[28]; this fragmentation scheme is referred to here as "main chain fragmentation" (Figure 4b). The total computational cost of whole protein increases even if one large fragment is included. Therefore, it is practical to reduce the cost of calculation by dividing the histidine fragment coordinated with the Zn^{2+} ions into the main and side chains and merging only the side chain with the Zn^{2+} ions (Figures 4c and 4d); the main/side chain (CB–CA) fragmentation, as seen in Figure 4c, is a method that has often been performed in the past[28–30], while the main/side chain (CB–CG) fragmentation, as seen in Figure 4d, is the first attempt to validate the accuracy of metalloprotein fragmentation. However, because this fragmentation process is very complicated to handle manually, we implemented the merge function to be efficiently executed by FMOe.

First, the structure was pretreated using the following procedure. This cyclic peptide structure (PDB ID: 6UFA) is a high-resolution X-ray crystal structure (0.77 Å resolution). The structure had no missing residues, and most of the hydrogen atoms for the cyclic peptide molecule were registered. However, hydrogen atoms are not present in water molecules. Therefore, the hydrogen atoms on the cyclic peptide were used as is, whereas only water molecules were hydrogenated using the AddH function of the MOE. Next, structural optimization was performed with the Amber10:EHT force field using energy

24

minimization via the MOE. The structure was highly reliable because these data were obtained using high-resolution X-ray crystallography. Therefore, the atoms in the molecular system were constrained during structure optimization under the following conditions. All atoms registered in the PDB were fixed, and only the complementary hydrogen atoms on the water molecules were unconstrained.

For the complex structure after structure optimization, side-chain fragmentation of histidine coordinated with Zn^{2+} ions was performed according to the manual fragmentation procedure described in Section 2.1 (Figures 4c and 4d). Subsequently, for the fragment to be merged, at least one atom was selected from each fragment, comprising the five fragments containing Zn^{2+} and four side chains of the histidines (Figure 5a). Finally, by pressing the "+" button, meaning the "*Merge Fragment*," on the right side of the "*Merge*" table in the "*FMOe Fragmentation*" window, the fragments containing the selected atoms were merged into one fragment (Figure 5a). Fragment information after merging can be found in the "*Merged Fragments*" table in the "*FMOe Fragmentation*" window (Figure 5b). The atomic details comprising the fragment can be confirmed using the following procedure. When the relevant fragment in the "*Merged Fragments*" table was selected, atoms that make up the fragment were selected on the molecular structure of the "*MOE*" window, and the atomic information of the fragment could be obtained (Figure 5b). A similar procedure can merge the Zn^{2+} ion and four histidines, His6 (HIS6), His18 (HIS18), D-His12 (DHI12), and D-His24 (DHI24), after main-chain fragmentation (Figure 4b). Once the fragment merge

