Chem. Pharm. Bull.
Communication

*In vitro and in silico* studies of potential coronavirus-specific 3C-like protease inhibitors

Shoji Maehara*, Manami Kobayashi, Mari Kuwada, Tominari Choshi, Hirohumi Inoue, Yuzou Hieda, Takashi Nishiyama, and Toshiyuki Hata

*Faculty of Pharmacy and Pharmaceutical Sciences, Fukuyama University, Sanzo, 1 Gakuen-cho, Fukuyama, Hiroshima 729-0292, Japan*

smaehara@fukuyama-u.ac.jp

*To whom correspondence should be addressed.
E-mail: smaehara@fukuyama-u.ac.jp
Faculty of Pharmacy and Pharmaceutical Sciences, Fukuyama University, Sanzo, 1 Gakuen-cho, Fukuyama, Hiroshima 729-0292, Japan*
Summary
We investigated similar compounds to ebselen and tideglusib, which exhibit strong activity against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), using MACCS keys. Four candidate compounds were identified. One of them, phenylbenzothiazol-3-one, showed coronavirus-specific 3C-like (3CL) protease inhibitory activity. The results indicated that a similarity score above 0.81 is a good indicator of activity for ebselen-and-tideglusib-like compounds. Subsequently, we simulated the ring-cleavage Michael reaction of ebselen at the Se center, which is responsible for its 3CL protease inhibitory activity, and determined the activation free energy of the reaction. The results showed that reaction simulation is a useful tool for estimating the activity of inhibitory compounds that undergo Michael addition reactions with the relevant cysteine S atom of 3CL proteases.

Keywords: severe acute respiratory syndrome coronavirus 2; 3C-like protease; phenylbenzothiazol-3-one; ebselen; molecular modeling
**Introduction**

Since the end of 2019, coronavirus disease (COVID)-19 caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has spread worldwide. The spread of COVID-19 has caused numerous problems around the world, such as increases in the numbers of seriously ill patients and patient deaths, overwhelmed medical services, and postinfection sequelae. An mRNA vaccine for this infectious disease has been developed, and its administration is underway, but its effects cannot be fully confirmed yet. To date, remdesivir and dexamethasone have been approved as therapeutic agents for severely ill patients, and an antibody cocktail therapy (Ronapreve) has been approved and is being used as a therapeutic agent for mildly ill patients. However, to date, there is no treatment for asymptomatic individuals. In addition, mutations in the spike gene of the causative virus SARS-CoV-2 can occur, leading to ongoing concerns about the acquisition of resistance to current vaccines that target spikes. Accordingly, the development of effective preventive methods and therapeutic agents for COVID-19 is an urgent and ongoing challenge.

In the present study, we datamined research papers concerning COVID-19, which has been widespread since its outbreak in 2019; SARS, which was first identified in 2002; and 3C-like (3CL) proteases specific to coronavirus. This protease is the major viral proteinase that regulates the activity of the coronavirus replication complex; therefore, it is an attractive target for treatment. Information on nitrogen-containing organic compounds with inhibitory activity was extracted. There are numerous reports on the causal relationship between structural similarity and activity. Accordingly, we focused on compounds similar to ebselen and tideglusib based on the similar property principle that we have obtained by organic synthesis. Ebselen, a seleno organic compound, exhibits glutathione peroxidase-like activity. It is one of the promising synthetic antioxidants. It exhibits bioactivity by forming a seleno sulfide bond with the thiol group of the cysteine (Cys) residue of the protein. Studies on the conformation of the 3CL protease of SARS-CoV-2 and ebselen using molecular modeling have confirmed the fitting of ebselen to the active site of this enzyme. In addition, molecular dynamics simulations and LCMS showed that ebselen is located in Cys145 of 3C-like protease in the catalytic cavity via a selenosulfide bond.

This report presents the 3CL protease inhibitory activities of these analogs and demonstrates potency prediction using the molecular modeling simulation method.

**Results and Discussion**

The published compounds considered in this study are ebselen, tideglusib, and PX-12, which are active against SARS-CoV-2 3CL protease and 45240, which is active against the 3CL protease of SARS-CoV, the causative SARS virus. The similarity between these chemical structures and those of the compounds we synthesized was evaluated using the MACCS keys program. We focused on three compounds (compounds 1, 2, and 3) which have a Tanimoto coefficient similarity score of 0.6 or higher (Fig. 1).

