



Collective residue interactions in trimer complexes of SARS-CoV-2 spike proteins analyzed by fragment molecular orbital method

Okuwaki, Koji ; Akisawa, Kazuki ; Hatada, Ryo ; Mochizuki, Yuji ; Fukuzawa, Kaori ; Komeiji, Yuto ; Tanaka, Shigenori

(Citation)

Applied Physics Express, 15(1):017001

(Issue Date)

2022-01

(Resource Type)

journal article

(Version)

Version of Record

(Rights)

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

Content from this work may be used under the terms of the Creative Commons Attribution 4.0 license. Any further distribution of this work must maintain attribution to the...

(URL)

<https://hdl.handle.net/20.500.14094/90008894>



LETTER • OPEN ACCESS

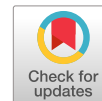
Collective residue interactions in trimer complexes of SARS-CoV-2 spike proteins analyzed by fragment molecular orbital method

To cite this article: Koji Okuwaki *et al* 2022 *Appl. Phys. Express* **15** 017001

View the [article online](#) for updates and enhancements.

You may also like

- [Emerging nanotechnology role in the development of innovative solutions against COVID-19 pandemic](#)
Zeeshan Ahmad Bhutta, Ayesha Kanwal, Moazam Ali et al.
- [Inactivation of SARS-CoV-2 in clinical exhaled breath condensate samples for metabolomic analysis](#)
Shuang Hu, Mitchell M McCartney, Juan Arredondo et al.
- [Binding of SARS-CoV-2/SARS-CoV spike protein with human ACE2 receptor](#)
Rajendra P Koirala, Bidhya Thapa, Shyam P Khanal et al.



Collective residue interactions in trimer complexes of SARS-CoV-2 spike proteins analyzed by fragment molecular orbital method

Koji Okuwaki¹ , Kazuki Akisawa¹, Ryo Hatada¹, Yuji Mochizuki^{1,2} , Kaori Fukuzawa^{2,3,4} , Yuto Komeiji⁵ , and Shigenori Tanaka^{6*}

¹Department of Chemistry and Research Center for Smart Molecules, Faculty of Science, Rikkyo University, 3-34-1 Nishi-ikebukuro, Toshima-ku, Tokyo 171-8501, Japan

²Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

³School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan

⁴Department of Biomolecular Engineering, Graduate School of Engineering, Tohoku University, 6-6-11 Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan

⁵Health and Medical Research Institute, AIST, Tsukuba Central 6, Tsukuba, Ibaraki 305-8566, Japan

⁶Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan

*E-mail: tanaka2@kobe-u.ac.jp

Received November 13, 2021; revised December 5, 2021; accepted December 14, 2021; published online December 24, 2021

In large biomolecular systems such as protein complexes, there are huge numbers of combinations of inter-residue interactions whose comprehensive analyses are often beyond the intuitive processing by researchers. Here we propose a computational method to allow for a systematic analysis of these interactions based on the fragment molecular orbital calculations, in which the inter-fragment interaction energies are comprehensively processed by the singular value decomposition. For a trimer complex of SARS-CoV-2 spike protein, three-body interactions among residues belonging to three chains are analyzed to elicit a small number of essential interaction modes or networks crucial for the structural stability of the complex. © 2021 The Author(s). Published on behalf of The Japan Society of Applied Physics by IOP Publishing Ltd

Supplementary material for this article is available [online](#)

Recently, the advance in cryogenic electron microscope (cryo-EM) technology^{1,2)} has significantly spurred the structure determinations of huge biomolecular systems such as membrane proteins and protein-DNA (or RNA) complexes, which were highly intractable previously. This trend necessarily demands, in the context of computational structural biology, the analyses of complicated interactions between many residues or nucleotides involved in the molecular complexes. On the other hand, the progress in ab initio calculation techniques for biomolecular systems as seen in the fragment molecular orbital (FMO) method^{3–6)} has enabled the accurate interaction analysis on large biomolecular complexes. These examples include the complexes associated with influenza⁷⁾ and measles⁸⁾ hemagglutinin proteins, and SARS-CoV-2 spike proteins,^{9–14)} whose multimer structures contain more than thousands of residues. More specifically, in a recent study on SARS-CoV-2 spike protein composed of a trimer of chains A–C (three monomers) with more than a thousand residues each, the electron-correlated FMO calculations at the perturbative Møller–Plesset second-order (MP2) to fourth-order (MP4) levels were carried out¹²⁾ on the Fugaku supercomputer with the world's highest performance currently.¹⁵⁾ It is striking that, since FMO-MP3 calculations¹⁶⁾ on these systems can be performed only by about twice computational cost compared with FMO-MP2 on the Fugaku supercomputer, we can obtain the MP2.5 level evaluations¹⁷⁾ for the inter-fragment interaction energies (IFIEs)^{5,6)} in terms of a scaled MP3 approach, which are comparable in accuracy to very accurate CCSD(T) level values. These accurate results on the IFIEs have then provided useful information about the molecular recognition in biomolecular complexes.

