Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease

Chunlong Ma,[1] Michael D. Sacco,[2] Brett Hurst,[3,4] Julia A. Townsend,[5] Yanmei Hu,[1] Tommy Szeto,[1] Xiujun Zhang,[2] Bart Tarbet, [3,4] Michael T. Marty,[5] Yu Chen,[2,*] Jun Wang[1,*]

[1]Department of Pharmacology and Toxicology, College of Pharmacy, The University of Arizona, Tucson, AZ, 85721, United States

[2]Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, 33612, United States

[3]Institute for Antiviral Research, Utah State University, Logan, UT, 84322, USA

[4]Department of Animal, Dairy and Veterinary Sciences, Utah State University, Logan, UT, 84322, USA

[5]Department of Chemistry and Biochemistry, The University of Arizona, Tucson, AZ, 85721, United States

*Corresponding authors:

Jun Wang, Tel: 520-626-1366, Fax: 520-626-0749, email: junwang@pharmacy.arizona.edu

Yu Chen, Tel: 813-974-7809, email: ychen1@usf.edu

Keywords: SARS-CoV-2, COVID-19, main protease, 3CL protease, calpain inhibitors, boceprevir, GC-376

A novel coronavirus SARS-CoV-2, also called novel coronavirus 2019 (nCoV-19), started to circulate among humans around December 2019, and it is now widespread as a global pandemic. The disease caused by SARS-CoV-2 virus is called COVID-19, which is highly contagious and has an overall mortality rate of 6.96% as of May 4, 2020. There is no vaccine or antiviral available for SARS-CoV-2. In this study, we report our discovery of inhibitors targeting the SARS-CoV-2 main protease (M^{pro}). Using the FRET-based enzymatic assay, several inhibitors including boceprevir, GC-376, and calpain inhibitors II, and XII were identified to have potent activity with single-digit to submicromolar IC_{50} values in the enzymatic assay. The mechanism of action of the hits was further characterized using enzyme kinetic studies, thermal shift binding assays, and native mass spectrometry. Significantly, four compounds (boceprevir, GC-376, calpain inhibitors II and XII) inhibit SARS-CoV-2 viral replication in cell culture with EC_{50} values ranging from 0.49 to 3.37 µM. Notably, boceprevir, calpain inhibitors II and XII represent novel chemotypes that are distinct from known M^{pro} inhibitors. A complex crystal structure of SARS-CoV-2 M^{pro} with GC-376, determined at 2.15 Å resolution with three monomers per asymmetric unit, revealed two unique binding configurations, shedding light on the molecular interactions and protein conformational flexibility underlying substrate and inhibitor binding by M^{pro}. Overall, the compounds identified herein provide promising starting points for the further development of SARS-CoV-2 therapeutics.

INTRODUCTION

An emerging respiratory disease COVID-19 started to circulate among human in December 2019. Since its first outbreak in China from an unknown origin, it quickly became a global pandemic. As of May 4, 2020, there are 239,740 deaths among 3,442,234 confirmed cases in 215 countries.[1] The etiological pathogen of COVID-19 is a new coronavirus, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), also called novel coronavirus (nCoV-2019). As the name indicates, SARS-CoV-2 is similar to SARS, the virus that causes severe respiratory symptoms in human and killed 774 people among 8098 infected worldwide in 2003.[2] SARS-CoV-2 shares ~82% of sequence identity as SARS and to a less extent for Middle East respiratory syndrome (MERS) (~50%).[3, 4] SARS-CoV-2 is an enveloped, positive-sense, single-stranded RNA virus that belongs to the β-lineage of the coronavirus.[5] The β-lineage also contains two other important human pathogens, the SARS coronavirus and MERS coronavirus. The mortality rate of SARS-CoV-2 is around 6.96% as of May 4, 2020, which is lower than that of SARS (~10%) and MERS (~34%).[2] However, current data indicate that SARS-CoV-2 is more contagious and has a larger R0 value than SARS and MERS,[6] resulting in higher overall death tolls than SARS and MERS. The SARS-CoV-2 virus is currently spreading at an alarming speed in Europe and the United States.

