Statistical interaction analyses between SARS-CoV-2 main protease and inhibitor N3 by combining molecular dynamics simulation and fragment molecular orbital calculation

Ryo Hatada1, Koji Okuwaki1, Kazuki Akisawa1, Yuji Mochizuki1,2, Yuma Handa3, Kaori Fukuzawa2,3,4, Yuto Komeiji5, Yosio Okiyama6, and Shigenori Tanaka7

Received December 6, 2020; revised December 28, 2020; accepted January 12, 2021; published online January 21, 2021

© 2021 The Author(s). Published on behalf of The Japan Society of Applied Physics by IOP Publishing Ltd

A combination of classical molecular dynamics (MD) simulation and ab initio fragment molecular orbital (FMO) calculation was applied to a complex formed between the main protease of the new coronavirus and the inhibitor N3 to calculate interactions within the complex while incorporating structural fluctuations mimicking physiological conditions. Namely, a statistical evaluation of interaction energies between N3 and amino acid residues was performed by processing a thousand of structure samples. It was found that relative importance of each residue is altered by the structural fluctuation. The MD-FMO combination should be promising to simulate protein related systems in a more realistic way.

Supplementary material for this article is available online

The COVID-19 disease caused by the new severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has still been the worldwide critical issue, as reported by the World Health Organization (WHO). A great amount of works in various fields have been conducted against this difficult situation. In the pharmaceutical industry, several proteins such as the main protease (Mpro) of SARS-CoV-2 have attracted considerable attention because they are potential targets of antiviral drugs, and hence a number of computational studies have been performed to find effective inhibitors against them. For example, Xu et al. showed that nelfinavir may be a promising inhibitor to Mpro, based on a homology modeling and docking simulations of a number of trial ligands. Jin et al. developed a new peptide-like inhibitor named N3 with virtual screening techniques and determined the crystal structure of the Mpro–N3 complex [published as 6LU7 in Protein Data Bank (PDB)]. By this availability of 6LU7 structure, several related studies such as molecular dynamics (MD) simulation were reported. Among them, we performed ab initio fragment molecular orbital (FMO) based analyses of interactions between the N3 inhibitor and amino acid residues in the pharmacophore, while using pair interaction energy (PIE) or interfragment interaction energy (IFIE) to identify important residues having site-specific interactions, e.g. hydrogen bond, with N3. We have successfully applied FMO to other SARS-CoV-2 protein systems including the spike protein and RNA polymerase—remdesivir complex, as in the cases of precedent studies for viral proteins.

Although we have performed even detailed PIE decomposition analyses (PIEDA) including an implicit hydration effect through the Poisson–Boltzmann (PB) model in Ref. 13, the structural fluctuation effect (which should be essential in physiological situation at room temperature) was not incorporated because we performed only a single structure calculation for the 6LU7 crystal data. Ishikawa et al. pioneered a statistical scheme in which classical MD simulations are first performed to generate structure samples and these series of structures are subjected to FMO calculations so that the structural fluctuation effects on IFIEs can be taken into account. We applied this MD–MO combinative scheme to DNA–uranil complex as well as the designed peptide–inorganic surface systems. Here, we took the same approach for the Mpro–N3 complex, while using a thousand of structure samples, ten times more than those in Refs. 30 and 31, by using a massively parallel computing resource of Fugaku supercomputer. The computational protocol used are summarized as follows.

We started with the crystal structure of the Mpro–N3 complex (PDB ID: 6LU7), from which we extracted a monomer block, and processed it in a standard manner with the MOE software; namely, we added missing hydrogens and optimized them with the AMBER10:EHT force field (FF) (refer also to Ref. 13). This structure is referred to as “static” hereafter. A thousand of “dynamic” structures were prepared by the MD simulation method with the AMBER16 program (with the AMBER10:EHT FF). The MD run durations were 50 ps for the thermal elevation from 0 to 310 K (NVT ensemble), 50 ps for the density relaxation (NPT), 1 ns for the equilibration (NPT), and 100 ns for the production process (310 K, 1 atm, NPT). From trajectories of
the production run, a thousand snapshots were sampled with an interval of 100 ps. This long-time and rich sampling should cover larger configurational spaces and thus provide more reliable statistical evaluations of interaction energies, in comparison with our previous studies.30–34 All the MD simulations were performed on the TSUBAME3.0 computer. Each sample structure was shaped to a droplet form with water molecules within 4 Å of Mpro; this criterion of water layer thickness was determined according to Refs. 39–41. The system ions (Na+) were retained to electrically neutralize the system.