25

was complete, a set of FMO calculation input files was created by clicking the "*Generate*" button at the bottom of the "*FMOe Fragmentation*" window (Figure 5b). The FMO-MP2/6-31G and FMO-MP2/6-31G* calculations were performed using the obtained FMO calculation input file. The results of the FMO calculation at MP2/6-31G for the main chain (Figure 4b), main/side chain (CB–CA) (Figure 4c), and main/side chain (CB–CG) (Figure 4d) fragmentation were registered in FMODB[26,27] with the codes (FMODB ID) VKK41, YNN22, and 5NRZZ, respectively; those at MP2/6-31G* for the main chain, main/side chain (CB–CA), and main/side chain (CB–CG) fragmentation were registered in FMODB[26,27] with the codes (FMODB ID) 5NRZZ, GYYM1, and 166QZ, respectively.
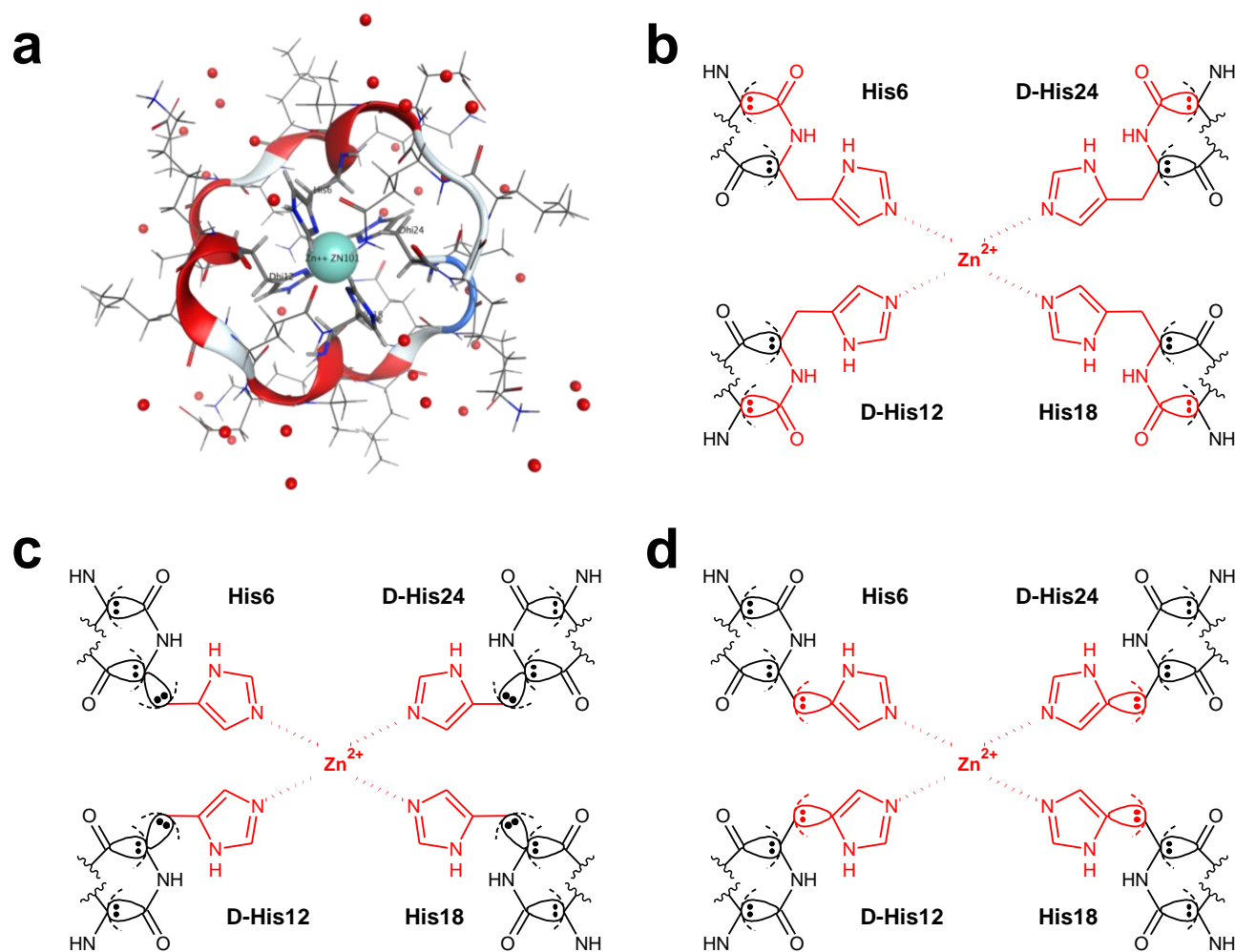
Figure 4. Fragmentation around Zn^{2+} of a cyclic peptide coordinated with Zn^{2+}. The three-dimensional structure of the cyclic peptide (PDB ID: 6UFA) is illustrated (a). The Zn^{2+} ion is a CPK model, and the water molecules are a ball model. The four histidines coordinated with the Zn^{2+} of the cyclic peptide are represented by a stick model, and other amino acid residues are represented by line models. The main chain fragmentation divides the CA carbon atom in the main chain of amino acid residues as a bond-detached atom (BDA) and the C carbon atom in the main chain as a bond-attached

27

atom (BAA) (b). The main/side chain (CB–CA) fragmentation splits the CA carbon atom in the main chain of amino acid residues as a BDA and the C carbon atom in the main chain as a BAA; in addition, the CB carbon atom in the side chain of the amino acid residue is divided as a BDA, and the CA carbon atom in the main chain is divided as a BAA (c). The main/side chain (CB–CG) fragmentation divides the CA carbon atom in the main chain of amino acid residues as a BDA and the C carbon atom in the main chain as a BAA; in addition, the CB carbon atom in the side chain of the amino acid residue is divided as a BDA, and the CG carbon atom in the main chain is divided as a BAA (d). The atoms shown in red are merged into the Zn^{2+} fragment.

28

Figure 5. Merge fragment flow. After automatic fragmentation, the atoms contained in the fragments to be merged are selected, and the "+" button, meaning the "*Merge Fragment*," in the "*FMOe Fragmentation*" window is clicked (a). The "*Merge*" table in the "*FMOe Fragmentation*" window and the "*Merged Fragments*" table reflect the fragment information after fragment merging (b).

