Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Identification of non-covalent 3C-like protease inhibitors against severe acute respiratory syndrome coronavirus-2 via virtual screening of a Korean compound library

Joo-Youn Lee, a, 1 Chih-Jung Kuo, b, 1 Jin Soo Shin, a Eunhye Jung, a Po-Huang Liang, c, d, 1, Yoon-Sik Jung, a, e, 1

a Infectious Diseases Therapeutic Research Center, Korea Research Institute of Chemical Technology, Daejeon 34114, Republic of Korea
b Department of Veterinary Medicine, National Chung Hsing University, Taichung 40227, Taiwan
c Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan
d Institute of Biochemical Sciences, National Taiwan University, Taipei 10617, Taiwan
e Department of Medicinal Chemistry and Pharmacology, University of Science and Technology, Daejeon 34113, Republic of Korea

ARTICLE INFO

Keywords:
Coronavirus disease 2019
Severe acute respiratory syndrome coronavirus 2
3C-like protease
Inhibitors
Antivirals

ABSTRACT

The outbreak of coronavirus (CoV) disease 2019 (COVID-19), a pandemic caused by the severe acute respiratory syndrome CoV-2 (SARS-CoV-2), has triggered the development of antivirals to treat patients and a vaccine to prevent the disease. To identify 3C-like protease inhibitors that can suppress SARS-CoV-2 replication, we performed a virtual screening of 500,282 compounds in a Korean compound bank. We then subjected the top computational hits to inhibitory assays against 3C-like protease in vitro, leading to the identification of a class of non-covalent inhibitors. Among these inhibitors, compound 7 showed an EC50 of 39.89 μM against SARS-CoV-2 and CC50 of 453.5 μM. This study provides candidates for the optimization of potent 3C-like protease inhibitors showing antiviral effects against SARS-CoV-2.

As of May 6, 2021, the coronavirus (CoV) disease 2019 (COVID-19) has caused 155,813,360 confirmed cases and 3,254,882 deaths (https://covid19.who.int/). The first case was reported in December 2019. The causative agent of COVID-19 is homologous to the severe acute respiratory syndrome (SARS)-associated CoV (SARS-CoV) that caused an outbreak in 2002–2003, and it was thereby named SARS-CoV-2 by the World Health Organization. Another outbreak that occurred in the Middle East in 2012 and spread to South Korea in 2015 was also caused by a human CoV known as the Middle East respiratory syndrome CoV (MERS-CoV). Targeting the enzymes essential for viral replication and the lifecycle of SARS-CoV-2 is a promising strategy for clinical therapy. At present, remdesivir treatment has shown a marginal (68%) benefit in clinical trials for patients with COVID-19, and it was thereby approved for use by the USA-Food and Drug Administration (FDA). It was previously shown to inhibit SARS-CoV and MERS-CoV, and it also antagonizes SARS-CoV-2 replication at a half-maximal effective concentration (EC50) of 0.7 μM by targeting the viral enzyme RNA-dependent RNA polymerase.

A key protease called 3C-like protease (3CL(pro)) undergoes autocleavage and then cleaves 11 sites on polyproteins generated by CoVs inside the host cells, which is essential for viral replication; therefore, it has been used as a target for developing antivirals against SARS-CoV, MERS-CoV, and SARS-CoV-2. 3CL(pro) is a chymotrypsin-like enzyme; however, it utilizes the His-Cys catalytic dyad to cleave...
conserved sequences of (Leu, Met, Phe)-Gln1(Ser, Ala, Gly) on the polypeptides of SARS-CoV. Due to the essential role of the protease, several rationally designed peptidomimetics based on its substrate specificity, as well as FDA-approved or experimental drugs such as disulfiram, ebenslen, tideglibus, TDZD-8, carmofur, and PX12 have been investigated and shown to inhibit 3CLpro activity and SARS-CoV-2 replication. Moreover, boceprevir, GC-376, and calpain inhibitors II and XII have been demonstrated to inhibit SARS-CoV-2 viral replication by targeting 3CLpro. As reported recently, the cysteine protease inhibitors MDI-28170 (calpain inhibitor III) and ONO 5334 (cathepsin K inhibitor) inhibit SARS-CoV-2 viral replication. Furthermore, presently used pharmaceuticals and herbal medicines have been screened for activity against SARS-CoV-2; several compounds have been identified as inhibitors of 3CLpro activity and SARS-CoV-2 replication. Moreover, by screening collections of 1068 and 2701 FDA drugs for inhibition of the 3CLpro and papain-like protease, respectively, we identified six drugs showing activity against SARS-CoV-2.

