Selective Covalent Targeting of SARS-CoV-2 Main Protease by Enantiopure Chlorofluoroacetamide

Daiki Yamane†, Satsuki Onitsuka†, Suyong Re†, Hikaru Isogai†, Rui Hamada†, Tadanari Hiramoto†, Eiji Kawanishi†, Kenji Mizuguchi‡, Naoya Shindo*†, Akio Ojida†*

†Graduate School of Pharmaceutical Sciences, Kyushu University, 3-1-1 Maidashi, Hikashi-ku, Fukuoka 812-8582, Japan
‡Artificial Intelligence Center for Health and Biomedical Research, National Institute of Biomedical Innovation, Health, and Nutrition, 7-6-8 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan

*Corresponding authors: shindo708@phar.kyushu-u.ac.jp (N.S.) ojida@phar.kyushu-u.ac.jp (A.O.)

Abstract
The pandemic of COVID-2019 has urged the development of antiviral agents against its causative pathogen SARS-CoV-2. The main protease (M\textsuperscript{pro}), a cysteine protease essential for viral replication, is a promising protein target. Here we report an irreversible SARS-CoV-2 M\textsuperscript{pro} inhibitor possessing chlorofluoroacetamide (CFA) as the warhead for covalent modification of M\textsuperscript{pro}. Ugi multi-component reaction employing chlorofluoroacetic acid allowed rapid generation of CFA derivatives, of which diastereomers displayed significantly different inhibitory activity against M\textsuperscript{pro}. We established a protocol for the optical resolution of chlorofluoroacetic acid, which enable the isolation of the stereoisomers of the best CFA compound 18. Kinetic analysis revealed that (R)-CFA is crucial for both binding affinity and the rate of irreversible inactivation. Our findings highlight the prominent influence of the CFA chirality on the covalent modification of cysteine, and provide the basis for improving the potency and selectivity in the development of novel CFA-based covalent inhibitors.

Keywords: SARS-CoV-2, main protease, chlorofluoroacetamide, Ugi multi-component reaction, cysteine
Introduction

The outbreak of the novel coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2),\textsuperscript{1,2} has been posing a serious threat to the global public health and economy. Although efficacious vaccines are now being administered worldwide, novel SARS-CoV-2 variants of interest and concern continue to emerge.\textsuperscript{3} Thus, the development of antiviral agents against SARS-CoV-2 is urgently needed to put an end to the current pandemic.\textsuperscript{4,5} SARS-CoV-2 has a positive-sense single-stranded genomic RNA that encodes polyproteins pp1a and pp1ab. Two cysteine proteases, the main protease (M\textsuperscript{pro}, also known as 3C-like protease, 3CL\textsuperscript{pro}) and the papain-like protease (PL\textsuperscript{pro}), are excised from these polyproteins by autocleavage and further digest polyproteins into nonstructural proteins (nsps), including crucial components of the viral replication–translation machinery. While both proteases are essential for viral replication, M\textsuperscript{pro} has been extensively studied as an attractive drug target against SARS-CoV-2 owing to the predominant role in polyprotein processing and the lack of a human homolog.\textsuperscript{6,7}