29

The FMO calculation times, total energies of the whole molecular system, PIEDA of the Zn^{2+} ion fragment, atomic charges of the Zn^{2+} ion and the four histidines, His6 (HIS6), His18 (HIS18), D-His12 (DHI12), and D-His24 (DHI24), coordinated with the Zn^{2+} ion, and the FMO-based electron densities were investigated to verify the calculation efficiency and accuracy of the Zn^{2+} ion fragmentation schemes. Table 1 lists the FMO calculation times for the three fragmentations. By fragmentation between the main and side chains, it can be confirmed that the calculation time was significantly reduced to 1/6 and 1/10 for the MP2/6-31G and MP2/6-31G* calculations, respectively, compared to the main chain fragmentation. In addition, the result of the FMO calculation for the main chain fragmentation, which has fewer fragmentation points, was used as a reference value for a high-precision calculation when verifying the accuracy of the physical quantities of the two side-chain fragmentations. In Table 2, the total energies of the main/side chain fragmentation schemes are compared with the results of the main-chain fragmentation. As a result, at each calculation level (HF and MP2 methods), we verified that the differences from the total energies from the main chain fragmentation were less than 0.90 and 0.36 hartree in the main/side chain (CB–CA) and main/side chain (CB–CG) fragmentations, respectively.

https://doi.org/10.26434/chemrxiv-2024-2xk5n **ORCID:** https://orcid.org/0000-0002-0742-3896 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Table 1 Total time of FMO calculations at MP2/6-31G and MP2/6-31G* levels for three fragmentation schemes, the main chain, the main/side chain (CB–CA), and the main/side chain (CB–CG), using 48 cores (Intel(R) Xeon(R) CPU E5-2670 v3 @ 2.30GHz).

Fragmentation	Total time (s)	
	MP2/6-31G	MP2/6-31G*
Main chain	22867.9	190394.6
Main/Side chain (CB–CA)	4152.3	22117.6
Main/Side chain (CB–CG)	3476.3	19379.9

Table 2 Total energies of FMO calculations at MP2/6-31G and MP2/6-31G* levels for main chain, main/side chain (CB–CA), and main/side chain (CB–CG) fragmentations.

Fragmentation	Total energy (hartree)		Diff. from main chain fragmentation data (hartree)	
	HF/6-31G	MP2/6-31G	HF/6-31G	MP2/6-31G
Main chain	−14273.134259	−14297.832157	-	-
Main/Side chain (CB–CA)	−14272.256573	−14296.936548	0.877686	0.895609
Main/Side chain (CB–CG)	−14272.792494	−14297.470763	0.341765	0.361394
Fragmentation	Total energy (hartree)		Diff. from main chain fragmentation data (hartree)	
	HF/6-31G*	MP2/6-31G*	HF/6-31G*	MP2/6-31G*
Main chain	−14278.153634	−14314.526500	-	-
Main/Side chain (CB–CA)	−14277.213611	−14313.567646	0.940023	0.958854
Main/Side chain (CB–CG)	−14277.784439	−14314.142674	0.369196	0.383826

As seen in Figures 6 and S3, interaction energy analysis with PIEDA of the Zn^{2+} ion fragment containing the four histidines was performed in the main chain fragmentation units to examine the accuracy of the fragmentation schemes in terms of interactions. Figures 6a, S3a, and S3b show that the qualitative trends of IFIE/PIEDA were consistent with the main chain fragmentation and the main/side chain fragmentations. As shown in Figures 6b and 6c, the differences in interaction energy from the main chain fragmentation to the two main/side chain fragmentation datasets were within ca. 3 kcal/mol. Next, from the FMO calculation results of the main/side chain (CB–CA) and main/side chain (CB–CG) fragmentations, Figure S4 shows the PIEDA between the fragments containing Zn^{2+} ions and each of the other fragments. The trends of the ES, EX, CT+mix, and DI components were almost identical for both fragmentation schemes. However, there was an exception in the CT+mix component of the four fragments, which showed a strong repulsive interaction of approximately +100 kcal/mol in the main/side chain (CB–CA) fragmentation. Here, we focused on these fragment pairs for the following reasons: The bare CT interaction energy should have negative values in nature. In fact, statistical analysis of IFIE/PIEDA for FMODB confirmed that for fragment pairs forming typical interactions such as hydrogen bonds, XH/π, and ion pairs, the CT+mix terms tend to be essentially attractive interaction energies in our previous work[27]. The CT+mix values between the four fragments and the Zn^{2+} ion fragment obviously differ from the tendency. The four fragments were the amino acid residues (Lys1, D-Lys7 (DLY7), Lys13, and D-Lys19 (DLY19))