Using an artificial intelligence approach, several drugs have been identified that inhibit SARS-CoV-2 replication in vitro. By performing quantitative high-throughput screening of 10,755 approved and investigational drugs and bioactive compounds, 23 small molecules have been found to inhibit SARS-CoV-2 3CLpro activity at a half-maximal inhibitory concentration (IC50) ranging from 0.26 to 28.85 µM, among which seven had an anti-SARS-CoV-2 effect. By analyzing the 3CLpro pharmacophore clusters for application in virtual screening of 2122 FDA drugs, we identified nelfinavir and boceprevir for drug repurposing for developing a potential therapy for COVID-19.

**Virtual screening of the KCB library to identify potential 3CLpro inhibitors.** To identify non-peptidomimetic chemical inhibitors of SARS-CoV-2 3CLpro from the KCB library, we performed ligand-based virtual screening using the three-dimensional (3D) shapes of the ligand-complexed SARS-CoV 3CLpro structures, 3 MJ5, 3SN8, 3V3M, 4OVZ, 4WO, and 4TWW, obtained from the Protein Data Bank (PDB) (http://www.rcsb.org). All ligand-based virtual screening was performed using the OpenEye Scientific software (OpenEye Scientific Software, Santa Fe, NM, USA). The database was prepared using OMEGA v3.0.1, which was used to generate a maximum number of 200 low-energy conformers for each molecule in the 500,282-membered KCB library. A 3D shape similarity search of ligands as a query was performed using ROCS v3.2.2, and the top 10,000 hits for each query molecule were selected based on the ROCS_TanimotoCombo score. We then performed a 3D electrostatic properties similarity search for the screened 10,000 compounds using EON v2.2.0, and the top 1,000 hits were selected based on the EON_EnergyCombo score for each of the six query molecules. The sum of 3D shape- and electrostatically similar candidates comprised 6,000 compounds that were finally selected for structure-based virtual screening.

**Table 1**

| Compound | Structure | IC50 (µM) | EC50 (µM) | Compound | Structure | IC50 (µM) | EC50 (µM) |
|----------|-----------|-----------|-----------|----------|-----------|-----------|-----------|
| 1        | ![Structure](image1.png) | 8.7 ± 0.8 | NA        | 2        | ![Structure](image2.png) | 10.9 ± 0.8 | NA        |
| 3        | ![Structure](image3.png) | 8.0 ± 1.4 | 140.9 ± 3.80 | 4        | ![Structure](image4.png) | 9.5 ± 0.7 | NA        |
| 5        | ![Structure](image5.png) | 12.1 ± 2.4 | NA        | 6        | ![Structure](image6.png) | 8.9 ± 0.8 | NA        |
| 7        | ![Structure](image7.png) | 5.7 ± 1.1 | 39.89 ± 0.17 | Remdesivir | NA | 11.1 |