Here, our problem consists of the fact that there are too many inter-residue interactions to be analyzed. For example,

if we consider the inter-residue interactions between two protein monomers with a thousand (10^3) residues each, we should take account of $(10^3)^2 = 10^6$ interaction pairs. Furthermore, if we simultaneously consider the inter-residue interactions in a trimer with a thousand residues each, the number of combinations of three-body interactions amounts to $(10^3)^3 = 10^9$, whose comprehensive analysis gives a big challenge in data science. It is noted in this context that we cannot confine our consideration to the short-range interactions between nearby residues due to the involvement of long-range electrostatic interactions. The purpose of the present study is thus to propose a novel prescription to analyze the FMO-based inter-residue interactions (or IFIEs) in huge trimer or higher complexes, in which the information of many combinations of three-body interactions is efficiently compressed to extract the essential interactions crucial for molecular recognition and structural stability of complexes. Employing the trimers of SARS-CoV-2 spike protein as a prototypical system, we perform the FMO2-MP2.5/6-31G* calculations¹²⁾ to obtain the IFIEs between all the amino-acid residues contained in the trimer, where “FMO2” means the two-body expansion approximation^{5,6)} conventionally adopted in the FMO method. We then use a data compression technique based on the singular value decomposition (SVD)^{18–21)} in which the important residues in each monomer (chain) are selected at the preliminary level of dimer interaction analysis and employed as bases to find the crucial interaction networks at the trimer level. While higher-accuracy analyses based on higher-cost basis set and FMO3 or higher expansion method^{5,6)} would be optionally feasible, the primary aim of this work is to provide a basic computational framework for coping with a huge number of IFIE combinations. With the aid of the proposed methodology, we will elucidate the differences in trimer interactions between



the closed and open structures of SARS-CoV-2 spike proteins, which would be useful for understanding the infection mechanism of COVID-19 viruses.

In the following, let us consider three-body interaction energies E_{ijk} among three fragments, i , j , and k . In this study we specifically consider three amino-acid residues labeled by i , j , k , that belong to each monomer, A, B, C, respectively, composing a trimer complex protein. For instance, each monomer contains more than 10^3 residues in the case of the spike protein of SARS-CoV-2,²²⁾ thus providing a challenge to the analysis of three-body interaction energies E_{ijk} .

The primary contribution to E_{ijk} comes from the combination of two-body interactions, which can be represented by the IFIEs E_{ij} in the FMO calculations. The details to obtain the IFIEs for the spike protein of SARS-CoV-2 are illustrated below. Here, using the IFIEs obtained by the present FMO2 calculations, the three-body interaction energies are simply expressed by geometric mean,

$$E_{ijk} = (E_{ij}E_{jk}E_{ki})^{1/3}, \quad (1)$$

while the direct three-body correction energies may be included on demand via the costly FMO3 calculations.⁵⁾

Since the number of elements for E_{ijk} exceeds 10^9 , we need significant data compression to perform a practical interaction energy analysis. Considering also that the two-body interactions, E_{ij} , E_{jk} , and E_{ki} , have more than 10^6 elements each, we first attempt the CP (CANDECOMP/PARAFAC) decomposition or the SVD^{18–21)} for them as

$$E_{ij} = \sum_{\mu} \lambda_{\mu} u_{\mu i} v_{\mu j}, \quad (2)$$

where λ_{μ} is the (decremental) eigenvalue for mode μ , and $u_{\mu i}$ and $v_{\mu j}$ refer to the associated eigenvectors describing the contributions from each fragment (amino-acid residue in this case) of two monomers (chains) so that $U = (u_{\mu i})$ and $V = (v_{\mu j})$ construct unitary matrices. Performing similar procedures for E_{jk} and E_{ki} as well, we only keep higher-value contributions (say, top 10) of the eigenvalues. Retaining only the highly contributing residues (say, top 10) to each eigenvector, we can thus select about 10^2 important residues for each monomer. In this way, the remaining number of the elements for E_{ijk} becomes about $(10^2)^3 = 10^6$.

Furthermore, for the three-body tensor interactions obtained above, we carry out the CP decomposition as²¹⁾

$$E_{ijk} = \sum_{\mu} \Lambda_{\mu} u_{\mu i} v_{\mu j} w_{\mu k}, \quad (3)$$

where Λ_{μ} is the (decremental) eigenvalue for mode μ , and $u_{\mu i}$, $v_{\mu j}$ and $w_{\mu k}$ refer to the associated eigenvectors to form unitary matrices, $U = (u_{\mu i})$, $V = (v_{\mu j})$ and $W = (w_{\mu k})$. The crucial residues for the three-body interactions are thus extracted as important contributors to the higher-eigenvalue modes in a systematic way. We consider the top 20 modes in the present work. The SVD (CP decomposition) was performed with TensorLy library (<http://tensorly.org/stable/index.html>).²³⁾

Both closed and open structures of the SARS-CoV-2 spike protein for FMO calculations were prepared from cryo-EM structures,^{24,25)} where the corresponding PDB IDs were