32

immediately after the histidine that were merged into the Zn^{2+} ion fragment. Notably, the CT+mix components of the four fragments in the main/side chain (CB–CA) fragmentation showed the CT+mix components of +100 kcal/mol. In contrast, the CT+mix components of the four fragments in the main/side chain (CB–CG) fragmentation were +30 kcal/mol, and the repulsive interactions were weaker than those in the main/side chain (CB–CA) fragmentation. These repulsions of the CT+mix components affect the accuracy of the total energy. This outcome may be due to the proximity of the BDAs of the CA carbon atom in the main chain and the CB carbon atom in the side chain. Therefore, it is recommended to exclude these fragment pairs from the IFIE/PIEDA analysis in the main/side chain fragmentation. This recommendation is because these fragment pairs show unusual values of the CT+mix terms due to fragment pairs with the BDAs next to each other. Similarly, in the conventional IFIE analysis of protein, IFIE values between covalent bond fragments, i.e., fragment pairs in which the BDA is next to the BAA, are excluded from the analysis due to unusual IFIE values caused by bond detachment in the FMO calculation[31,32].

33
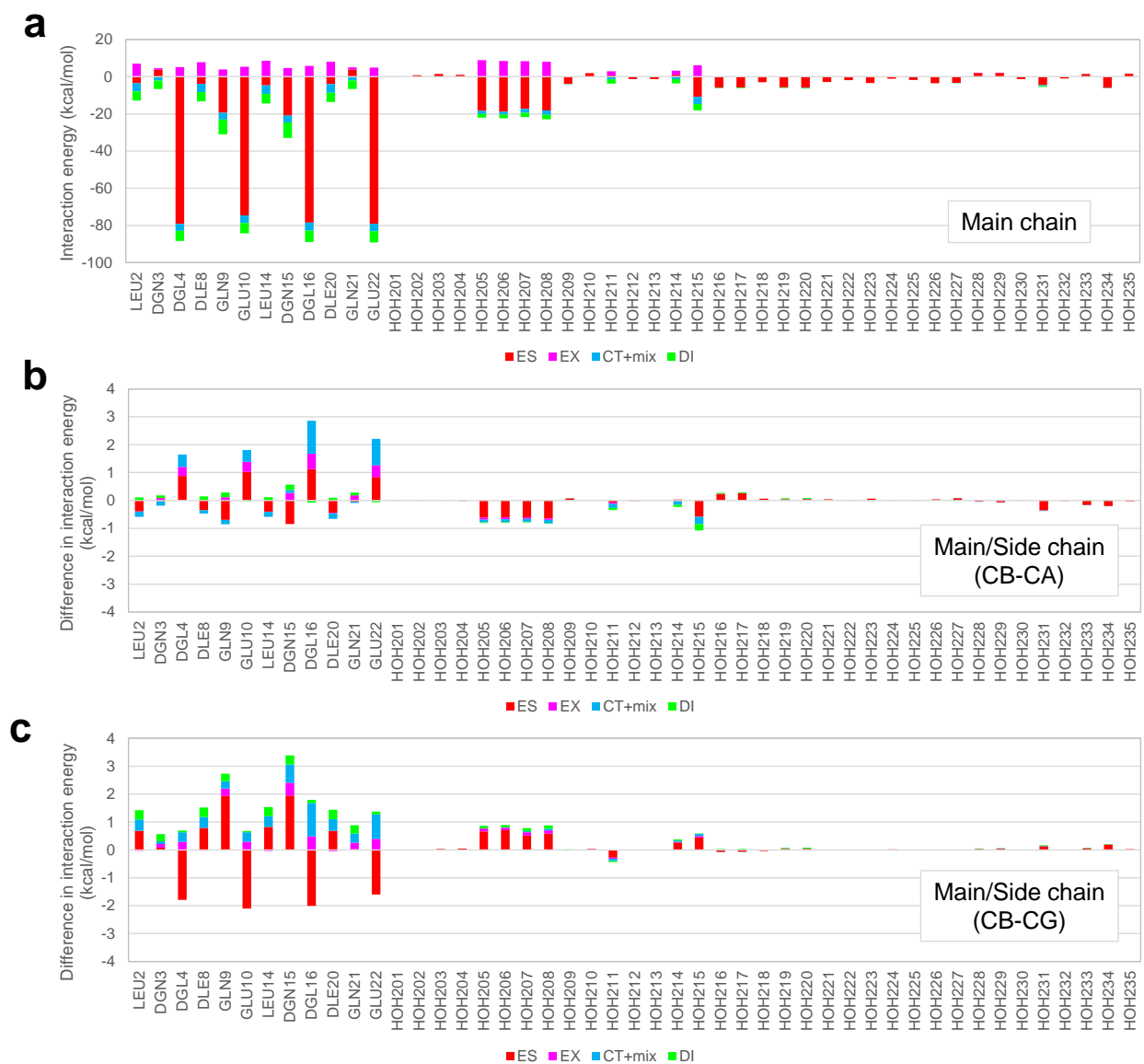
Figure 6. Interaction energy analysis of Zn^{2+} ion fragment, including the four histidines for the cyclic peptide containing Zn^{2+} (PDB ID: 6UFA). Interaction energies with PIEDA for main chain fragmentation data (FMODB ID: 8NN1Y) are shown in (a). Differences in interaction energies between the main chain fragmentation data and the main/side chain (CB–CA) fragmentation