**Remdesivir**

Furthermore, by analyzing the 3CLpro crystal structures, 3 MJ5, 3SN8, 3V3M, 4OVZ, 4WO, and 4TWW, obtained from the Protein Data Bank (PDB) (http://www.rcsb.org). All ligand-based virtual screening was performed using Schrodinger Suite 2020–1 (Schrodinger LLC, New York, NY, USA) and the X-ray crystal structure of SARS-CoV-2 3CLpro in complex with the peptide inhibitor obtained from PDB (ID 6LU7). The protein structure was modified using Protein Preparation Wizard in Maestro, and the receptor grid box was generated at a cubic size of 25 × 25 × 25 Å centered on a complexed ligand. The 6,000 compounds identified via ligand-based virtual screening and the 500,282 members of the KCB library. A 3D shape similarity search of ligands as a query was performed using ROCS v3.2.2, and the top 10,000 hits were selected based on the ROCS_TanimotoCombo score. We then performed a 3D electrostatic properties similarity search for the screened 10,000 compounds using EON v2.2.0, and the top 1,000 hits were selected based on the EON_EnergyCombo score for each of the six query molecules. The sum of 3D shape- and electrostatically similar candidates comprised 6,000 compounds that were finally selected for structure-based virtual screening.

**Structure-based virtual screening was performed using Schrodinger Suite 2020–1 (Schrodinger LLC, New York, NY, USA) and the X-ray crystal structure of SARS-CoV-2 3CLpro in complex with the peptide inhibitor obtained from PDB (ID 6LU7). The protein structure was modified using Protein Preparation Wizard in Maestro, and the receptor grid box was generated at a cubic size of 25 × 25 × 25 Å centered on a complexed ligand. The 6,000 compounds identified via ligand-based virtual screening and the 500,282 members of the KCB library were subjected to ligand preparation using the LigPrep module by applying an OPLS3e force field. During the process, tautomers and ionization states at pH 7.0 ± 2.0 were generated using the Epik module. Docking for each chemical library database was performed using the Glide v8.6 program in the Standard Precision mode. Based on visual inspection, 253 compounds were selected for performing 3CLpro assays to determine the active hits.

**Similarity-based virtual screening using two-dimensional fingerprints.** To identify more analogues of the active hits, we performed a second round of virtual screening based on the fingerprint similarity method. The similarity between single pairs of compounds was calculated using Pipeline Pilot 2020 (Dassault Systèmes BIOVIA, San Diego, CA, USA). The KCB library database was screened using a Tanimoto
According to the in silico virtual screening and the 3CL\textsuperscript{pro} inhibition assay, compounds 3 and 7 were identified as the most potent inhibitors in the group. To determine their potential binding modes with 3CL\textsuperscript{pro}, computer modeling was performed for these two compounds. In compound 3 (Fig. 1A), the triazolopyrimidineone ring was found to bind to the active site adjacent to Cys145 via electrostatic interactions, whereas the Gly143 backbone NH formed additional hydrogen bonds with the nitrogen of the ring. The amide bond formed two hydrogen bonds with the backbone NH groups of Thr26 and Gly143. The 1,4-difluoro substituted phenyl ring was located toward Thr24 and Thr26. The 3,5-dimethyl substituted phenyl group was bound to the hydrophobic pocket formed by His41, Met49, and Met165. In compound 7 (Fig. 1B), the benzoimidazopyrimidineone ring was found to bind to the active site adjacent to Cys145 via electrostatic interactions. The carbonyl group of the amide bond formed hydrogen bonds with the backbone NH groups of Gly143 and Ser144. The 1,3-difluoro substituted phenyl ring was located toward Leu27 and Met49 and was shorter than the 1,3-difluoro-phenyl ring in compound 3. The chlorophenyl ring was found to bind to the same hydrophobic pocket that consisted of His41, Met49, and Met165 in compound 3.

