Intermolecular interaction among Remdesivir, RNA and RNA-dependent RNA polymerase of SARS-CoV-2 analyzed by fragment molecular orbital calculation

Koichiro Katoa, Teruki Honmab, Kaori Fukuzawac,d*

a. Science Solutions Division, Mizuho Information and Research Institute, Inc., 2-3 Kanda Nishiki-cho, Chiyoda-ku, Tokyo 101-8442, Japan
b. RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
c. Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan
d. Department of Biomolecular Engineering, Graduate School of Engineering, Tohoku University, 6-6-11 Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan

ABSTRACT

COVID-19, a disease caused by a new strain of coronavirus (SARS-CoV-2) originating from Wuhan, China, has now spread around the world, triggering a global pandemic, leaving the public eagerly awaiting the development of a specific medicine and vaccine. In response, aggressive efforts are underway around the world to overcome COVID-19. In this study, referencing the data published on the Protein Data Bank (PDB ID: 7BV2) on April 22, we conducted a detailed analysis of the interaction between the complex structures of the RNA-dependent RNA polymerase (RdRp) of SARS-CoV-2 and Remdesivir, an antiviral drug, from the quantum chemical perspective based on the fragment molecular orbital (FMO) method. In addition to the hydrogen bonding and intra-strand stacking between complementary strands as seen in normal base pairs, Remdesivir bound to the terminus of an primer-RNA strand was further stabilized by diagonal π-π stacking with the -1A base of the complementary strand and an additional hydrogen bond with an intra-strand base, due to the effect of chemically modified functional group. Moreover, stable OH/π interaction is also formed with Thr687 of the RdRp. We quantitatively revealed the exhaustive interaction within the complex among Remdesivir, template-primer-RNA, RdRp and co-factors, and published the results in the FMODB database.

Keywords
Coronaviruses, SARS-CoV-2, RNA dependent RNA polymerase, Remdesivir, template-primer RNA, Interaction analysis, Fragment molecular orbital method
1. Introduction

COVID-19 is a disease caused by a new strain of coronavirus SARS-CoV-2 that originated in Wuhan (Hubei, China) in December 2019 [1]. As of May 19, COVID-19 has infected 4,731,458 people and claimed the lives of 316,169 of them worldwide, while showing no signs of containment[2]-[3]. The urgent development of effective drugs and vaccines is essential for humanity to overcome this contagious disease. Being a single stranded positive sense RNA virus, SARS-CoV-2 contains more than 10 types of viral structural and non-structural proteins (SP and NSP) [4]. Researchers across the world have been unremittingly analyzing the x-ray crystal structures and cryo-electron microscopy (cryo-EM) structures of these proteins, and the data on these structures is being continuously published on the Protein Data Bank (PDB). Especially, the Protein Data Bank Japan (PDBj)[5] has posted representative and all entries by category in an easy-to-understand manner on its feature page. The spike (S) glycoproteins protruding from the viral surface play important roles in the infection process from the virus to the host cell [6][7], hence constituting a critical target for vaccines and effective drugs. Other promising potential drug targets include NSPs, such as main protease (Mpro)[8] and RNA-dependent RNA polymerase (RdRp)[9][10]. Additionally, drugs such as Remdesivir[11][12], Favipinavir[13], Nelfinavir/Cepharanthine[14][15], Lopinavir/Ritonavir (KALETRA)[16], and Ivermactin[17] have been attracting attention as candidate compounds for use as effective cures.

In silico analysis using molecular simulations can effectively utilize the published structural data for drug development. Various techniques are being utilized including docking, classical molecular dynamics, and quantum chemistry. In particular, the fragment molecular orbital (FMO) method[18]-[20] is a quantum chemical calculation technique that divides macromolecules such as proteins and nucleic acids into fragments, thereby enabling the high-speed computation of electronic states of whole systems. This method is also effective in analyzing interaction energy between fragments; hence, it has been used frequently in recent years for the quantitative evaluation of complex structures in fields related to drug development[21][22] and structural biology[23][24]. Regarding the in silico analysis of COVID-19-related proteins, the results of classical molecular dynamics simulation have been made available on the Internet[25]-[27]. In addition, docking analyses[15][28] and FMO calculations[28]-[30] are being actively performed. The FMO database (FMODB) developed by our collaborative research team including the authors launched a COVID-19 feature page on April 17[31]. Utilizing a supercomputer, the team has been performing FMO computations one after another and publishing the results. As of May 13, the database has published the FMO calculation
results of 76 complex structures of COVID-19-related proteins.