34

(FMODB ID: GYYM1) and the main/side chain (CB–CG) fragmentation (FMODB ID: 166QZ) are shown in (b) and (c), respectively. The fragment units in the interaction analysis were adjusted to the main chain fragmentation scheme to compare the data with the main chain fragmentation data.

Tables S1 and S2 summarize the atomic charges of the Zn^{2+} ion and the four histidines, His6 (HIS6), His18 (HIS18), D-His12 (DHI12), and D-His24 (DHI24), based on the FMO calculation results (HF/6-31G and HF/6-31G* levels). Note that the atomic charge trends between the data using the 6-31G and 6-31G* basis functions were consistent. Figures S5–S7 show the net atomic charges for three fragmentation schemes and the differences in the atomic charges between the main chain fragmentation and the main/side chain fragmentation data. As shown in Table S2 and Figures S5 and S6, the trends of the Mulliken and natural population analysis charges were consistent regardless of the fragmentation method. Moreover, most of the atoms have a difference of ca. $0.05e$ or less when compared to the atomic charge of the main-chain fragmentation method. As an exception, the C, CB, CG, and CD2 carbon atoms of each histidine in the main/side chain (CB–CA) fragmentation and the CA, C, CB, and CD2 carbon atoms of each histidine in the main/side chain (CB–CG) fragmentation had a difference of ca. $0.05e$ or more. They were the BDAs, BAAs, or their neighbors. In contrast, the value of the Merz–Kollman charge varied significantly depending on the fragmentation method from Table S2 and Figure S7. It was inferred that this outcome

35

was due to the fitting performed to reproduce the electrostatic potential and the difference in the atoms constituting the fragment. Therefore, in the future, when assigning an atomic charge using an electrostatic potential charge (e.g., RESP or Merz–Kollman) to construct an FMO force field, careful attention should be paid to the fragmentation treatment of the metal-containing proteins.

Finally, the FMO-based electron density was analyzed to verify that the electronic structure calculations were accurate. Figures S8a–S8c show the FMO-based electron densities ρ of the cyclic peptides obtained using the main chain fragmentation scheme; Figures S8d–S8i show the differences in the electron density $\Delta\rho$ between the main chain and the main/side chain fragmentation data. We compared the FMO-based electron density of the main chain fragmentation with that of the main/side chain fragmentation using $\Delta\rho$. At a high electron density level (ρ and $\Delta\rho = 0.05e$), there was no significant difference in electron density around the Zn^{2+} ion between the main chain and the two main/side chain fragmentations (Figures S8d and S8g). At low levels of electron density (ρ and $\Delta\rho = 0.01e$ or $0.005e$), slight and or significant differences were observed in the electron density between the main chain and main/side chain fragmentations around the BDAs and BAAs of the four histidines (Figures S8e, S8f, S8h, and S8i). The main/side chain (CB–CG) fragmentation differed less from the main/side chain (CB–CA) fragmentation. In the case of the main/side chain (CG–CA) fragmentation, the spatial variation in electron density was more widely affected than that of the main/side chain (CB–CA) fragmentation. This result is consistent with the atomic

36

charge analysis data, and the change in the electron density distribution explains the differences in the atomic charges of the imidazole ring of the histidine. Therefore, main/side chain (CB–CG) fragmentation was confirmed to be the calculation method with the fewest splitting errors, even in electron density.

From the above results, the accuracy verification of the total energy, IFIE, atomic charge, and electron density compared with the calculation efficiency and main chain fragmentation in this molecular system indicates that main/side chain (CB–CG) fragmentation is an appropriate fragmentation method. However, metal-containing proteins have variations in the amino acid residues coordinated with the metal and in their sites (main chain and side chain). Therefore, the fragmentation of main/side chain (CB–CG) is not always a general-purpose method. In particular, the electronic state of a metal is essential in a reaction, so it is necessary to develop a fragmentation method that can adequately handle the electronic state. In the future, we will accumulate FMO calculation examples for metal-containing proteins and aim to prepare rules for appropriate fragmentation methods.