The electronic energy of two-body FMO expansion14–17 can be expressed as

$$E_{FMO} = \sum_i E_i' + \sum_{i<j} \Delta E_{ij},$$

where $i$ and $j$ are fragment indices. The first term of Eq. (1) is the fictitiously isolated monomer energy, and the second term corresponds to the fragment interaction energy termed as PIE16 or IFIE17,18 In PIEDA,25,26 each interaction energy is further decomposed as

$$\Delta E_{ij} = E(ES) + E(EX) + E(CT) + E(DI),$$

where ES, EX, CT and DI mean “ElectroStatic”, “EXchange repulsion”, “Charge-Transfer and mix terms”, and “DIspersion”, respectively. The EX, EX, and CT terms are evaluated at the Hartree–Fock level,42 whereas the DI term is usually obtained at the second-order Möller–Plesset perturbation (MP2) level including the electron correlation effect.42–45 We adopted the partial renormalization,46 to reduce the tendency of overestimation in the MP2 based DI energy.19 As shown in Fig. 1, the N3 ligand has a covalent bond to Cys145, and N3 was segmented to 5 fragments.15 The numbers of fragments were 311 for the Mpro—N3 complex and about 1400 for the water layer, respectively. The basis set used was 6-31 G*,47 a standard choice in recent FMO studies. A series of FMO-MP2/6-31 G* calculations with the PIEDA option were performed with the ABINIT-MP program17 on the supercomputer Fugaku. Typical job timing for a single structure sample was 0.6 h with 192 nodes (or a half rack) of Fugaku, and the concurrent job processing led to only 6 h for the completion of all the structure samples.

In interpretation of IFIE and PIEDA, care is necessary for the fact that the fragmentation of usual FMO is made at the single bond between the carbonyl carbon and the alpha carbon (with side chain)48 and thus that a misleading assignment shift of interacting fragment can occur when the carbonyl oxygen forms a hydrogen bond. This matter was carefully handled as was done in Ref. 13. We obtained the averages and standard deviations (or 1σ) of IFIE and PIEDA data calculated for the “dynamic” structures.

Below, we present the results and discussion. Figure 2 illustrates a superimposed view of 10 ns interval sampled structures of the N3 ligand and crucial residues. Because of the covalent bond to Cys145 of Mpro, fluctuations of Fragments 3 and 4 are small, whereas those of the terminal Fragments 1 and 5 are larger. Refer also to Fig. S1 (available online at stacks.iop.org/APEX/14/027003/mmedia) in Supporting Information for the root mean square deviation of trajectories.

The total IFIE between Mpro and N3 is $-199.9$ kcal mol$^{-1}$ for the “static” structure, and the corresponding statistical value of the “dynamic” structures (of a thousand samples) is $-145.9 \pm 10.8$ kcal mol$^{-1}$, as found in Table S1; see a comparison with another “dynamic” results obtained by the first hundred samples if necessary. As discussed later, the overall decrease of stabilization might be attributed to the increased average distances between the N3 fragments and the interacting residues relative to the “static” or crystal structure. Crystal structures may be generally denser than fluctuated structures in physiological conditions, leading to the closer and stronger interactions with ligands in the pharmacophore. Namely, there are potential overestimations in IFIEs in the present system of Mpro—N3 for the X-ray crystal structure. The IFIE value of each Fragment of N3 is plotted in Fig. 3 (refer to Table S1 again). The stabilization of Fragment 4, which was found to be the most important moiety in Ref. 13, is sizably decreased relative to the “static” value, or the overestimation by the “static” structure is confirmed. As a result, the importance of Fragment 3 becomes comparable to that of Fragment 4 in a statistical view with “dynamic” structures. These two Fragments of N3 are close to a covalent bond to Cys145, and their mobile abilities are rather restricted relative to other Fragments.