6VXX and 6VYB, respectively. Note that the open structure is assumed to be responsible for the infection into human cells via the binding of receptor binding domain (RBD)²⁶⁾ to angiotensin-converting enzyme 2 (ACE2). Because the data on both structures lacked several amino-acid residues due to their relatively low resolutions, homology modeling was conducted to reproduce the missing parts with the MOE program.²⁷⁾ The total number of residues was 3363 for the trimer. The positions of generated hydrogen atoms were optimized also with MOE. For these processed 6VXX (closed) and 6VYB (open) structures, a molecular dynamics (MD)-based relaxation was performed with the AMBER18 program.²⁸⁾ These processed structures of 6VXX and 6VYB are illustrated in Fig. 1, where chains A, B, and C are colored red, green, and blue, respectively, and the dark-colored parts correspond to RBD (Thr333-Pro527 of each chain). All three RBDs are facing inward in the closed structure, whereas the RBD of chain-B is facing outward in the open structure. RBDs are crucial in binding with ACE2 located at the host cell surface. These two protein models were then subjected to a series of FMO calculations. Further details on the protein structure preparations are described in the supplementary material (SM) available online at stacks.iop.org/APEX/15/017001/mmedia.

In the present study, the second and third-order Møller-Plesset (MP2 and MP3) perturbation calculations with 6-31G* basis set were performed for the spike proteins¹²⁾ with an integral-direct parallelism in ABINIT-MP program^{5,6)} in the framework of two-body expansion FMO2 method. The incremental correlation energies of MP3¹⁶⁾ were halved and added to the MP2 energies, thus obtaining the MP2.5 energies.¹⁷⁾ Details of the FMO calculations on the supercomputer Fugaku at the RIKEN Center for Computer Science (R-CCS) are provided in SM including the conditions of job execution.

Then, we have carried out the three-body (trimer) SVD analysis on the basis of the FMO-IFIE values for closed and open structures of SARS-CoV-2 spike proteins. First, for each pair combination of the three chains of A, B and C, we have performed the dimer SVD based on Eq. (2) in order to select the important residues for protein-protein interactions. We retained the top-10 eigenmodes with higher eigenvalues, and selected the top-10 residues with higher amplitudes in each eigenvector, whose rationale may be found below. Removing the duplication of residues, we have thus chosen 63, 56 and 59 individual residues for chains A, B and C, respectively, in the case of the closed form, and 69, 60 and 63 individual residues for chains A, B and C, respectively, in the case of the open form. Next, the trimer SVD was carried out in terms of Eq. (3), where only the key residues selected above were employed to construct the basis space.

The distributions of the eigenvalues in the trimer SVD are shown for the highest 20 values in Fig. S1 of SM for both the cases of closed and open forms. As seen in this figure, the eigenvalues of SVD become very small at the 20th mode, and we have confirmed that approximately two-thirds of accumulated eigenvalue sums of the 20 modes are contributed by the highest 10 modes. Then, selecting the top-3 eigenvectors with higher eigenvalues, we show the top-10 residue components with higher amplitudes in each eigenvector in Tables S1–S6 of SM. It is noted here that the top-10 and top-20

residues with higher amplitudes give approximately two-thirds and 80% contributions, respectively, for the first three modes in the case of closed structure. Figures 2 and 3 then depict the key residues in each chain for the top-3 eigenvectors visually on the three-dimensional chain structures. In order to see the locational relation among the key residues belonging to three chains, we also show the three-dimensional map for the whole complex structures in Fig. S2 of SM.

Let us first focus on the second and third eigenmodes for the closed structure. For the second eigenmode with the eigenvalue of 607, the contributions of Glu1031 and Arg1039 are dominant in all three chains of A, B and C. For the third eigenmode with the eigenvalue of 565, the contributions of Arg1091 and Asp1118 are dominant in all

three chains of A, B and C. These facts thus imply that these key residues belonging to the central helix, β -hairpin and β -sheet domains (see Table S7 in SM) are essential for keeping the stability of trimer structure primarily due to the salt bridges among charged residues. Similarly, we observe the dominance of the same residues in the cases of the first and second eigenmodes for the open structure as well.

As for the first eigenmode with the eigenvalue of 782 for the closed structure, we find as the important residues Arg983, Asp745, Asp737 and Arg44 in the chain A, Arg765, Arg457, Lys458, Asp737 and Lys417 in the chain B, and Arg319 and Asp614 in the chain C. In contrast to the second and third eigenmodes addressed above, this eigenmode illustrates the collective residue interactions over widely spread domains of the three chains including the