M\textsuperscript{pro} recognizes glutamine and a hydrophobic amino acid residue at the S1 and S2 pockets in the active site, respectively, and cleaves amide bonds mainly within the Leu-Gln↓(Ser, Ala, Gly) sequence (↓: cleavage site).\textsuperscript{8} This substrate specificity is conserved across other coronaviruses, including the causative pathogens of past outbreaks, SARS-CoV-1 and Middle-East respiratory syndrome coronavirus (MERS-CoV).\textsuperscript{9,10} Therefore, current efforts to develop SARS-CoV-2 M\textsuperscript{pro} inhibitors are greatly aided by the previous compound designs, especially ones targeting SARS-CoV-1 M\textsuperscript{pro}.\textsuperscript{5,6} The majority of SARS-CoV-2 M\textsuperscript{pro} inhibitors reported to date are peptidomimetics based on the Leu-Gln sequence, combined with an electrophilic warhead, such as α-ketoamide,\textsuperscript{11} aldehyde,\textsuperscript{12-15} ketone,\textsuperscript{16-19} and vinyl sulfone,\textsuperscript{8} for the covalent capture of the sulfhydryl group of catalytic Cys145. PF-00835231, developed by Pfizer, is one of the representative compounds in this category, of which prodrug is now under clinical trials (Figure 1A).\textsuperscript{12} The ketone electrophile of this compound allows reversible covalent engagement with M\textsuperscript{pro}. Other than peptidomimetics, several dipeptide compounds generated by Ugi multi-component reaction (MCR), are reported as SARS-CoV-1 and SARS-CoV-2 M\textsuperscript{pro} inhibitors (Figure 1B).\textsuperscript{20-22}
Covalent inhibition of proteins has been a powerful strategy for achieving potent and sustained pharmacological efficacy.\(^{23}\) In particular, a number of acrylamide-based covalent inhibitors targeting the noncatalytic cysteine of cancer-associated proteins have been clinically approved or under clinical trials for cancer treatment.\(^{24,25}\) Nonetheless, care should be taken to ensure a high degree of target selectivity for minimizing the toxicity risk due to nonspecific off-target labeling.\(^{26}\) Because medications for infectious disease treatment would require a broader margin of safety compared to anticancer agents, this point is crucial for designing covalent M\(^{pro}\) inhibitors. One strategy for achieving high target selectivity is the use of weakly reactive, or “latent” electrophiles,\(^{27,28}\) which are stable toward off-target biomolecules but sufficiently reactive under the microenvironment in the binding pocket of the target. Reversible modification of the targeted amino acid residue represents another useful strategy to eliminate undesired inhibitor–protein adducts.\(^{29,30}\) We have recently introduced α-chloroacetamide (CFA) as a weakly reactive, cysteine-directed warhead useful for covalent inhibition of tyrosine kinases\(^ {31,32}\) and a cysteine protease.\(^ {33}\) The CFA group exhibited weaker intrinsic reactivity toward cysteine thiol than the acrylamides, and the gel-based activity-based protein profiling (ABPP)
experiments revealed that the CFA-based inhibitors exhibited the higher reaction selectivity toward the targeted tyrosine kinase compared to the structurally related acrylamide-based inhibitors. We also demonstrated that the CFA–thiol adduct was gradually hydrolyzed under neutral aqueous conditions to reversibly generate an unmodified free thiol, whereas the CFA–target protein adduct stably existed in the binding pocket sequestered from aqueous solvents. Given these desirable reaction features, we envisioned that CFA-based compound would serve as selective covalent SARS-CoV-2 M\textsuperscript{pro} inhibitors.

In the present study, we report the development of CFA-based novel covalent inhibitors against SARS-CoV-2 M\textsuperscript{pro}. Ugi MCR using chlorofluoroacetic acid allowed rapid synthesis of CFA-appended dipeptides and fast structure–activity relationship (SAR) study by a fluorescence-based enzymatic assay to provide 18 as the optimized CFA compound. Interestingly, we found that the absolute configuration of the chiral center of CFA unit largely affected the inhibitory activity of 18. To gain insight into this phenomenon, we established a simple synthetic procedure for enantiomerically pure chlorofluoroacetic acid and prepared the set of four stereoisomers of 18. Among them, \((R,R)-18\) inhibited M\textsuperscript{pro} with a dramatically higher potency than the other isomers. The enzyme kinetics analysis revealed that the \(R\) configuration of CFA unit of 18 contributes to increasing both the binding affinity and the reaction rate with M\textsuperscript{pro}, suggesting the importance of the chiral design of CFA unit for enzyme inhibition.

Results and Discussion

We chose the dipeptide scaffold, accessible through Ugi MCR, as the starting point of the study. Ugi MCR, a representative multi-component reaction, would generate CFA derivatives in one single step from amine, aldehyde, isocyanide, and chlorofluoroacetic acid. As shown in Figure 1B, several Ugi MCR-generated dipeptides have been reported as M\textsuperscript{pro} inhibitors. ML188-(R), a SARS-CoV-1 M\textsuperscript{pro} inhibitor, was discovered by a high-throughput screen campaign and subsequent structural optimization.\textsuperscript{20} Based on this molecular architecture, the group of Chen and Wang recently reported Jun8-76-3A as a SARS-CoV-2 M\textsuperscript{pro} inhibitor.\textsuperscript{21} Cocrystal structures revealed that the 3-pyridyl group of these compounds interacts with the S1 pocket, which shows a strict selectivity for a glutamine residue of substrates, through hydrogen bonding with His163. 4-tert-Butylphenyl or 4-biphenyl group occupies the S2 pocket that prefers hydrophobic amino acid residues such as leucine and phenylalanine. In addition, the furan-2-carboxamide moiety forms a bifurcated interaction with the backbone NH of Gly143 in the S1’ pocket. While these compounds reversibly bind to M\textsuperscript{pro}, London et al. reported an acrylamide-based covalent inhibitor 1, although the potency did not significantly differ from a noncovalent furan-2-carboxamide 2.\textsuperscript{22}