At present, Remdesivir is one of the most promising drugs for COVID-19, with countries rushing to officially approve its use. Recent studies have revealed the cryo-EM structure of the complex comprising Remdesivir and its target protein RdRp (nsp12)[32]-[34], cofactors (nsp7 and nsp8), and template-primer RNA strands[9]. This paper reports the FMO calculation results of the cryo-EM structure and the results of a quantitative energy analysis, revealing intermolecular interactions between Remdesivir, RdRp, and template-primer RNA.

2. Materials and Methods

The structural data was referenced in the PDBj (PDBID: 7BV2)[9]. Molecular Operating Environment (MOE)[35] software was used for modeling the molecular structures. First, the structures were corrected by the structure preparation function, and hydrogens were added by protonate3D. For the N- and C-termini of proteins, -NH2 and -COOH were capped, respectively. Next, the fixed atomic coordinates were released in stages in order from water molecules, ions, side chains, to main chains, thereby achieving structural optimization under constraint conditions (tether=1). Amber14:EHT was used for the force field, with a cutoff radius of 12 Å. When FMO calculations were performed, the water molecules and ions were deleted. The theoretical level of the FMO calculation was FMO2-MP2/6-31G*[36][37]. The computation program used was ABINIT-MP[20]. Inter-molecular interactions in the vicinity of Remdesivir were analyzed using inter-fragment interaction energy (IFIE), and IFIE were decomposed by pair interaction energy decomposition analysis (PIEDA) into four energy components; electrostatic (ES), exchange repulsion (EX), charge transfer with mixed term (CT+mix) and dispersion (DI) contributions[38]. The unit of fragment division was amino-acid residues for proteins, whereas for RNA, the base moiety and sugar-phosphate backbone were regarded as separate fragments. Although the monophosphate form of Remdesivir forms a covalent bond with RNA, the phosphate and sugar parts of Remdesivir were divided, and the remaining sugar and base parts were treated as one fragment (Figure S1). The molecular complex used for the computations comprised Remdesivir, template-primer-RNA duplex, and RdRp-cofactor (nsp12-nsp7-nsp8), consisting of 1,006 residues and 26 bases including Remdesivir. This is to say that the whole molecular size used in the calculation to be 16,770 atoms and 1,057 fragments (Figure 1).

3. Results
Figure 1 (a) shows the structure used in the FMO calculations. From the perspective of steric conformation, the base part of Remdesivir forms not only two hydrogen bonds with the facing +1U’ (as with normal A-U base pairs) but also an additional hydrogen bond with the adjacent intra-strand -1U (Figure 1 (b)). Additional contact is also observed between the nitrile substitution (-C≡N) on the sugar backbone and Thr687 in addition to hydrogen bond between the hydroxyl group of sugar and Asn691 (Figure 1 (c)). Figure 2 shows the results of FMO calculations for these structures. Figure 2 (a) graphs the IFIEs between Remdesivir and amino-acid or RNA fragments, whereas Figure 2 (b) represents the PIEDA energies that are decomposed values of the former. In the protein part, ones whose IFIE absolute value is ± 2 kcal/mol or above are shown. In the RNA part, ones in the vicinity of Remdesivir are shown. The superscript b in the RNA part signifies a base fragment. Based on the surrounding structure of Remdesivir shown in Figure 1, and compared with the FMO energies, strong IFIE interactions are noted with amino-acid fragments Thr687, Asn691, and Asp760. The PIEDA energies of these residues reveal their interaction properties. A major ES contribution and a CT contribution were observed for Thr687 and Asn691, indicating hydrogen bonding property. Meanwhile, the only interaction property for Asp760 was electrostatic (ES) caused by polarized electric charges within the fragments. It should be noted that the ratio of DI contribution was relatively large in Thr687, suggesting OH/π interaction by the π electron in the nitrile group of Remdesivir. The IFIEs of the base fragments revealed strong attractive interactions at the base parts +1U’, -1A’, and -1U. The PIEDA energies of +1U’ base indicated that there is hydrogen bonding interaction with the major ES contribution and non-negligible CT and DI contributions. -1A’ has a π-π stacking interaction mainly by the DI contribution, and -1U has both hydrogen bonding interaction by the ES and CT contributions and π-π stacking interaction by the DI contribution. The strength of these interactions is structurally mapped on Figure 3. Many of the interactions with Remdesivir were hydrogen bonds; therefore, ES (red) was dominant. However, DI (green) was dominant only for -1A’, with the diagonal π-π stacking between the purine rings stabilizing the overall structure. As pointed out in the original article regarding cryo-EM structural analyses[9], there are five Remdesivir binding residues on the RdRp: Lys545, Arg555, Ser682, Asn691, and Asp760. In this study, the FMO calculation results identified Thr687, Asn691, and Asp760 to be critical. As Lys545 was electrostatically repulsive and Arg555 was situated distantly, their contribution was negligible. Ser682 formed a hydrogen bond with +1U of the complementary strand, rather than with Remdesivir. The exhaustive IFIE and PIEDA values of all fragment pairs constituting the RdRp-RNA-Remdesivir complex have been
made available on our FMODB database for viewing and downloading (FMODBID: [31]).