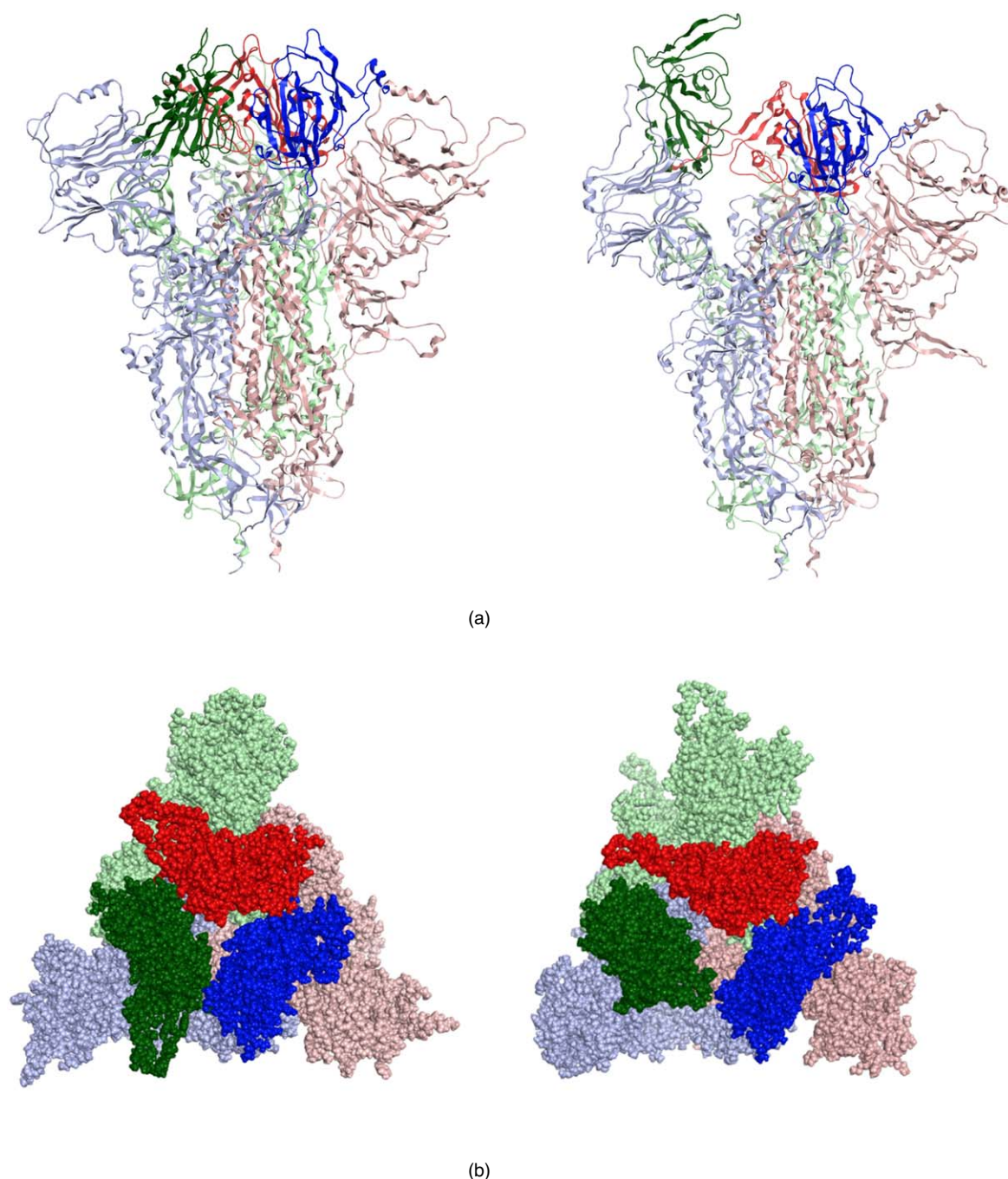


Fig. 1. (Color online) Closed (Left, PDB entry: 6VXX) and open (Right, PDB entry: 6VYB) structures of SARS-CoV-2 spike protein trimers. Red, green, and blue colors indicate chains A, B, and C, respectively. (a) Side view. (b) Top view. The dark-colored parts refer to the receptor binding domain (RBD).