**Antiviral activities of the 3CL\textsuperscript{pro} inhibitors.** After identifying a collection of candidate compounds, we next tested their inhibitory activity against 3CL\textsuperscript{pro} using a fluorogenic peptide, Dabcyl-KTSAVLQSGFRKME-Edans, which contains a fluorescence quenching pair; the fluorescence increased when the peptide was cleaved by the protease.\textsuperscript{24} Subsequently, the IC\textsubscript{50} values of the active compounds were determined using reaction mixtures containing 2.5 nM 3CL\textsuperscript{pro} and 6 \textmu M fluorogenic substrate in a buffer of 20 mM Bis-Tris (pH 7.0) in the absence and presence of various concentrations of the inhibitors. All experiments were performed in triplicate and the values are presented as the mean \pm standard deviation (Table 1). The most active were compounds 3 and 7, and their IC\textsubscript{50} values were determined to be 8.0 \pm 1.4 and 5.7 \pm 1.1 \textmu M, respectively.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

The chemical library used in this study was kindly provided by the Korea Chemical Bank (www.chembank.org) of the Korea Research Institute of Chemical Technology, Daejeon, South Korea. The SARS-CoV-2 resource (NCCP No. 43326) for this study was provided by the National Institute of Health, Cheongwon-gun, South Korea.

**Funding**

This work was supported by the Ministry of Science and ICT, Korea [grant number 2020K1A4A7A02094996] to YSJ; and the Ministry of Science and Technology, Taiwan [grant numbers MoST109-2327-B-002-009, MoST109-2745-B-001-001] to PHL. The authors also thank the support from KRICT to YSJ and Academia Sinica, Taiwan to PHL.
**Author contributions**

J.Y.L. performed the in-silico virtual screening and molecular docking studies; C.J.K. conducted the 3CLpro assay; J.S.S. and E.J. carried out the anti-viral EC50 measurements; Y.S.J and P.H.L. wrote the paper.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bmcl.2021.128067.

**References**

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, et al. \( \text{Nature} \) 2012;492:166.
2. Zhang B, Jiang XM, Su H, Li J, Zhao Y, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, et al. \( \text{Lancet} \) 2020;367:1814.
3. Cheung MT, Cheng VC, Chan KH, et al. \( \text{Nature} \). 2020;582:289-292.
4. Jin Z, Zhao Y, Sun Y, Zhang B, Wang H, Wu Y, Zhu Y, Zhu C, Hu T, Du X, Duan Y, Yu J, Yang X, Yang X, Yang K, Liu K, Liu X, Guddat LW, et al. \( \text{Nat Struct Mol Biol} \). 2020;27:529-532.
5. Ma C, Sacco MD, Hurst B, Townsend JA, Hu Y, Szeto T, Zhang X, et al. \( \text{Cell Res.} \) 2020;30:678-692.
6. Riva L, Yuan S, Yin X, Martin-Sancho L, Matsunaga N, Pache L, Burgstaller-Muehlbacher S, De Jesus PD, Terriere P, Hull MV, Chang MW, Chan JF, Cao J, Poon VK, Herbert KM, Cheng K, Nguyen TH, Rubanov A, Fei Y, Nguyen C, Choi A, Rathnasighe R, Schotsaert M, Miotto L, Dejozer M, Zwaaka TP, Sit KY, Martinez-Sobrido L, Liu WC, White KM, Chapman ME, Lendy EK, Gyinne JR, Albrecht R, Ruppini E, Mescar AD, Johnson JR, Benner C, Sun R, Schulze PG, Su AI, et al. \( \text{Science} \). 2020;368:1331-1335.
7. Jin Z, Dai W, Xu X, 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 LW, Xw W, Xiao G, Qin C, et al. \( \text{Nature} \). 2020;582:289-292.
8. Jin Z, Zhao Y, Sun Y, Zhang B, Wang H, Wu Y, Zhu Y, Zhu C, Hu T, Du X, Duan Y, Yu J, Yang X, Yang X, Yang K, Liu K, Liu X, Guddat LW, et al. \( \text{Nat Struct Mol Biol} \). 2020;27:529-532.
9. Ma C, Sacco MD, Hurst B, Townsend JA, Hu Y, Szeto T, Zhang X, et al. \( \text{Cell Res.} \) 2020;30:678-692.