There is currently no antiviral or vaccine for SARS-CoV-2. The SARS-CoV-2 viral genome encodes a number of structural proteins (e.g. capsid spike glycoprotein), non-structural proteins (e.g. 3-chymotrypsin-like protease (3CL or main protease), papain-like protease, helicase, and RNA-dependent RNA polymerase), and accessary proteins. Compounds that target anyone of these viral proteins might be potential antiviral drug candidates.[7, 8] In this study, we focus on the viral 3CL protease, also called the main protease (M^{pro}), and aim to develop potent M^{pro}

inhibitors as SAR-CoV-2 antivirals. The M^{pro} plays an essential role in coronavirus replication by digesting the viral polyproteins at more than 11 sites, and it appears like a high profile target for antiviral drug discovery.[9-12] The M^{pro} has an unique substrate preference for glutamine at the P1 site (Leu-Gln↓(Ser,Ala,Gly)), a feature that is absent in closely related host proteases, suggesting it is feasible to achieve high selectivity by targeting viral M^{pro}. As such, we developed the Fluorescence Resonance Energy Transfer (FRET)-based enzymatic assay for the SARS-CoV-2 M^{pro} and applied it to screen a focused library of protease inhibitors. Here we report our findings of several novel hits targeting SARS-CoV-2 M^{pro} and their mechanism of action. The in vitro antiviral activity of the hits was also evaluated in cell culture using infectious SARS-CoV-2 virus. Overall, our study provides a list of drug candidates for SARS-CoV-2 with a confirmed mechanism of action, and the results might help speed up the drug discovery efforts in combating COVID-19. The compounds identified herein represent one of the most potent and selective SARS-CoV-2 M^{pro} inhibitors so far with both enzymatic inhibition and cellular antiviral activity.[9, 11, 12] The X-ray crystal structure of SARS-CoV-2 M^{pro} with GC-376 showed that the compound can adapt two configurations R and S, offering a molecular explanation of the high-binding affinity of the aldehyde-containing inhibitors. Significantly, the discovery of calpain II and XII inhibitors as potent SARS-CoV-2 antivirals suggest that it might be feasible to design dual inhibitors against the viral M^{pro} and the host calpains, both of which are important for viral replication.

RESULTS

Establishing the FRET-based assay for the SARS-CoV-2 main protease (M^{pro})

The M^{pro} gene from SARS-CoV-2 strain BetaCoV/Wuhan/WIV04/2019 was inserted into pET-29a(+) vector and expressed in BL21(DE3) *E. Coli.* with a His-tag in its C-terminus. The M^{pro} protein was purified with Ni-NTA column to high purity (Fig. 1A). To establish the FRET assay condition, we designed a FRET based substrate using the sequence between viral polypeptide NSP4-NSP5 junction from SARS-CoV-2: Dabcyl-KTSAVLQ/SGFRKME(Edans). We then tested the M^{pro} proteolytic activity in buffers with different pH. We found that M^{pro} displays highest activity in pH 6.5 buffer (Fig. 1B), which contains 20 mM HEPES, 120 mM NaCl, 0.4 mM EDTA, and 4 mM DTT and 20% glycerol. As such, all the following proteolytic assay was conducted using this pH 6.5 buffer. Next, we characterized the enzymatic activity of this SARS-CoV-2 M^{pro} by measuring the K_m and V_{max} values. When 100 nM M^{pro} was mixed with various concentration of FRET substrate (0 to 200 μM), the initial velocity was measured and plotted against substrate concentration. Curve fitting with Michaelis-Menton equation gave the best-fit values of K_m and V_{max} as 32.8 ± 3.5 μM and 29.4 ± 1.1 RFU/s, respectively (Fig. 1C).

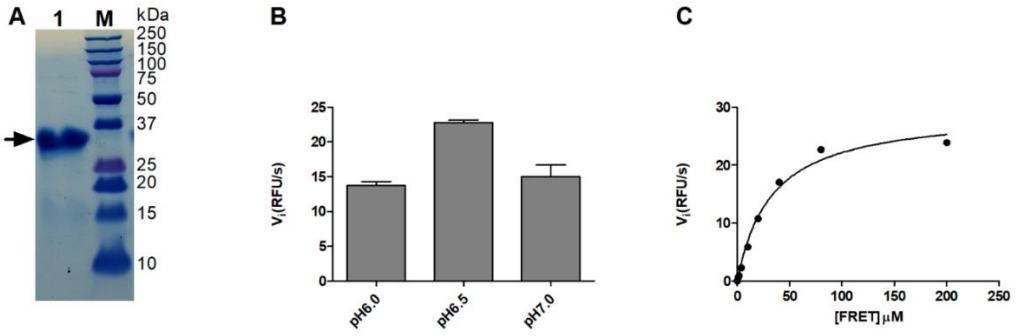


Figure 1: SARS-CoV-2 M^{pro} expression and characterization. (A) SDS-PAGE of His-tagged-Main protease (M^{pro}) (lane 1); Lane M, protein ladder; the calculated molecular weight of the His-tagged-M^{pro} is 34,992 Da. (B) Reaction buffer optimization: 250 nM His-tagged-M^{pro} was diluted into three reaction buffers with different pH values. (C) Michaelis-Menten plot of