4. CONCLUSION

In this study, we reported a library with MOE for creating input files for FMO calculations performed using ABINIT-MP and analyzed the results. This approach makes it possible to complete the construction of molecular structures, generate input files for FMO calculations, and explore the results. As an example

https://doi.org/10.26434/chemrxiv-2024-2xk5n ORCID: https://orcid.org/0000-0002-0742-3896 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

of using FMOe, we performed IFIE/PIEDA interaction analysis by ligand fragmentation for the complex between the tripeptide inhibitor nirmatrelvir, with a Cys145 covalent bond, and SARS-CoV-2 M^{pro}. This example is challenging to handle using the current automatic fragmentation function of ABINIT-MP. Nevertheless, using FMOe, we performed FMO calculations relatively easily and identified the key amino acid residues of M^{pro} in nirmatrelvir binding and their interactions with the functional groups of the inhibitor. Next, the accuracy of the fragmentation method for Zn^{2+} ions was verified for a cyclic peptide coordinated with a Zn^{2+} ion. Here, we attempted the following fragmentations: the four histidines, His6, His18, D-His12, and D-His24, coordinated with the Zn^{2+} ion were divided into the main chain and the side chain with the BDA at the CB carbon atom and the BAA at the CA or CG carbon atom, and the side chains of histidines and the Zn^{2+} ion were merged into one fragment. Consequently, main/side chain (CB–CG) fragmentation was calculated efficiently and accurately.

FMOe simplifies complicated fragmentation, such as that of covalent inhibitors, and is also expected to efficiently handle medium-molecular-weight drug candidate compounds for diversifying drug discovery modalities (e.g., cyclic peptides). In addition, the examination of metal-containing proteins, where an amino acid residue is coordinated with a metal (e.g., an enzyme), was accelerated. In the future, it is expected that the accumulation of knowledge of the FMO calculation results of these complicated molecular systems will lead to the establishment of general-purpose fragmentation rules that enable more

38

convenient automatic fragmentation processing. Therefore, we plan to simplify the workflow and consequently accelerate the research. This library is available in the public domain at https://github.com/drugdesign/FMOE.

Data availability

The source code of FMOe is available from https://github.com/drugdesign/FMOE. The FMO calculation data is available from https://drugdesign.riken.jp/FMODB/.

Supporting information

The interaction energy analysis between nirmatrelvir (9I7) and each amino acid residue of SARS-CoV-2 M^{pro} using ligand fragmentation is displayed in Figures S1 and S2. Interaction energy analysis of Zn^{2+} ion fragment, including the four histidines for the cyclic peptide containing Zn^{2+} ion, is shown in Figures S3 and S4 based on the main chain and the main/side chain fragmentation schema, respectively. Atomic charges of His6, His18, D-His12 (DHI12), D-His24 (DHI24), and Zn^{2+} ion using Mulliken, natural population analysis, and Merz–Kollman approaches at the FMO-HF/6-31G and FMO-HF/6-31G* levels are listed in Tables S1 and S2, respectively. Mulliken, natural population analysis, and Merz–Kollman approaches for each fragmentation scheme at the HF/6-31G* level are displayed in Figures S5, S6, and S7, respectively. FMO-based electron density analysis of a Zn^{2+} ion and its coordinated histidines is demonstrated in Figure S8 (PDF).

39

Author information

Author Contributions

[#]HM and YK contributed equally to this work. Conceptualization: HM, YK, CW; Formal Analysis: CW; Software: HM, YK, CW, KK; Writing – original draft: CW, YO; Writing – review & editing: HM, YK, CW, KK, YO, KF, TH; Funding acquisition: CW, KF, TH; Supervision: KF, TH.

Notes

Acknowledgments

Abbreviations

BAA: bond attached atom; BDA: bond detached atom; CPF: check point file; CT: charge transfer; DI: dispersion interaction; ES: electrostatic; EX: exchange repulsion; FMO: fragment molecular orbital; FMODB: FMO database; GUI: graphical user interface; HF: Hartree-Fock; IFIE: Inter-fragment interaction energy; MO: molecular orbital; MOE: Molecular Operating Environment; MP2: Second-order Møller–Plesset perturbation theory; PDB: protein data bank; PIEDA: pair interaction energy decomposition analysis; SVL: Scientific Vector Language

References

(1) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules. *Chem. Phys. Lett.* **1999**, *313* (3–4), 701–706. DOI: 10.1016/S0009-2614(99)00874-X.

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

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

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

(5) Hitaoka, S.; Harada, M.; Yoshida, T.; Chuman, H. Correlation Analyses on Binding Affinity of Sialic Acid Analogues with Influenza Virus Neuraminidase-1 Using Ab Initio MO Calculations on Their Complex Structures. *J. Chem. Inf. Model.* **2010**, *50* (10), 1796–1805. DOI: 10.1021/ci100225b.