For the candidate compounds, activity tests were performed using 3CL protease assay kits (BPS Bioscience), and inhibition curves were plotted. Their IC50 values were calculated from the obtained inhibition curves. The 3CL protease inhibitory activities of compounds 1 and 3 are low (IC50 > 500 μM). However, compound 2, which has a similarity score of 0.81 (vs. ebselen) shows 3CL protease inhibitory activity (Fig.
This result suggests that a MACCS key similarity score of 0.81 or higher is an indicator of SARS-CoV-2 3CL protease inhibitory activity.

Therefore, based on candidate compound 1 and 2, which have high similarity scores, the 3CL protease inhibitory activities of compounds in which the Se atom of ebselen is replaced with C or S and in which the phenyl group is replaced with hydrogen were measured. The results are shown in Fig. 3.

Thus, the compound in which X is a S atom and the R group is a phenyl ring has the highest 3CL protease inhibitory activity of the compounds considered.

The activity of ebselen is based on a mechanism in which the Se atom undergoes an addition reaction with the S atom of a cysteine residue and the Se-containing five-membered ring opens. This mechanism has been demonstrated by X-ray crystal analysis of 3CL protease bound to ebselen as an inhibitor. The 3CL protease inhibitory activity of the ring-cleaved version of compound 2, which showed the strongest activity in the above activity test, was thus measured.

Interestingly, the ring-closed compound 2 (IC\textsubscript{50} = 93 μM) shows higher inhibitory activity than the ring-opened compound 7 (IC\textsubscript{50} = 419 μM; Fig. 4). Thus, it was concluded that the five-member-ring-closed compound is attacked by the S atom of a cysteine residue in the 3CL protease, and that a Michael reaction in which ring-opening and S addition happen simultaneously occurs.

Subsequently, we performed a molecular modeling simulation of this ring-opening reaction and calculated its predicted activation energy. We determined that the ease with which the ring-opening reaction proceeds is the source of the 3CL protease inhibitory activity.

First, a ring-cleavage reaction simulation was performed based on ebselen. As a result of modeling using density functional theory (DFT) at the ωB97XD/6-31 + G(d) level and analyzing the data with the artificial force induced reaction (AFIR) method, we were able to propose the pathway of the reaction (Fig. 5).

Next, ring-cleavage simulations were also performed for Se(ebselen)-S and S(compound 2)-S, and the activation free energies were calculated. As a result, it was found that the reaction of ebselen has the lowest activation energy of 15.4 kcal/mol and that of candidate compound 2 has an activation energy of 24.3 kcal/mol. It was assumed that the candidate compound having a C atom instead of Se would not undergo this reaction (Table 1).

The 3CL protease inhibitory activity of compound 2 was reported to be 0.49 μM for IC\textsubscript{50} in different evaluation systems. In our result compound 2 has a lower 3CL protease inhibitory activity than ebselen. It is suggested that the activation energy for the simulated five-membered-ring-cleavage reaction is an indicator of 3CL protease inhibitory activity for such compounds. This correlation may aid the design and preparation of derivatives of compound 2 as coronavirus-specific 3CL protease inhibitors.

Conclusions

The 3CL protease inhibitory activity of ebselen and its analogs is due to the Se atom of the five-membered ring. No activity is observed when the Se atom is replaced with a C or an O atom, but activity is observed when it is replaced with an S atom. We also established that these activities are not due to reaction of the ring-opened state, but rather to the addition of S with the simultaneous opening of the five-membered ring.
We also established that the activity of phenyl-benzothiazol-3-one (2) is lower than that of ebselen because the activation activity of the corresponding ring-opening reaction is higher for 2 than for ebselen. This demonstrates a new approach to activity prediction for such compounds.

Accordingly, we have identified compound 2 as a lead compound. This compound has a simple chemical structure and it is not difficult to synthesize derivatives. Moreover, a recent study revealed that the ebselen derivative MR-6-7 has potent 3CL protease inhibitory activity. Based on this structure, we propose a new candidate compound. The similarity score of this new candidate compound is calculated to 0.83. (Fig. 6) Thus, it may exhibit potent 3CL protease inhibitory activity.