4. Discussion

Figure 4 shows the summary of interactions in the vicinity of Remdesivir, a nucleic acid analog. In previous studies, we reported on FMO calculation results regarding interactions between nucleic acid bases. In double-strand DNA, the IFIE of Watson-Crick hydrogen bond was -40 – -50 kcal/mol for A-T pairs and around -20 kcal/mol for C-G pairs, and the 1,2-stacking energy was approximately -10 – -20 kcal/mol[39]-[41]. In RNA complexes, stacking within a single-strand RNA was an interaction weaker than -10 kcal/mol[23]. Compared with these energy values, the 1,2-stacking energy was -27.43 kcal/mol and the single-strand stacking energy was -17.38 kcal/mol in the Remdesivir base analog interaction. Therefore, the stacking interaction of the base analog Remdesivir was found to be considerably stronger than those of RNA adenine that exists naturally. This is attributed to the carbon-nitrogen atomic substitutions introduced into the adenine skeleton of Remdesivir. Furthermore, regarding the interaction between the sugar moiety of Remdesivir and the RdRp amino-acid residues, the stabilization energy due to the hydrogen bonding and OH/π interactions observed between the nitrile substituent (-C≡N) and Thr687 was -8.13 kcal/mol. This suggests that the sugar moiety of Remdesivir also acquired additional interactions as compared to naturally occurring RNA nucleotides. On the other hand, +1U’ is affected by the twist in the Remdesivir bases, and the intra-strand stacking with +2U’ no longer assumes a ladder-like structure, with the interaction subjected to be repulsive.

The above results demonstrate that performing FMO calculations for the experimentally determined cryo-EM structure of the RdRp-RNA complex helps deepen the structural biological understanding of the complex. Besides, we succeeded in quantitatively evaluating the intermolecular interactions in the vicinity of Remdesivir. We hope that these findings will prove beneficial for the development of COVID-19 therapeutic drugs.

Acknowledgements

The authors thank Dr. Daisuke Takaya, Dr. Chiduru Watanabe and Dr. Kikuko Kamisaka at RIKEN for data registration in the FMO database (FMODB). KF thank to Prof. Yuji Mochiduki (Rikkyo University) and Prof. Shigenori Tanaka (Kobe University) for discussion on FMO calculation od COVID-19. This research was
partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP20am0101113. A part of this research was done in activities of the FMO drug design consortium (FMODD), and the results were obtained using Oakforest-PACS supercomputer system in the HPCI project (hp200101).

References
[1] Alexander E. Gorbalenya, Susan C. Baker, Ralph S. Baric et. al., The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nature Microbiology 5, 536–544(2020).
[2] Coronavirus disease (COVID-19) situation reports. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/
[3] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. 10.1016/S1473-3099(20)30120-1 (2020). doi:10.1016/S1473-3099(20)30120-1 Medline
[4] Masters, Paul S. The molecular biology of coronaviruses, Advances in Virus Research, Volume 66, 2006, Pages 193-292,
[5] Protein Data Bank Japan, https://pdbj.org/
[6] Fang Li, Structure, Function, and Evolution of Coronavirus Spike Proteins, Annual Review of Virology, 3, 237-261 (2016).
[7] Walls, Alexandra C.; Park, Young-Jun; Tortorici, M. Alejandra, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein, Cell, 180, 1-12 (2020).
[8] Zhenming Jin, Xiaoyu Du, Yechun Xu, et. al., Structure of Mpro from COVID-19 virus and discovery of its inhibitors, Nature https://doi.org/10.1038/s41586-020-2223-y (2019).
[9] W. Yin, C. Mao, X. Luan, et. al., Science, 10.1126/science.abc1560 (2020).
[10] Y. Gao, L. Yan, Y. Huang, et. al., Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science 10.1126/science.abb7498 (2020).
[11] E. P. Tchesnokov, J. Y. Feng, D. P. Porter, M. Götte, Mechanism of Inhibition of Ebola Virus RNA-Dependent RNA Polymerase by Remdesivir. Viruses 11, 326 (2019). doi:10.3390/v11040326pmid:30987343
[12] M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, M. Xu, Z. Shi, Z. Hu, W. Zhong, G. Xiao, Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 30, 269–271 (2020).
doi:10.1038/s41422-020-0282-0 Medline
[13]Furuta Y, Komeno T, Nakamura T. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. Proc Jpn Acad, Ser B, Phys Biol Sci. 2017; 93:449-463.