(6) Watanabe, C.; Fukuzawa, K.; Okiyama, Y.; Tsukamoto, T.; Kato, A.; Tanaka, S.; Mochizuki, Y.; Nakano, T. Three- and Four-Body Corrected Fragment Molecular Orbital Calculations with a Novel Subdividing Fragmentation Method Applicable to Structure-Based Drug Design. *J. Mol. Graph. Model.* **2013**, *41*, 31–42. DOI: 10.1016/j.jmgm.2013.01.006.

(7) Heifetz, A.; Trani, G.; Aldeghi, M.; MacKinnon, C. H.; McEwan, P. A.; Brookfield, F. A.; Chudyk, E. I.; Bodkin, M.; Pei, Z.; Burch, J. D.; Ortwine, D. F. Fragment Molecular Orbital Method Applied to Lead Optimization of Novel Interleukin-2 Inducible T-Cell Kinase (ITK) Inhibitors. *J. Med. Chem.* **2016**, *59* (9), 4352–4363. DOI: 10.1021/acs.jmedchem.6b00045.

(8) Watanabe, C.; Watanabe, H.; Fukuzawa, K.; Parker, L. J.; Okiyama, Y.; Yuki, H.; Yokoyama, S.; Nakano, H.; Tanaka, S.; Honma, T. 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. *J. Chem. Inf. Model.* **2017**, *57* (12), 2996–3010. DOI: 10.1021/acs.jcim.7b00110.

(9) 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: Enhanced Performance and Applicability*; Mochizuki, Y., Tanaka, S., Fukuzawa K., Eds.; Springer Singapore, **2021**; pp 93–125. DOI: 10.1007/978-981-15-9235-5_7.

(10) Barca, G. M. J.; Bertoni, C.; Carrington, L.; Datta, D.; De Silva, N.; Deustua, J. E.; Fedorov, D. G.; Gour, J. R.; Gunina, A. O.; Guidez, E.; Harville, T.; Irle, S.; Ivanic, J.; Kowalski, K.; Leang, S.

S.; Li, H.; Li, W.; Lutz, J. J.; Magoulas, I.; Mato, J.; Mironov, V.; Nakata, H.; Pham, B. Q.; Piecuch, P.; Poole, D.; Pruitt, S. R.; Rendell, A. P.; Roskop, L. B.; Ruedenberg, K.; Sattasathuchana, T.; Schmidt, M. W.; Shen, J.; Slipchenko, L.; Sosonkina, M.; Sundriyal, V.; Tiwari, A.; Galvez Vallejo, J. L.; Westheimer, B.; Włoch, M.; Xu, P.; Zahariev, F.; Gordon, M. S. Recent Developments in the General Atomic and Molecular Electronic Structure System. *J. Chem. Phys.* **2020**, *152* (15), 154102. DOI: 10.1063/5.0005188.

(11)   Nakano, T.; Mochizuki, Y.; Fukuzawa, K.; Amari, S.; Tanaka, S. CHAPTER 2 - Developments and Applications of ABINIT-MP Software Based on the Fragment Molecular Orbital Method. In *Modern Methods for Theoretical Physical Chemistry of Biopolymers*; Starikov, E. B., Lewis, J. P., Tanaka, S., Eds.; Elsevier Science: Amsterdam, **2006**; pp 39–52. DOI: 10.1016/B978-044452220-7/50066-6.

(12)   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. *Recent Advances of the Fragment Molecular Orbital Method: Enhanced Performance and Applicability*, **2021**; pp 53–67. DOI: 10.1007/978-981-15-9235-5_4.

(13)   *PAICS - parallelized ab initio calculation system based on FMO*. http://www.paics.net/index_e.html (accessed 2022-09-05).

(14)   Facio. https://zzzfelis.sakura.ne.jp/ (accessed 2024-01-30).

(15)   Kitaura, K. FU. https://cbi-society.org/home/documents/FU.html (accessed 2023-01-30).

(16)   BioStation Viewer. https://fmodd.jp/biostationviewer-dl/.

(17)   PAICS View. http://www.paics.net/paics_view_e.html (accessed 2024-01-30).

(18)   Watanabe, C.; Tanaka, S.; Okiyama, Y.; Yuki, H.; Ohyama, T.; Kamisaka, K.; Takaya, D.; Fukuzawa, K.; Honma, T. Quantum Chemical Interaction Analysis between SARS-CoV-2 Main Protease and Ensitrelvir Compared with Its Initial Screening Hit. *J. Phys. Chem. Lett.* **2023**, *14* (15), 3609–3620. DOI: 10.1021/acs.jpclett.2c03768.