© 2021 The Author(s). Published on behalf of

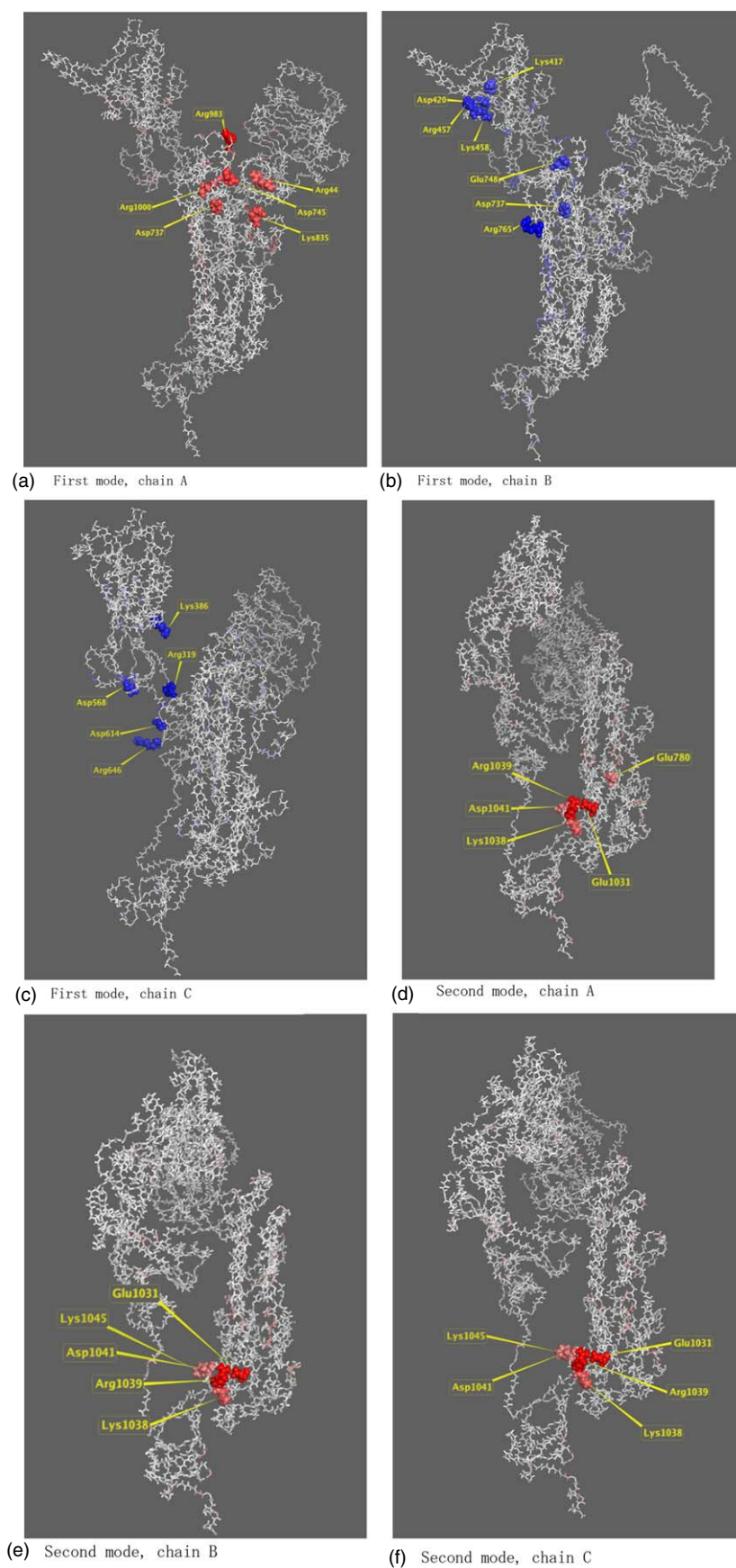
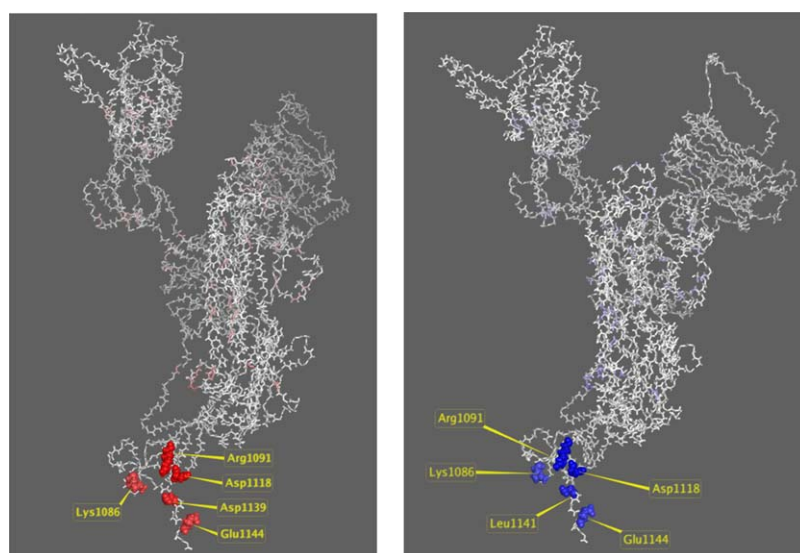


Fig. 2. Key residues with higher amplitudes in each eigenvector of SVD analysis for the closed structure of SARS-CoV-2 spike protein trimer. Red and blue colors refer to the positive and negative values of vector components, respectively, whose amplitudes are indicated by their deepness. Results for (a)–(c) first, (d)–(f) second, and (g)–(i) third eigenmodes with higher eigenvalues are depicted for the chains A–C.



(g) Third mode, chain A

(h) Third mode, chain B



(i) Third mode, chain C

(Continued.)