**Experimental**

**Calculation of similarity score using the MACCS Keys**

The chemical structures of the test compounds were converted to simplified molecular input line entry system format. The Tanimoto coefficient was used to as a similarity score. The Rdkit program was used to calculate the similarity score by using MACCS keys.

**Activation free energy calculation for the five-membered-ring-cleavage reaction using the AFIR method**

A reaction paths for the five-membered ring cleavage were explored by the AFIR method, then these paths were improved using by the local updated planes (LUP) method. As a result of the LUP path, a candidate structures for the equilibrium and transition states were obtained respectively. Finally, all structures were fully optimized by the ωB97XD/6-31 + G(d). The Gibbs free energies were evaluated at 298.15 K and 1 atm. All calculations were performed using the GRRM14 combined with the Gaussian09.

**SARS-CoV-2 3CL protease inhibition**

SARS-CoV-2-specific 3CL-protease-MBP-tagged assay kits were purchased from BPS Bioscience and used according to the manufacturer’s protocol. First, 4 ng/μL of 3CL protease-MBP-tagged enzyme in 30 μL of assay buffer was preincubated with 10 μL of the test compounds (final concentrations of 0, 0.5, 1.5, 5.0, 15, 150, and 500 μM in a final volume of 50 μL) for 30 min at 25°C. The enzymatic reaction was initiated by adding 10 μL of substrate solution. The reaction mixture was incubated for 15 h at 25°C. Wells with 1% DMSO, 4 ng/μL enzyme, and 50 mM substrate served as positive controls with no enzyme inhibition. Wells with GC376 (final concentrations of 2.5, 7.5, 25, 75, and 250 μM) served as the standard inhibitor and negative control. The fluorescence intensity of each reacted fraction was measured using an Infinite 200 Pro microplate-reading fluorimeter (TECAN). Based on the obtained fluorescence intensity, dose-response curves were created using the statistical software R to determine the IC$_{50}$. Compounds 1, 2, 3, 4, and 7 were synthesized at Fukuyama University. Compound 5 and 6 were obtained from Tokyo Chemical Industry (compound 5) and Kanto Chemical Co., Inc. (compound 6).
Acknowledgments
This research project was conducted as part of the 2020 Emergency Support: New Coronavirus Infection Control Subsidy program with support from the Nakatani Foundation. For more information on the Nakatani program, see https://www.nakatani-foundation.jp/achievements/covid-19_achievements_list/.