(19)   Molecular Operating Environment (MOE), 2022.02 Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2022.

(20)   *FMOe*. https://github.com/drugdesign/FMOe.

(21)   Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; Dantonio, A.; Di, L.; Eng, H.; Ferre, R.; Gajiwala, K. S.; Gibson, S. A.; Greasley, S. E.; Hurst, B. L.; Kadar, E. P.; Kalgutkar, A. S.; Lee, J. C.; Lee, J.; Liu, W.; Mason, S. W.; Noell, S.; Novak, J. J.; Obach, R. S.; Ogilvie, K.; Patel, N. C.; Pettersson, M.; Rai, D. K.; Reese, M. R.; Sammons, M. F.; Sathish, J. G.; Singh, R. S. P.; Steppan, C. M.; Stewart, A. E.; Tuttle, J. B.; Updyke, L.; Verhoest, P. R.; Wei, L.; Yang, Q.; Zhu, Y. An Oral SARS-CoV-2 Mpro Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science* **2021**, *374* (6575), 1586–1593. DOI: 10.1126/science.abl4784.

(22)   Kneller, D. W.; Li, H.; Phillips, G.; Weiss, K. L.; Zhang, Q.; Arnould, M. A.; Jonsson, C. B.; Surendranathan, S.; Parvathareddy, J.; Blakeley, M. P.; Coates, L.; Louis, J. M.; Bonnesen, P. V.; Kovalevsky, A. Covalent Narlaprevir- and Boceprevir-Derived Hybrid Inhibitors of SARS-CoV-2 Main Protease. *Nat. Commun.* **2022**, *13* (1), 2268. DOI: 10.1038/s41467-022-29915-z.

(23)   Mochizuki, Y.; Koikegami, S.; Nakano, T.; Amari, S.; Kitaura, K. Large Scale MP2 Calculations with Fragment Molecular Orbital Scheme. *Chem. Phys. Lett.* **2004**, *396* (4–6), 473–479. DOI: 10.1016/j.cplett.2004.08.082.

(24)   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. **2020**. DOI: 10.26434/chemrxiv.13200176.v3.

(25)   Nakano, T.; Kaminuma, T.; Sato, T.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. Fragment Molecular Orbital Method: Application to Polypeptides. *Chemical Physics Letters* **2000**, *318* (6), 614–618. DOI: 10.1016/S0009-2614(00)00070-1.

(26)   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. *CBIJ* **2019**, *19*, 5–18. DOI: 10.1273/cbij.19.5.

(27)   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. *J. Chem. Inf. Model.* **2021**, *61* (2), 777–794. DOI: 10.1021/acs.jcim.0c01062.

(28)   Imai, K.; Takimoto, D.; Saito, R.; Watanabe, C.; Fukuzawa, K.; Noriyuki, K. FMO Calculations for Zinc Metalloprotease: Fragmentation of Amino-Acid Residues Coordinated to Zinc Ion. *CBIJ* **2022**, *22*, 21–25. DOI: 10.1273/cbij.22.21.

(29)   Kato, K.; Yamamoto, A.; Watanabe, C.; Fukuzawa, K. Application of Model Core Potentials to Zn- and Mg-Containing Metalloproteins in the Fragment Molecular Orbital Method. *CBIJ* **2023**, *23* (0), 14–25. DOI: 10.1273/cbij.23.14.

(30)   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 Calculations and Interaction Energy Analysis of SARS-CoV-2-Related Proteins. *J. Chem. Inf. Model.* **2021**, *61 (9)*, 4594–4612. DOI: 10.1021/acs.jcim.1c00694.

(31)   Fedorov, D. G.; Kitaura, K. CHAPTER 1 - Theoretical Development of the Fragment Molecular Orbital (FMO) Method. In Modern Methods for Theoretical Physical Chemistry of Biopolymers; Starikov, E. B., Lewis, J. P., Tanaka, S., Eds.; Elsevier Science: Amsterdam, **2006**; pp 3–38. DOI: 10.1016/B978-044452220-7/50065-4.

(32)   Kurisaki, I.; Fukuzawa, K.; Komeiji, Y.; Mochizuki, Y.; Nakano, T.; Imada, J.; Chmielewski, A.; Rothstein, S. M.; Watanabe, H.; Tanaka, S. Visualization Analysis of Inter-Fragment Interaction Energies of CRP–cAMP–DNA Complex Based on the Fragment Molecular Orbital Method. *Biophysical Chemistry* **2007**, *130* (1–2), 1–9. DOI: 10.1016/j.bpc.2007.06.011.