As the initial step, we prepared the CFA analog of 1. Ugi MCR using racemic chlorofluoroacetic acid proceeded smoothly to afford CFA dipeptide 3, of which diastereomers could be separated by silica gel chromatography (Figure 2A). (Hereafter, the diastereomer that elutes first in normal-phase chromatography is referred to as A, and the later B). The inhibitory activity of compounds against SARS-CoV-2 M\textsuperscript{pro} was assessed by an enzymatic assay using the purified recombinant protein and a fluorogenic substrate Ac-Abu-Tle-Leu-Gln-(4-methylcoumaryl-7-amide) (MCA). This substrate is based on the optimized tetrapeptide sequence reported
by Drag et al.\textsuperscript{8} and allows robust fluorescent reading compared to the FRET-based substrate possessing an Edans/Dabcyl donor–acceptor pair.\textsuperscript{14} We found that the diastereomers 3A and 3B displayed significantly different inhibitory activity against SARS-CoV-2 M\textsuperscript{pro} (Figure 2B). Whereas 3A was virtually inactive (IC\textsubscript{50} > 20 μM), 3B exhibited higher inhibitory activity (IC\textsubscript{50} = 1.05 μM) than acrylamide 1 (IC\textsubscript{50} = 2.62 μM). This result was intriguing because it indicated that the M\textsuperscript{pro} inhibitory activity of 3 is sensitive to the chirality of both CFA and the new α-amino acid stereocenter formed by Ugi MCR. Previous studies have shown that the R configuration of the amino acid α-carbon of ML188 and Jun8-76-3A is crucial for M\textsuperscript{pro} inhibition. We also assessed the M\textsuperscript{pro} inhibitory activity of acetamide 4 and difluoroacetamide 5 and found that both compounds were inactive (IC\textsubscript{50} > 20 μM). These data, taken together, strongly suggest that the covalent attachment at CFA contributes to the M\textsuperscript{pro} inhibitory activity of 3B in a stereospecific manner.

Figure 2. A) Preparation of the diastereomers of 3 by Ugi MCR. B) Inhibitory activity of Ugi MCR-generated dipeptides against recombinant SARS-CoV-2 M\textsuperscript{pro}. The reported IC\textsubscript{50} values for Jun8-76-3A, 1, and 2 are 0.31, 2.95, 2.72 μM, respectively.

Encouraged by this result, we next conducted SAR study to improve the potency of CFA dipeptides (Figure 3A). We first evaluated the effect of the para-substituent on the aniline-derived benzene ring (R\textsuperscript{1}). As this moiety is supposed to occupy the S2 pocket that mainly accommodates a leucine residue, we screened several hydrophobic substituents. The diastereomers of 6–9 were separable by silica gel column chromatography and assayed individually. In all cases, the diastereomer A was inactive (IC\textsubscript{50} > 20 μM). Both a smaller or a larger alkyl group compared to tert-butyl at this position reduced the potency (6B and 7B, IC\textsubscript{50} = 2.42 and 2.68 μM, respectively). 8B possessing a trifluoromethyl group was also less active (IC\textsubscript{50} = 3.44 μM) than 3B. On the other hand, the M\textsuperscript{pro} inhibitory activity of 9B carrying a pentafluorosulfonyl (SF\textsubscript{5}) group, a thermally and chemically
stable bioisostere for a tert-butyl group, was found slightly higher (IC$_{50}$ = 0.90 μM) than that of 3B. We next investigated the effect of the heterocycle derived from the aldehyde building block (R$^2$). When an additional nitrogen atom was installed to the 3-pyridyl group, we observed increased M$^\text{pro}$ inhibitory activity except for 3-pyridazinyl derivative 10 (a diastereomixture, IC$_{50}$ > 20 μM). 4-pyridazine 11 and 2-pyrazine 13 were assayed as mixtures of diastereomers and showed IC$_{50}$ values of 0.75 and 0.30 μM, respectively. 5-Pyrimidine 12 was the only compound of which diastereomers were separable, and 12B exhibited the highest activity in the series (IC$_{50}$ = 0.25 μM). The effect of the isocyanide-derived substituent R$^3$ was also investigated. A tert-butyl group at this position, as can be seen in the structure of ML188-(R), reduced the potency (14B, IC$_{50}$ = 3.44 μM). Replacement of the 3-fluorophenyl group with a 3-pyridyl group slightly increased activity (16, IC$_{50}$ = 0.73 μM), although the diastereomers were inseparable. Introduction of another benzene ring to 16 afforded a separable diastereomixture of 17, however, with a slight loss of activity (IC$_{50}$ = 1.43 μM). Finally, considering the M$^\text{pro}$ inhibitory activity and the separability of diastereomers, we synthesized CFA 18 by employing 4-SF$_5$-phenyl group at R$^1$, 5-pyrimidine at R$^2$, and 2-(3-fluorophenyl)ethyl group at R$^3$ positions, respectively (Figure 3B). 18B exhibited the most potent M$^\text{pro}$ inhibitory activity among the tested CFA derivatives with IC$_{50}$ value of 0.22 μM, although this value was not significantly different from that of 12B. At this point, we determined the relative configuration of the diastereomers of 18. X-ray crystallography revealed that 18B has a syn-configuration, where the 5-pyrimidinyl group and the CFA fluorine atom orient toward the same direction in the drawn structure (Figure S1). Consequently, 18A was determined to be an anti-isomer.
Figure 3. Structure–activity relationship of CFA dipeptides. A) Screening of R\(^1\), R\(^2\), and R\(^3\) substituents for the improvement of the inhibitory activity against SARS-CoV-2 M\(^{pro}\). ND, not determined. B) Best CFA dipeptide 18 and the relative configuration of each diastereomer determined by X-ray crystallography of 18B (see Figure S1).