[14]N. Yamamoto, S. Matsuyama, T. Hoshino, N. Yamamoto. Nelfinavir inhibits replication of severe acute respiratory syndrome coronavirus 2 in vitro. bioRxiv preprint doi: https://doi.org/10.1101/2020.04.06.026476.

[15]H. Ohashi, K. Watashi, W. Saso, et. al. Multidrug treatment with nelfinavir and cepharamthine against COVID-19. bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.039925doi:

[16]Chu, C. M., Cheng, V. C. C., Hung, I. F. N.. Role of lopinavir/ritonavir in the treatment of SARS: Initial virological and clinical findings. Thorax 59 (3), 252–256 (2004).

[17]L. Calya, J. D. Drucea, M. G. Catton. The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 in vitro. Antiviral Res. 178, 104787 (2020).

[18]Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules. Chem. Phys. Lett. 313, 701–706 (1999).

[19]Fedorov, D. G.; Nagata, T.; Kitaura, K. Exploring Chemistry with the Fragment Molecular Orbital Method. Phys. Chem. Chem. Phys. 14, 7562-7577 (2012).

[20]S. Tanaka, Y. Mochizuki, Y. Komeiji, et. al. Electron-correlated fragment-molecular-orbital calculations for biomolecular and nano systems. Phys. Chem. Chem. Phys., 16, 10310-10344 (2014).

[21]C. Watanabe, H. Watanabe, K. Fukuzawa, et. al. Theoretical analysis of activity cliffs among benzofuranone class Pim1 inhibitors using the fragment molecular orbital with molecular mechanics Poisson-Boltzmann surface area (FMO+MM-PBSA) method. J. Chem Info Model, 57, 2996-3010 (2017). DOI: 10.1021/acs.jcim.7b00110

[22]Y. Sheng, H. Watanabe, K. Maruyama, et. al. Toward good correlation between FMO interaction energies and experimental IC50 for ligand binding: A case study of p38 MAP kinase. Comput Struct Biotech,16, 421-434 (2018)

[23]S. Iwasaki, W. Iwasaki, M. Takahashi, et. al. The Translation Inhibitor Rocaglamide Targets a Bimolecular Cavity between elf4A and Polypurine RNA, Molecular Cell, 73, 1-11 (2018). DOI: 10.1016/j.molcel.2018.11.026.