RBD, S1^D domain, N-terminal domain (NTD), central β -strand, downward helix and heptad repeat region (see Table S7 in SM), reflecting the significance of long-range electrostatic interactions. It is thus supposed that this interaction mode represents the association between the motion of RBD and the stability of the whole complex, which should be essential for the viral infection through the binding of RBD to ACE2 receptor. In particular, Asp614 in chain C is known²⁹⁾ to play a key role, whose mutation into Gly614 (D614G) significantly enhances the infection efficiency through the structural changes of spike protein. In addition, the mutations of Lys417 belonging to RBD are considered to be crucial^{13,30)} in recent variants of concern.^{31,32)} Then, this interaction mode corresponds to the third eigenmode for the open structure, where Lys786 and Lys854 in the chain A, Arg765 and Arg1019 in chain B, and Asp614 and Arg646 in the chain C show dominant interactions. It is remarkable that the interactions associated with Arg44, Arg319, Lys417, Arg457, Lys458, Asp737 and Asp745 are substantially reduced in this case through the transformation from the closed to open structures.¹²⁾ Overall, such a finding of

essential interaction networks would be helpful for the elucidation of those allosteric effects induced by the occurrence of amino-acid mutations.

By using the decomposition technique proposed in the present work, one can easily identify a small number of important interaction networks in biomolecular complexes (e.g. trimers) from a huge number of combinations of inter-residue interactions which may amount to the order of billions if naively treated. In the present FMO-based study we have applied the proposed prescription to the trimer complexes of SARS-CoV-2 spike protein, and succeeded in the efficient data compression concerning the extraction of inter-residue interaction network crucial for the structural stability and dynamics associated with the viral transmission into host cells. It is then remarked in the framework of FMO methodology that the intended applications incorporating the pair interaction energy decomposition analysis (PIEDA)³³⁾ and FMO3 expansion⁵⁾ are fairly straightforward. For example, the similar SVD analysis focusing on the electron-correlated, dispersion interaction part would be interesting in order to discuss the significance of relatively short-range

© 2021 The Author(s). Published on behalf of

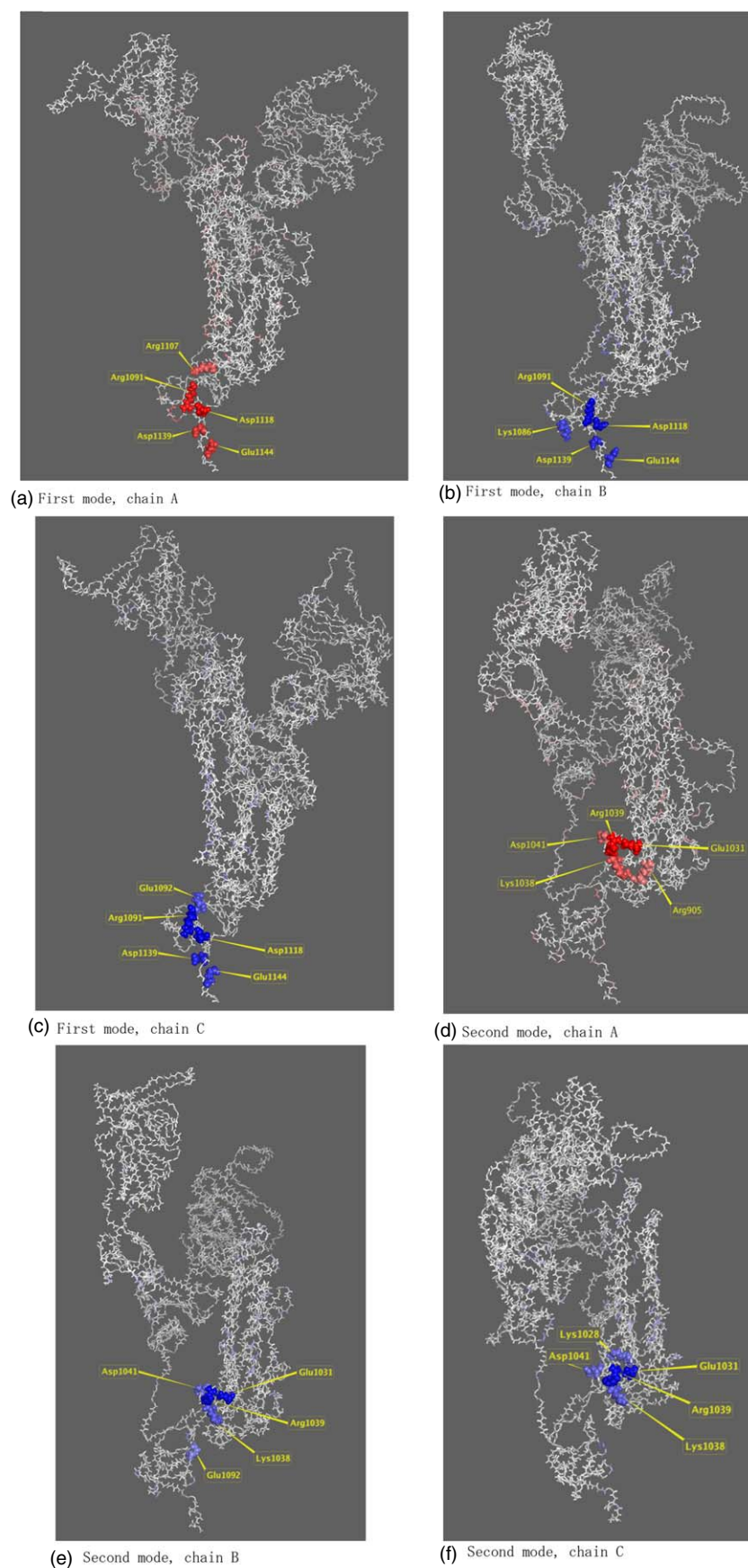
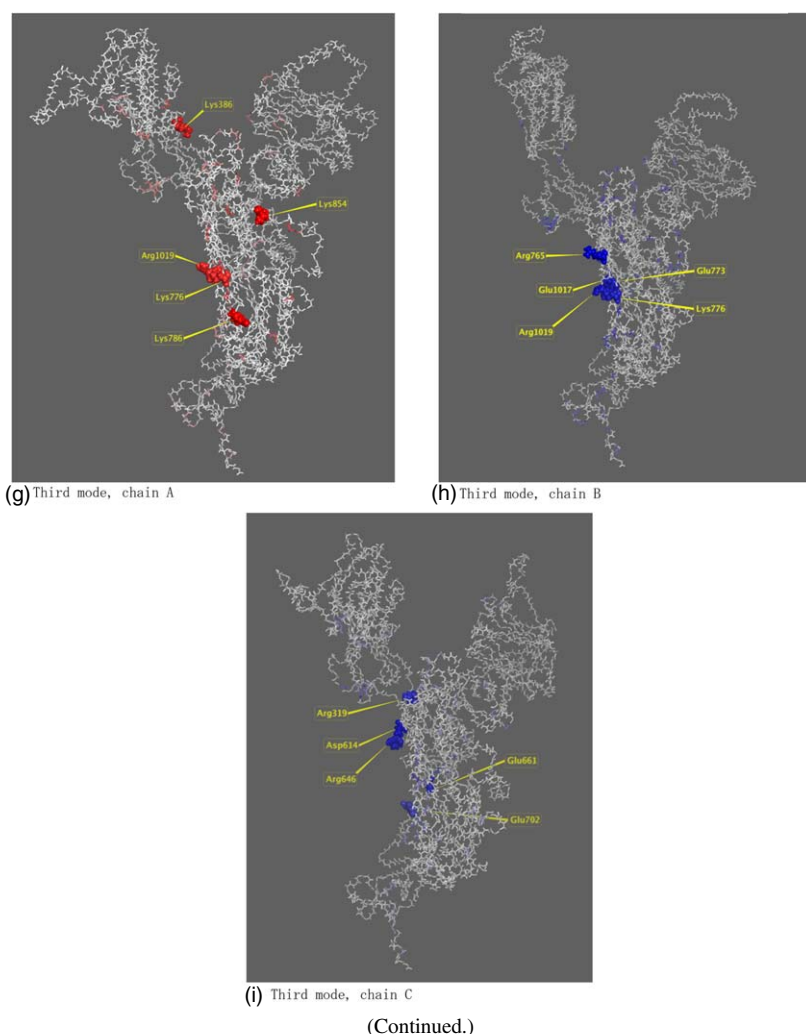