To obtain the four stereoisomers of 18 in pure form, we next focused on the optical resolution of chlorofluoroacetic acid.\(^{36}\) Wakselman previously reported the preparation of optically active chlorofluoroacetyl chloride based on the chromatographic separation of the diastereomers of CFA 19 derived from (S)-(-)-1-phenethylamine.\(^{37}\) We sought the optical resolution of chlorofluoroacetic acid by a similar approach, however, the isolation of diastereomers of 19 was found laborious in our attempts due to poor separation by silica gel chromatography. Instead, we established a practical procedure to obtain enantiomerically pure chlorofluoroacetic acid employing (R)-(-)-2-phenylglycinol (20) as the resolving agent (Scheme 1). 20 and commercially available ethyl chlorofluoroacetate were mixed without solvent to afford a diastereomeric mixture of CFA 21, which was directly charged on silica gel and subjected to chromatography. The diastereomers 21A and 21B exhibited much improved separation compared to those of 19, and each diastereomer was isolated in multigram scale by a single chromatography. The absolute configuration at CFA was determined to be R for 21B by X-ray crystallography (Figure S2). Pure 21A and 21B were subjected to hydrolysis in refluxing 3M H\(_2\)SO\(_4\)/1,4-dioxane to provide (S)- and (R)-chlorofluoroacetic acid, respectively. Finally, Ugi MCR using each enantiomer of chlorofluoroacetic acid enabled isolation of the four stereoisomers of 18.

Scheme 1. The optical resolution of chlorofluoroacetic acid using (R)-(-)-2-phenylglycinol (20) as the chiral resolving agent. See Supporting Information for the detail.

Having the isolated stereoisomers of 18 in hand, we evaluated the effect of the chirality of each stereocenter on the M\(^{pro}\) inhibitory activity (Figure 4). In the fluorescent enzymatic assay, (R,R)-18 (the former R indicates the absolute configuration at the Ugi MCR-generated stereocenter, and the latter at CFA) displayed dramatically
higher potency with IC\textsubscript{50} value of 0.056 \(\mu\)M compared to the others (Figure 4B and 4D). Among the other three isomers, (S,R)-18 possessing an (R)-CFA warhead showed low micromolar activity (IC\textsubscript{50} = 2.24 \(\mu\)M), whereas isomers obtained from (S)-chlorofluoroacetic acid were weak inhibitors (IC\textsubscript{50} = 10.62 and 14.04 \(\mu\)M for (R,S)- and (S,S)-18, respectively). To elucidate how the chirality at CFA affected the inhibitory activity, we performed kinetic analysis of M\textsuperscript{pro} inhibition by (R,R)-18 and its diastereomers (S,R)- and (R,S)-18.\textsuperscript{18,38} All three compounds inhibited the hydrolysis of the fluorogenic substrate by M\textsuperscript{pro} in a time-dependent manner, indicative of the irreversible mode of action (Figure 4C and S3–S6). The kinetic parameters for the irreversible inactivation of M\textsuperscript{pro} are summarized in Figure 4D. As expected, (R,R)-18 inactivated M\textsuperscript{pro} with much faster rate (\(k_{\text{inact}}/K_i = 4.167\) M\textsuperscript{−1}s\textsuperscript{−1}) than the diastereomers (S,R)- and (R,S)-18 (\(k_{\text{inact}}/K_i = 66.0\) and 20.0 M\textsuperscript{−1}s\textsuperscript{−1}, respectively). (R,R)- and (S,R)-18, both possessing an (R)-CFA warhead, showed similar maximal inactivation rate constant \(k_{\text{inact}}\) (0.0056 and 0.0043 s\textsuperscript{−1}, respectively), whereas the binding constant \(K_i\) (describing the concentration required to achieve the half of the inactivation rate of \(k_{\text{inact}}\)) was significantly different (1.34 and 64.7 \(\mu\)M, respectively). The large difference in \(K_i\) value by the absolute configuration at the Ugi MCR-generated stereocenter was consistent to the previous studies, where the R configuration at this carbon was crucial for the M\textsuperscript{pro} inhibitory activities of ML188 and Jun8-76-3A.\textsuperscript{20,21} Comparing (R,R)- and (R,S)-18, both \(k_{\text{inact}}\) and \(K_i\) values were significantly different. Taken together, these kinetics data demonstrated that the chirality at CFA is important for both in terms of reversible binding and the rate of covalent modification of M\textsuperscript{pro}. We also determined the kinetic parameters of acrylamide 1. Despite the higher intrinsic reactivity of acrylamide toward thiols compared to that of CFA,\textsuperscript{31} the \(k_{\text{inact}}\) value of 1 was significantly smaller than that of (R)-CFA and comparable to (S)-CFA. This may be due to the acrylamide disposition in 1–M\textsuperscript{pro} binding complex is not optimal for Michael addition of Cys145.