[24]K. Koiwai, K. Inaba, K. Morohashi,et. al. An integrated approach unravels a crucial structural property for the function of the insect steroidogenic Halloween protein Noppera-bo. J. Biol. Chem., in press.
[25] D. E. Shaw Research, "Molecular Dynamics Simulations Related to SARS-CoV-2," D. E. Shaw Research Technical Data, 2020.
[26] folding@home, https://foldingathome.org/
[27] T. Komatsu, Y. Koyama, N. Okimoto, G. Morimoto, Y. Ohno, M. Taiji. COVID-19 related trajectory data of 10 microseconds all atom molecular dynamics simulation of SARS-CoV-2 dimeric main protease. Mendeley Data, 10.17632/vpps4vhryg.1
[28] B. Nutho, P. Mahalapbutr, K. Hengphasatporn, et. al. Why Are Lopinavir and Ritonavir Effective against the Newly Emerged Coronavirus 2019? Atomic Insights into the Inhibitory Mechanisms. Biochemistry 59, 1769-1779 (2020). DOI: 10.1021/acs.biochem.0c00160.
[29] R. Hatada, K. Okuwaki, Y. Mochizuki et. al. Fragment molecular orbital based interaction analyses on COVID-19 main protease - inhibitor N3 complex (PDB ID:6LU7). ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.11988120.v1
[30] H. Lim, A. Baek, J. Kim et. al. Hot spot profiles of SARS-CoV-2 and human ACE2 receptor protein-protein interaction obtained by densityfunctional tight-binding fragment molecular orbital method. Nat. Res. In review. DOI:10.21203/rs.3.rs-17218/v1
[31] FMO database (FMODB) https://drugdesign.riken.jp/FMODB/
[32] J. Ziebuhr, The coronavirus replicase. Curr. Top. Microbiol. Immunol. 287, 57–94 (2005). doi:10.1007/3-540-26765-4_3
[33] A. J. te Velthuis, Common and unique features of viral RNA-dependent polymerases. Cell. Mol. Life Sci. 71, 4403–4420 (2014). doi:10.1007/s00018-014-1695-z pmid:25080879
[34] S. Venkataraman, B. V. L. S. Prasad, R. Selvarajan, RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. Viruses 10, 76 (2018). doi:10.3390/v10020076pmid:29439438
[35] Molecular Operating Environment (MOE); Chemical Computing Group Inc.: 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.
[36] Mochizuki, Y.; Koikegami, S.; Nakano, T.; Amari, S.; Kitaura, K. Large Scale MP2 Calculations with Fragment Molecular Orbital Scheme. Chem. Phys. Lett. 396, 473–479 (2004).
[37] Mochizuki, Y.; Nakano, T.; Koikegami, et. al. A Parallelized Integral-Direct Second-Order Møller-Plesset Perturbation Theory Method with a Fragment Molecular Orbital Scheme. Theor. Chem. Acc. 112, 442–452 (2004).
[38] Fedorov, D. G.; Kitaura, K. Pair Interaction Energy Decomposition Analysis. J.
Comput. Chem. 28, 222–237 (2007).

[39] K. Fukuzawa, I. Kurisaki, C. Watanabe, et. al. Explicit solvation modulates intra- and inter-molecular interactions within DNA: electronic aspects revealed by the ab initio fragment molecular orbital (FMO) method. Comp. Theor. Chem., 1054, 29-37 (2015).

[40] H. Yamada, Y. Mochizuki, K. Fukuzawa, Y. Okiyama and Y. Komeiji. Fragment molecular orbital (FMO) calculations on DNA by a scaled third-order Møller-Plesset perturbation (MP2.5) scheme. Comp. Theor. Chem., 1101, 46-54 (2017)

[41] A. Rossberg, T. Abe, K. Okuwaki, et. al. Destabilization of DNA through interstrand crosslinking by UO$_2^{2+}$ Chem. Commun., 55, 2015-2018 (2019). DOI: 10.1039/c8cc09329f

Figure/Table Legends

Figure. 1
(a) The overall structure of the 7BV2 complex used for the FMO calculations. Remdesivir is represented as an atomic-colored CPK model. Both the double-strand structure and atoms of RNA are shown in green. The magnified images of the Remdesivir surrounding structure are atomic-colored on both the RNA side (b) and the RdRp side (c). Atomic colors are as follows: carbon: gray; oxygen: red; nitrogen: blue; phosphorus: pink; and hydrogen: white. The base names for the RNA template strand are marked with an apostrophe (’).

Figure. 2
In this graph of the IFIEs (a) and their energy decomposition analysis (PIEDA) (b), only the names of proteins with an interaction energy of ± 2 kcal/mol or above and the names of RNA in the vicinity of Remdesivir are listed. The superscript b beside RNA signifies a fragment in the sugar-base part.

Figure. 3
This figure visualizes the main components of PIEDA energy in the vicinity of Remdesivir. Remdesivir is the yellow-colored ball and stick model, and the other structures are represented as the stick models contrasted depending on the intensity of interaction. When the main component of the interaction with Remdesivir is electrostatic interaction (ES), the attractive and repulsive energies are colored in red,
and blue, respectively. When the main component is dispersion interaction (DI), the energy is colored in green.

Figure. 4
A diagram of interaction energies between Remdesivir, nearby residues, and bases. The square frames represent RNA bases, and the rounded frames signify RdRp residues. The value beside each arrow represents an interaction energy value, with the unit of kcal/mol. The red lines specifically indicate stabilization by hydrogen bonding. The green lines indicate stabilization by π-π stacking interaction between -1U and -1A’ and stabilization by OH/π interaction with Thr687.

Supplementary Material
Figure S1
Fragmentation for Remdesivir.