---

**Fig. 1.** (Color online) Fragmentation of N3 ligand. The fragmentation point is identified by orange triangle indicating the direction of bond detachment atom (BDA) to bond attachment atom (BAA).17 There is a covalent bond to Cys145.13 Refer to related descriptions in Ref. 13 when needed.
being consistent with the situation observed in Fig. 2. Interestingly, the relative importance of Fragment 5 is enhanced in the “dynamic” structures. A plausible reason could be considered as follows. Fragment 5 is almost invisible on the electron density map of the crystal (see Fig. S2), so the atomic coordinates of this moiety should be relatively unreliable in the crystal structure. Therefore, the stabilization of interaction energy suggests that Fragment 5 is more precisely positioned by considering “dynamic” structures through MD.

Next, we discuss residue specific discussion. Figure 4 shows IFIE plots of leading residues whose stabilization energies are larger than $-10$ kcal mol$^{-1}$ (see Table S2 compiling the respective IFIE values and contact distances). As can be seen, His163 is the outstanding residue with the largest stabilization in both “static” and “dynamic” evaluations. Other notable residues in stabilizations are Met165, Leu167, and Gln189. The contribution from these four residues were consistently pointed out in Refs. 3, 6–8, and 13 (ours). A considerable amount of decrease in stabilization (ca. 15 kcal mol$^{-1}$) is found for His163 (associated with Fragment 4), while those of Met165, Leu167, and Gln189 are lesser. These relations are in accord with the differences in the contact distances listed in Table S2; about 0.6 Å is found for His163. Keller et al. investigated the structural changes of 6LU7 by MD simulation and observed that a loop consisting of Glu166—Gly170 in Mpro moves away from the N3 ligand. The present results are consistent with their observation. More detailed Fragment-wise results of IFIE, are provided in Figures S3–S7 as well as Tables S3–S7; refer to these numerical data when necessary.
Figure 5 plots the averaged PIEDA values (compiled in Table S8) of the leading residues in the interactions between Mpro and N3. The ES term is the main source of stabilization of hydrogen bonds, and its amount of His163 is the largest one. As can be seen in Fig. 2, the N–H part of the ring in His163 is directed toward the carbonyl carbon of Fragment 4, and thus the ES contribution may be enhanced. In comparison with the “static” PIEDA results in Ref. 13, contributions from the CT term are decreased. On the other hand, the DI term plays a supplementary role in stabilization of the “dynamic” situation, especially for Met165 and Gln189. His41 is interacting with Fragment 5 primally by hydrogen bond associated with its N–H part in the ring. This situation is contrasted to another case in which His41 rather has ππ stacking stabilizations with aromatic rings of different (not N3) ligands. 10)

As we have seen so far, there are several differences between the “static” and “dynamic” structures in interactions between Mpro and N3. We analyzed the correlation between “dynamic” IFIE values and interacting distances, and the results are given in Fig. S8 and Table S9. As expected, the values of correlation coefficient have sizable values of 0.6–0.7, except for some cases. This fact implies that thermal fluctuations in a realistic condition certainly affect the interactions among the N3 ligand and surrounding residues.

In the present letter, we have reported a combinative approach of classical MD and ab initio FMO for the Mpro—N3 complex, which indicated the importance of statistical evaluation of interactions by employing a thousand of fluctuating structures. The degree of reality to simulate physiological condition was thus enhanced in interaction analyses. Recent publications 6,9–12,49–51) related with SARS-CoV-2 proteins have shed light on the “dynamic” aspects, and our work has been just along this line. The developments of biomaterials 52–58) have attracted attention and interest, and theoretical analyses and simulations could contribute to efficient designs and optimizations (e.g. Refs. 32–34). As demonstrated here, the combination of MD and FMO with an amount of sampled structures has now been made practical by using massive computing resources such as the Fugaku supercomputer. We would hope that this approach will be applied to various systems in the nano-biotechnology field of applied physics. 59)
Acknowledgments For FMO calculations with ABINIT-MP, computational resources of the supercomputer Fugaku (the evaluation environment in the trial phase) were provided by the RIKEN Center for Computational Science (R-CCS) under a special project against COVID-19. Technical supports on Fugaku were provided by Drs. Kota Sakakura (FOCUS) and Hirokazu Watanabe (BPC Systems Inc.). All MD simulations were performed on TSUBAME3.0 at Tokyo Institute of Technology under a funding by AMED/HINDS (JP20am010113). The present work was supported also by Rikkyo SFR as well as JSPS Kakenhi (JP17H06353, JP18K03825, and JP19K12100).