**Figure 4.** Comparison of the four stereoisomers of 18. A) The absolute configurations of each stereoisomer (\(R = (3\)-fluorophenyl)ethyl). The R and S configurations at CFA were highlighted in green and magenta,
The selectivity of (R,R)-18 toward SARS-CoV-2 Mpro was evaluated. We first assessed the inhibitory activity of (R,R)-18 toward other cysteine proteases (Figure 5A). (R,R)-18 potently inhibited SARS-CoV-1 Mpro with IC50 value of 0.094 μM. This was not surprising because Mpro is highly conserved between SARS-CoV-1 and SARS-CoV-2 (96% identity of the amino acid sequence). (R,R)-18 did not inhibit SARS-CoV-2 PLpro (<1% inhibition at 20 μM), indicating its specificity toward Mpro between two proteases required for polyprotein processing. We also evaluated the inhibition of two human cysteine proteases, cathepsins B and L. Both proteases were not inhibited by 20 μM of (R,R)-18, whereas Ugi MCR-generated, noncovalent furan-2-carboxamide compound Jun8-76-3A, moderately inhibited cathepsin B (55% inhibition at 20 μM). To assess the Mpro selectivity of CFA dipeptides among the proteome, we designed a clickable activity-based probe (ABP) by replacing the F on the (3-fluorophenyl)ethyl group of (R,R)-18 by an alkyne (Figure 5B). To our delight, the ethynyl moiety was well tolerated and probe 22 exhibited comparable inhibitory activity against SARS-CoV-2 Mpro (IC50 = 0.11 μM) to that of (R,R)-18 (IC50 = 0.056 μM). To test the capability as an ABP, 22 was reacted with recombinant SARS-CoV-2 Mpro (37 °C, 1 h) and subsequently conjugated to rhodamine-azide by copper-catalyzed azide-alkyne cycloaddition (CuAAC). The covalent 22–Mpro adduct was visualized by in-gel fluorescence analysis after SDS-PAGE (Figure 5B). The gel image showed that Mpro could be detected from as low as 25 nM by 2.5 μM of 22. Mpro (50 nM) was labeled by 22 in a concentration-dependent manner, and the fluorescent intensity reached the plateau at 2.5 μM of 22. 22 successfully labeled Mpro spiked into the lysate of A431 cells (derived from human lung cancer), demonstrating its capability for covalent modification of Mpro in the complex human proteome (Figure 5C and S7). We also prepared acrylamide probe 23 based on the structure of 1, and compared the proteome-wide activity of 22 and 23 under live cell conditions (Figure 5D and S8). A431 cells were treated with 22 or 23 (37 °C, 1 h) and subsequently processed for the visualization of the labeled proteome after cell lysis. Acrylamide 23 exhibited a higher degree of concentration-dependent indiscriminate protein labeling compared to CFA 22, in concordance to the higher intrinsic reactivity of acrylamide toward thiols than that of CFA. Overall, these data demonstrated that the CFA-dipeptides displayed good selectivity toward Mpro and low indiscriminate reactivity in the human cell proteome.
**Figure 5.** Selectivity of CFA-dipeptides toward M<sup>pro</sup>. A) Inhibitory activity of (R,R)-18 and Jun8-76-3A toward other viral and human cysteine proteases. The data are represented as the average of two independent experiments. B) Fluorescent labeling of recombinant SARS-CoV-2 M<sup>pro</sup> by CFA probe 22. M<sup>pro</sup> was reacted with the designated concentration of 22 (37 °C, 1 h) and analyzed by in-gel ABPP after CuAAC with rhodamine azide. C) Reactivity profile of 22 toward M<sup>pro</sup> spiked into A431 cell lysate (37 °C, 1 h). The samples of right two lanes were preincubated with the designated inhibitor (37 °C, 30 min) before addition of 22. Red arrow indicates SARS-CoV-2 M<sup>pro</sup>. D) Reactivity profiles of 22 and acrylamide probe 23 toward the live cell proteome of A431 cells (37 °C, 1 h).