Conflict of Interest
The authors declare no conflict of interest.
References
1) Anand K., Ziebuhr J., Wadhani P., Mesters J. R., Hilgenfeld R., Science., 300, 1763–1767 (2003).
2) Jin Z., Du X., Xu Y., Deng Y., Liu M., Zhao Y., Zhang B., Li X., Zhang L., Peng C., Duan Y., Yu J., Wang L., Yang K., Liu F., Jiang R., Yang X., You T., Liu X., Yang X., Bai F., Liu H., Liu X., Guddat L. W., Xu W., Xiao G., Qin C., Shi Z., Jiang H., Rao Z., Yang H., Nature, 582, 289–293 (2020).
3) Kotouge R., Nishiyama T., Yamauchi A., Ono K., Hatae N., Oikawa T., Hibino S., Choshi T., Heterocycles, 100, 25–45 (2020).
4) Tsuruta Y., Inoue H., Anal. Biochem., 265, 15–21 (1998).
5) Nakamura Y., Feng Q., Kumagai T., Torikai K., Ohigashi H., Osawa T., Noguchi N., Niki E., Uchida K., J. Biol. Chem., 277, 2687–2694 (2001).
6) Menéndez C. A., Byléhn F., Perez-Lemus G. R., Alvarado W., de Pablo J. J., Sci. Adv., 6, 345–356 (2020).
7) Ampordonai K., Meng X., Shang W., Jin Z., Rogers M., Zhao Y., Zhao R., Liu Z. J., Yang H., Zhang L., O’Neill P. M., Samar Hasnain S., Nat. Commun., 12, 1–7 (2021).
8) Kuo C. J., Liu H. G., Lo Y. K., Seong C. M., Lee K. I., Jung Y. S., Liang P. H., FEBS Lett., 583, 549–555 (2009).
9) Maeda S., Morokuma K., J. Chem. Theory Comput., 7, 2335–2345 (2011).
10) Maeda S., Morokuma K., J. Chem. Phys., 132, 241102 (2010).
11) Sun L. Y., Chen C., Su J., Li J. Q., Jiang Z., Gao H., Chigan J. Z., Ding H. H., Zhai L., Yang K. W., Bioorg. Chem., 112, 1–10 (2021).
12) Maeda S., Harabuchi Y., Osada Y., Taketsugu T., Morokuma K., Ohno K., <http://grrm.chem.tohoku.ac.jp/GRRM/>.
13) Maeda S., Ohno K., Morokuma K., Phys. Chem. Chem. Phys., 15, 3683–3701 (2013).
14) Ohno K., Maeda S., J. Phys. Chem. A, 110, 8933–8941 (2006).
15) Maeda S., Ohno K., Maeda S., Ohno K., JPCA, 109, 5742–5753 (2005).
16) Ohno K., Maeda S., Chem. Phys. Lett., 384, 277–282 (2004).
17) Gaussian 09, Revision E.01. Frisch M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., Scalmani G., Barone V., Mennucci B., Petersson G. A., Nakatsuji H., Caricato M., Li X., Hratchian H. P., Izmaylov A. F., Bloino J., Zheng G., Sonnenberg J. L., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Vreven T., Montgomery J. A., Jr., Peralta J. E., Ogliaro F., Bearpark M., Heyd J. J., Brothers E., Kudin K. N., Staroverov V. N., Kobayashi R., Normand J., Raghavachari K., Rendell A., Burant J. C., Iyengar S. S., Tomasi J., Cossi M., Rega N., Millam J. M., Klene M., Knox J. E., Cross J. B., Bakken V., Adamo C., Jaroni L., Gomperts R., Stratmann R. E., Yazyev O., Austin A. J., Cammi R., Pomelli C., Ochterski J. W., Martin R. L., Morokuma K., Zakrzewski V. G., Voth G. A., Salvador P., Dannenberg J. J., Dapprich S., Daniels A. D., Farkas Ö., Foresman J. B., Ortiz J. V., Cioslowski J., Fox D. J., Gaussian, Inc., Wallingford CT, 2009.
Table 1. Activation free energies for the ring-cleavage reactions of ebselen and compound 2 obtained by DFT simulations and their corresponding IC$_{50}$ values for 3CL protease inhibition

| Reactant     | Type | Activate ΔG (kcal/mol) | 3CL Protease inhibition (IC$_{50}$, μM) |
|--------------|------|------------------------|----------------------------------------|
| ebselen      | Se-S | 15.4                   | 0.67                                   |
| Compound 2   | S-S  | 24.3                   | 93.0                                   |
| Compound | Structure | Similarity Score |
|----------|-----------|-----------------|
| ebselen  | ![Structure](image) | 0.74            |
| tideglusib | ![Structure](image) | 0.41            |
| PX-12    | ![Structure](image) | 0.26            |
| 45240    | ![Structure](image) | 0.64            |
| Compound-1 | ![Structure](image) | 0.81            |
| Compound-3 | ![Structure](image) | 0.24            |
| Compound-2 | ![Structure](image) | 0.09            |
| Compound-4 | ![Structure](image) | 0.24            |

**Figure 1.** Literature compounds considered in this study and candidate compounds 1–4, with their corresponding similarity scores.
Figure 2. Dose-response curves, similarity scores, and IC₅₀ values for candidate compounds 1–3.
Figure 3. Dose-response curves and structures of candidate compounds 1, 2, 5, and 6.
Figure 4. Structures and IC$_{50}$ values of compound 2 and its ring-cleaved derivative compound 7.
**Figure 5.** Free energy diagram for the reaction of ebselen and compound 2 with a cysteine residue.
Figure 6. Structures of MR6-7-2 and a corresponding novel candidate compound.