Fig. 3. Key residues with higher amplitudes in each eigenvector of SVD analysis for the open structure of SARS-CoV-2 spike protein trimer. Red and blue colors refer to the positive and negative values of vector components, respectively, whose amplitudes are indicated by their deepness. Results for (a)–(c) first, (d)–(f) second, and (g)–(i) third eigenmodes with higher eigenvalues are depicted for the chains A–C.



intermolecular interactions such as π - π , CH- π , and cation- π interactions. Furthermore, the proposed methodology on the basis of tensor decomposition^{20,21)} has more general applicability also including the analysis on experimental or observational results with high complexities.

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 (g9330001). Technical supports on Fugaku were provided by Drs. Kota Sakakura (FOCUS) and Hiromasa Watanabe (HPC Systems Inc.). The IFIE data curation was supported by Dr. Hirotaka Kondo (VRAIDE G.K.). MD-based structure relaxations were performed on TSUBAME3.0 at Tokyo Institute of Technology under a funding by AMED/BINDS (JP20am0101113). The present work was also funded by Rikkyo SFR, JSPS Kakenhi (Nos. 17H06353, 18K03825, 19K12010, 21K06098) and Quantum Leap Flagship Program (MEXT Q-LEAP, Grant Number JPMXS0120330644).

ORCID iDs Koji Okuwaki <https://orcid.org/0000-0002-4510-5717>
 Yuji Mochizuki <https://orcid.org/0000-0002-7310-5183>
 Kaori Fukuzawa <https://orcid.org/0000-0001-5357-8250>
 Yuto Komeiji <https://orcid.org/0000-0002-9708-6175>
 Shigenori Tanaka <https://orcid.org/0000-0002-6659-2788>

- 1) E. Callaway, *Nature* **525**, 172 (2015).
- 2) J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, and S. H. W. Scheres, *eLife* **7**, e42166 (2018).
- 3) D. G. Fedorov and K. Kitaura (ed.) *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems* (CRC Press, Boca Raton, FL, 2009).
- 4) D. G. Fedorov, T. Nagata, and K. Kitaura, *Phys. Chem. Chem. Phys.* **14**, 7562 (2012).
- 5) S. Tanaka, Y. Mochizuki, Y. Komeiji, Y. Okiyama, and K. Fukuzawa, *Phys. Chem. Chem. Phys.* **16**, 10310 (2014).
- 6) Y. Mochizuki, S. Tanaka, and K. Fukuzawa (ed.) *Recent Advances of the Fragment Molecular Orbital Method* (Springer, Singapore, 2021).
- 7) 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).
- 8) F. Xu, S. Tanaka, H. Watanabe, Y. Shimane, M. Iwasawa, K. Ohishi, and T. Maruyama, *Viruses* **10**, 236 (2018).
- 9) H. Lim, A. Baek, J. Kim, M. S. Kim, J. Liu, K. Y. Nam, J. H. Yoon, and K. T. No, *Sci. Rep.* **10**, 1 (2020).
- 10) C. Watanabe, Y. Okiyama, S. Tanaka, K. Fukuzawa, and T. Honma, *Chem. Sci.* **12**, 4722 (2021).
- 11) K. Watanabe, C. Watanabe, T. Honma, Y.-S. Tian, Y. Kawashima, N. Kawashita, T. Takagi, and K. Fukuzawa, *J. Phys. Chem. Lett.* **12**, 4059 (2021).
- 12) K. Akisawa, R. Hatada, K. Okuwaki, Y. Mochizuki, K. Fukuzawa, Y. Komeiji, and S. Tanaka, *RSC Adv.* **11**, 3272 (2021).
- 13) K. Akisawa, R. Hatada, K. Okuwaki, S. Kitahara, Y. Tachino, Y. Mochizuki, Y. Komeiji, and S. Tanaka, *Jpn. J. Appl. Phys.* **60**, 090901 (2021).
- 14) T. Ishikawa, H. Ozono, K. Akisawa, R. Hatada, K. Okuwaki, and Y. Mochizuki, *J. Phys. Chem. Lett.* **46**, 11267 (2021).
- 15) <https://top500.org/>.
- 16) Y. Mochizuki, K. Yamashita, K. Fukuzawa, K. Takematsu, H. Watanabe, N. Taguchi, Y. Okiyama, M. Tsuboi, T. Nakano, and S. Tanaka, *Chem. Phys. Lett.* **493**, 346 (2010).
- 17) M. Pitonak, P. Neogrady, J. Cerny, S. Grimme, and P. Hobza, *Chem. Phys. Chem.* **10**, 282 (2009).
- 18) K. Maruyama, Y. Sheng, H. Watanabe, K. Fukuzawa, and S. Tanaka, *Comput. Theor. Chem.* **1132**, 23 (2018).

- 19) S. Tanaka, C. Watanabe, T. Honma, K. Fukuzawa, K. Ohishi, and T. Maruyama, *J. Mol. Graph. Model.* **100**, 107650 (2020).
- 20) C. J. Appellof and E. R. Davidson, *Anal. Chem.* **53**, 2053 (1981).
- 21) M. E. Kilmer, L. Horesh, H. Avron, and E. Newman, *Proc. Natl. Acad. Sci.* **118**, e2015851118 (2021).
- 22) L. Guruprasad, *Proteins* **89**, 569 (2021).
- 23) J. Kossaifi, Y. Panagakis, A. Anandkumar, and M. Pantic, *J. Mach. Learn. Res.* **20**, 1 (2019).
- 24) A. C. Walls, Y.-J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, and D. Velesler, *Cell* **181**, 281 (2020).
- 25) J. Lan et al., *Nature* **581**, 215 (2020).
- 26) F. Li, W. Li, M. Farzan, and S. C. Harrison, *Science* **309**, 1864 (2005).
- 27) Molecular Operating Environment (MOE) (2019.01), Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite 910, Montreal, QC, Canada, H3A 2R7, 2021.
- 28) D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, *J. Comput. Chem.* **26**, 1668 (2005).
- 29) S. M. C. Gobeil et al., *Cell Rep.* **34**, 108630 (2021).
- 30) A. Khan, T. Zia, M. Suleman, T. Khan, S. S. Ali, A. A. Abbasi, A. Mohammad, and D. Q. Wei, *J. Cell. Physiol.* **236**, 7045 (2021).
- 31) W. T. Harvey et al., *Nature* **424**, 409 (2021).
- 32) Science Brief: Emerging SARS-CoV-2 Variants [<https://cdc.gov/coronavirus/2019-ncov/more/science-and-research/scientific-brief-emerging-variants.html>].
- 33) D. G. Fedorov and K. Kitaura, *J. Comput. Chem.* **28**, 222 (2007).