**Conclusion**

SARS-CoV-2 M<sup>pro</sup> plays a vital role in the process of viral replication and is an attractive protein target for the development of antiviral agents to treat COVID-19. In this study, we developed a novel irreversible M<sup>pro</sup> inhibitor utilizing CFA as the warhead for the covalent capture of the catalytic cysteine. Ugi MCR afforded CFA derivatives in one-pot from chlorofluoroacetic acid, aniline, aldehyde, and isocyanide, allowing for rapid screening of the substituents derived from each building block. The diastereomers of CFA derivatives showed significantly different inhibitory activity against M<sup>pro</sup>, suggesting the importance of the chirality at CFA. We established a practical protocol for the optical resolution of chlorofluoroacetic acid, which enabled the isolation of the four stereoisomers of the optimized CFA compound 18 bypassing the use of expensive chiral HPLC. Among the stereoisomers, (R,R)-18 displayed significantly higher M<sup>pro</sup> inhibitory activity (IC<sub>50</sub> = 0.056 μM).
than those of the reported noncovalent and covalent inhibitors based on the same Ugi MCR-generated dipeptide scaffold. The kinetic analysis of \( M^{pro} \) inhibition revealed that the \( R \) configuration of CFA is crucial for the activity both in terms of binding affinity and the rate of covalent modification of \( M^{pro} \). These results indicated that the electrophilic reactivity of CFA could be enhanced under the microenvironment in the active site of \( M^{pro} \) in a stereospecific manner. \((R,R)-18\) showed high selectivity toward SARS-CoV-1 and SARS-CoV-2 \( M^{pro} \) among other cysteine proteases, and ABPP experiments using an alkynylated probe 22 demonstrated the low indiscriminate reactivity of CFA toward the human cell proteome. Overall, these data suggest the potential of \((R,R)-18\) as the lead for further development of CFA-based irreversible SARS-CoV-2 \( M^{pro} \). More importantly, our findings highlight the prominent influence of the CFA chirality on the covalent modification of the targeted cysteine in the enzyme binding pocket. This would provide an extra opportunity for improving the potency and selectivity in the development of novel CFA-based covalent inhibitors for various cysteine-containing protein targets.

The high electronegativity of fluorine makes C–F bonds highly polarized, and fluorine atom(s) in small molecule ligands can participate in noncovalent interactions with the protein in diverse ways. Assuming the same binding mode of CFA dipeptides in \( M^{pro} \) active site to the reported Ugi MCR-based inhibitors, we speculate that the fluorine atom of CFA makes hydrogen bonding(s) with the backbone NH of Gly143 and/or Cys145, which are thought to form the oxyanion hole in the S1’ pocket. This interaction may determine the orientation of the C–Cl bond of CFA in the binding pocket, rendering \( (R) \)-CFA more suitable for the \( S_N2 \) substitution reaction by the cysteine thiol compared to \((S)\)-CFA. Computational docking studies are ongoing to examine this possibility in more depth and will be reported in due course.

**Methods**

**Enzymatic Assays**

All cysteine proteases were purchased from BPS Bioscience (San Diego, CA): SARS-CoV-2 main protease \( (M^{pro}) \) (100823), SARS-CoV-1 \( M^{pro} \) (100807), SARS-CoV-2 papain-like protease \( (PL^{pro}) \) (100735), human cathepsin B (80001), and human cathepsin L (80005). All enzymatic assays were performed in 96-well microplates (PerkinElmer, 6052260). \( M^{pro} \) enzymatic assays were performed in the reaction buffer containing 20 mM Tris-HCl (pH 7.3), 100 mM NaCl, 1 mM EDTA, and 1 mM DTT. 25 nM SARS-CoV-2 \( M^{pro} \) or 200 nM SARS-CoV-1 \( M^{pro} \) (20 \( \mu \)L) and the testing compound at various concentrations (20 \( \mu \)L) were added to each well and incubated for 30 min at ambient temperature. The enzymatic reaction was initiated by adding the fluorogenic substrate (Ac-Abu-Tle-Leu-Gln-MCA, Peptide Institute 3250-v) (50 \( \mu \)M, 10 \( \mu \)L) and allowed to proceed for 6 h at ambient temperature. The fluorescence intensity of each well was measured using PerkinElmer EnSpire multimode plate reader (PerkinElmer) at an excitation wavelength of 355 nm and an emission wavelength of 460 nm. Percent activity or inhibition was calculated based on control wells containing no inhibitor (100% activity/0% inhibition) and blank wells containing no \( M^{pro} \) (0% activity/100% inhibition) and plotted versus inhibitor concentration. IC\(_{50}\) values were generated by curve fitting using KaleidaGraph (Synergy Software).