ORCID iDs Koji Okuizaki https://orcid.org/0000-0002-4510-5177 Yugi Mochizuki https://orcid.org/0000-0002-7510-9887 Kaori Fukuzawa https://orcid.org/0000-0001-5357-8250 Yuto Kornei https://orcid.org/0000-0002-9708-6759 Yoshio Okiyama https://orcid.org/0000-0002-6679-2788 Shigenori Tanaka

1) [https://who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports] (Accessed: 24/11/2020).
2) Z. Xu, C. Peng, Y. Shi, Z. Zhu, K. Mu, X. Wang, and W. Zhu, (2020), bioRxiv:2020.01.27.921627.
3) Z. Jin et al., Nature 582, 289 (2020).
4) Z. Li et al., Proc. Natl Acad. Sci. 117, 27381 (2020).
5) S. K. Emnomhi, K. Raja, I. Sebastian, and J. Joseph, J. Biomol. Struct. Dyn. (2020).
6) L. Mittal, A. Kumari, M. Srivastava, M. Singh, and S. Ashiana, J. Biomol. Struct. Dyn. (2020).
7) R. Yu, L. Chen, R. Lan, R. Shen, and P. Li, Int. J. Antimicrob. Agents 56, 106032 (2020).
8) A. I. Owis, M. S. El-Hawary, D. Amir, O. M. Aly, U. R. Abdelmohsen, and M. S. Kamel, RSC Adv. 10, 19570 (2020).
9) B. Nutlo, P. Malapalibut, K. Hengphasatporn, N. C. Pattaranggoon, N. Simanom, Y. Shigeta, S. Hammadobu, and T. Rangtromgpton, Biochemistry 59, 1769 (2020).
10) R. Yoshino, N. Yasuo, and M. Sekijima, Sci Rep. 10, 12493 (2020).
11) A. Gahlawat, N. Kumar, R. Kumar, H. Sandhu, I. P. Singh, S. Singh, A. Sjostedt, and P. Garg, J. Chem. Inf. Model. 60, 5781 (2020).
12) D. Suarez and N. Diaz, J. Chem. Inf. Model. 60, 5815 (2020).
13) R. Hatada, K. Okuizaki, Y. Mochizuki, Y. Handa, K. Fukuzawa, Y. Komeiji, Y. Okiyama, Y. Mochizuki, and S. Tanaka, J. Chem. Inf. Model. 60, 3593 (2020).
14) D. G. Fedorov and K. Kitaara (ed.) The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems (CRC Press, Boca Raton, FL, 2009).
15) K. Kitaara, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. 333, 701 (1999).
16) D. G. Fedorov, T. Nagata, and K. Kitaara, Phys. Chem. Chem. Phys. 14, 7562 (2012).
17) S. Tanaka, Y. Mochizuki, Y. Komeiji, Y. Okiyama, and K. Fukuzawa, Phys. Chem. Chem. Phys. 16, 10310 (2014).
18) S. Amari, M. Aizawa, J. Zhang, K. Fukuzawa, Y. Mochizuki, Y. Iwasawa, K. Nakata, H. Chuman, and T. Nakano, J. Chem. Inf. Model. 46, 221 (2006).
19) K. Akiwata, Y. Okiyama, K. Fukuzawa, Y. Mochizuki, and S. Tanaka, J. Chem. Inf. Model. 46, 1327 (2021).
20) K. Kato, T. Honma, and K. Fukuzawa, J. Mol. Graph. Model. 100, 107695 (2020).
21) T. Iwata, K. Fukuzawa, K. Nakajima, S. Aida-Hyugui, Y. Mochizuki, H. Watanabe, and S. Tanaka, Comput. Biol. Chem. 32, 198 (2008).
22) K. Takematsu, K. Fukuzawa, K. Omagari, S. Nakajima, K. Nakajama, Y. Mochizuki, T. Nakano, H. Watanabe, and S. Tanaka, J. Phys. Chem. B 113, 4991 (2009).
23) A. Yoshioka, K. Fukuzawa, Y. Mochizuki, K. Yamashita, T. Nakano, Y. Okiyama, E. Nobusawa, K. Nakajima, and S. Tanaka, J. Mol. Graph. Model. 30, 110 (2011).