SARS-CoV-2 \( PL^{pro} \) enzymatic assays were performed in the same reaction buffer to \( M^{pro} \) enzymatic assays. 500
nM PLpro (20 μL) and the testing compound (125 μM, 20 μL) were added to each well and incubated for 30 min at ambient temperature. The enzymatic reaction was initiated by adding the substrate (Z-Arg-Leu-Arg-Gly-Gly-MCA, Sigma-Aldrich SML2966) (50 μM, 10 μL) and allowed to proceed for 6 h at ambient temperature. Percent inhibition was calculated as described above.

Cathepsin enzymic assays were performed in the reaction buffer containing 20 mM sodium acetate (pH 5.5), 1 mM EDTA, and 5 mM DTT. 750 pM cathepsin B or 750 pM cathepsin L (20 μL) and the testing compound (125 μM, 20 μL) were added to each well and incubated for 30 min at ambient temperature. The enzymatic reaction was initiated by adding the fluorogenic substrate (Z-Leu-Arg-MCA, Peptide Institute 3210-v) (25 μM, 10 μL) and allowed to proceed for 1 h at ambient temperature. Percent inhibition was calculated as described above.

**Kₘ determination of Ac-Abu-Tle-Leu-Gln-MCA**

Kₘ value of SARS-CoV-2 Mpro for the fluorogenic substrate Ac-Abu-Tle-Leu-Gln-MCA was determined at 12 different substrate concentrations ranging from 156.3 to 2.6 μM. 20 nM Mpro solution (25 μL) was added to each well of a 96-well plate. The enzymatic reaction was initiated by adding substrate solution (25 μL) and allowed to proceed for 20 min at 37 °C. The product formation was continuously monitored using PerkinElmer EnSpire multimode plate reader at an excitation wavelength of 355 nm and an emission wavelength of 460 nm. The rate of the product formation was plotted versus the substrate concentration, and Kₘ value was obtained by non-linear regression using the Michaelis–Menten equation (Figure S9). Kₘ value of 126.8 μM was obtained for Ac-Abu-Tle-Leu-Gln-MCA and used to determine the kinetic parameters for the irreversible inactivation of SARS-CoV-2 Mpro.

**Determination of the kinetic parameters for Mpro inactivation**

125 μM Ac-Abu-Tle-Leu-Gln-MCA (20 μL) and the inhibitor at various concentrations (20 μL) were added to each well of 96-well microplates. The enzymatic reaction was initiated by adding 50 nM SARS-CoV-2 Mpro (10 μL) and allowed to proceed for 60 min at 37 °C. The fluorescence intensity was continuously measured using PerkinElmer EnSpire multimode plate reader at an excitation wavelength of 355 nm and an emission wavelength of 460 nm. The raw fluorescence data were zero-corrected by subtracting the fluorescence value at 0 sec and converted to product concentration [P] using a calibration curve obtained by 7-amino-4-methylcoumarin. [P]was plotted versus time and the initial state velocity vᵢ at each inhibitor concentration was determined by linear regression of the initial three data points. The observed first-order rate constant kₐbs at each inhibitor concentration was determined by non-linear regression using equation (1). The second-order rate constant kₐact/Kᵢ was determined by plotting kₐbs versus [I] and non-linear regression using equation (2). The vᵢ values were plotted against [I]/(1 + [S]/Kₘ) and non-linear regression using equation (3) afforded Kᵢ value, where vᵢ is the rate in the absence of the inhibitor. The kₐact value was calculated from the kₐact/Kᵢ and Kᵢ values.
\[ [P] = \frac{v_i}{k_{obs}} \left(1 - e^{-k_{obs}t}\right) + \text{offset} \quad (1) \]

\[ k_{obs} = \frac{k_{\text{inact}} \times [I]}{[I] + K_i(1 + [S]/K_m)} \quad (2) \]

\[ v_i = \frac{v_0}{1 + [I]/K_i} \quad (3) \]

**SARS-CoV-2 M\textsuperscript{pro} labeling and in-gel fluorescence analysis**

Recombinant SARS-CoV-2 M\textsuperscript{pro} was diluted in the buffer containing 50 mM HEPES (pH7.5), 150 mM NaCl, and 1 mM DTT. DMSO solution of 22 was diluted in the same buffer, and 25 \( \mu \)L each of M\textsuperscript{pro} and 22 solution in buffer was combined to initiate the reaction. After incubated for 1 h at 37 °C, to the mixture was added 100 mM \( N \)-ethylmaleimide solution (1 \( \mu \)L) and further incubated for 20 min. The reaction mixture was subjected to copper-catalyzed azide–alkyne cycloaddition (CuAAC) reaction with 16.5 \( \mu \)M rhodamine azide, 0.8 mM TCEP (Sigma-Aldrich), 80 \( \mu \)M tris(3-hydroxypropyltriazolylmethyl)amine (THPTA, TCI) and 1.25 mM CuSO\(_4\). The mixture was incubated for 1 h at 37 °C. After addition of 15 \( \mu \)L 5×SDS-PAGE loading buffer, the mixture was boiled for 5 min at 95 °C and each sample (12 \( \mu \)L) was resolved on 5-15% SDS-PAGE gel (Nacalai tesque). The in-gel fluorescence was analyzed using LAS-4000 lumino image analyzer (Fujifilm).

**SARS-CoV-2 M\textsuperscript{pro} labeling in A431 cell lysate**

A431 cells were purchased from the American Type Culture Collection (ATCC, VA) and grown at 37 °C under a humidified 5% CO\(_2\) atmosphere in a culture medium containing low-glucose DMEM (Gibco) supplemented with 10% FBS (HyClone), penicillin (50 IU/mL) and streptomycin (50 \( \mu \)g/mL). The cells were grown to ~90% confluence in 10 mL of growth medium in 100-mm culture dishes (Corning Primaria). The medium was aspirated off, and the cells were washed twice with DPBS (10 mL). To culture dishes were added the buffer containing 50 mM HEPES (pH7.5), 150 mM NaCl, and 1 mM DTT (1 mL), and the cells were harvested using a plastic scraper in DPBS and collected by centrifugation (4 °C, 200g, 5 min), and the supernatant was removed. The pellets were resuspended in the ice-cold buffer (200 \( \mu \)L) and lysed using a probe sonicator. The suspension was centrifuged (4 °C, 17,730g, 10 min) and the supernatant was collected for use as lysate.

Protein concentration of the lysate was determined using DC protein assay kit (BioRad) and normalized to 0.4 mg/mL by dilution with the buffer. Different concentrations of SARS-CoV-2 M\textsuperscript{pro} was spiked into the lysate and treated with 22 and processed for in-gel fluorescence analysis as described above. For competition experiments, the lysate spiked with M\textsuperscript{pro} was preincubated with the designated inhibitor for 30 min at 37 °C before adding 22.
Protein labeling in live A431 cells

A431 cells were grown to ~80% confluence in 4 mL of growth medium in 60-mm culture dishes (Corning Primaria). The growth medium was aspirated off, and the cells were washed twice with DPBS (3 mL), followed by treatment with 22 in culture medium (4 mL, FBS(−)). After incubation at 37 °C in a CO₂ incubator for 1 h, the medium was aspirated off. The cells were washed twice with cold DPBS (3 mL) and lysed with cold RIPA buffer (50mM Tris-HCl pH 7.6, 150mM NaCl, 1w/v% Nonidet P40 substitute, 0.5w/v% sodium deoxycholate, 0.1w/v% SDS, 50 µL) containing protease inhibitor cocktail (Roche) and N-ethylmaleimide (2 mM). The lysed cells were collected using a plastic scraper, transferred to a microfuge tube, and centrifuged (4 °C, 17,730g, 10 min). The supernatant was transferred to a microfuge tube and stored at −30 °C.

The cell lysates were thawed on ice. Protein concentrations of each sample were determined using DC protein assay kit (BioRad) and normalized to 0.5 mg/mL (A431 cells) by dilution with DPBS. The solution (52 µL) was subjected to CuAAC, SDS-PAGE, and in-gel fluorescence analysis as described for M⁺Mo labeling.

Acknowledgement

This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas “Chemistry for Multimolecular Crowding Biosystems” (JSPS KAKENHI Grant No. JP17H06349) and 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 JP18am0101091. N.S. acknowledges Grant-in-Aid for Scientific Research B (JSPS KAKENHI Grant No. 19H02854) and AMED under Grant Number JP21ak0101121 for their financial supports. A.O. acknowledges AMED under Grant Number JP20fk0108519. We thank Dr. Taisuke Matsumoto (Institute for Materials Chemistry and Engineering, Kyushu University) for X-ray diffraction experiments.
References