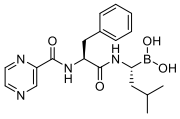
100 nM His-tagged- M^{pro} with the various concentrations of FRET substrate in pH 6.5 reaction buffer.

Primary screening of a focused protease library against the SARS-CoV-2 M^{pro}.

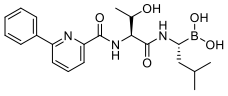
With the established FRET assay condition, we screened a collection of protease inhibitors from the Selleckchem bioactive compound library to identify potential SARS-CoV-2 M^{pro} inhibitors. The protease inhibitors are grouped based on their targets and mechanism of action and include proteasome inhibitors (**1-8**); HIV protease inhibitors (**9-14**); γ-secretase inhibitors (**15-22**); HCV NS3-4A protease inhibitors (**23-29**); DPP-4 inhibitors (**30-35**); miscellaneous serine protease inhibitors (**36-39**); cathepsin and calpain protease inhibitors (**40-43**); miscellaneous cysteine protease inhibitors (**44-48**); matrix metalloprotease inhibitors (**49-51**); and miscellaneous protease inhibitors (**52-55**). The inhibitors were pre-incubated with 100 nM of M^{pro} at 30 °C for 30 minutes in the presence of 4 mM 1,4-dithiothreitol (DTT) before the addition of 10 µM FRET substrate. The addition of DTT was to quench non-specific thiol reactive compounds and to ensure the M^{pro} is in the reducing condition. All compounds were tested at 20 µM, except compound **26**, which was tested at 2 µM due to its fluorescent background. Encouragingly, four inhibitors (**24**, **28**, **29** and **43**) showed more than 60% inhibition against M^{pro} at 20 µM. Among the hits, simeprevir (**24**), boceprevir (**28**), and narlaprevir (**29**) are HCV NS3-4A serine protease inhibitors, and compound MG-132 (**43**) inhibits both proteasome and calpain.

Table 1. List of protease inhibitors tested against SARS-CoV-2 M^{pro} in the primary FRET assay.

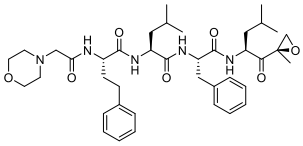
Proteosome inhibitors



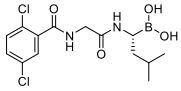
Bortezomib (PS-341) (**1**)



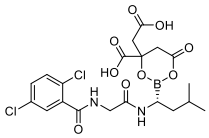
CEP-18770 (Delanzomib) (**2**)



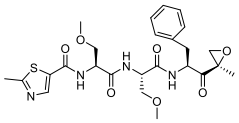
Carfilzomib (PR-171) (**3**)



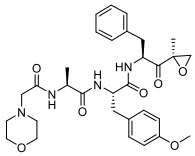
MLN2238 (**4**)



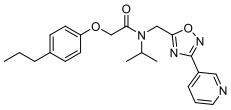
MLN9708 (**5**)



Oprozomib (ONX 0912) (**6**)

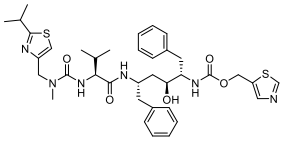


ONX-0914 (PR-957) (**7**)

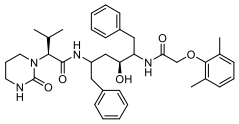


PI-1840 (**8**)

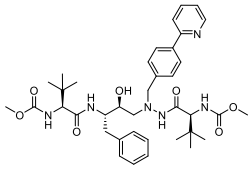
HIV protease (aspartic protease) inhibitors



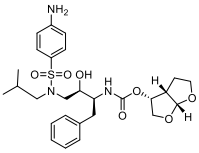
Ritonavir (**9**)



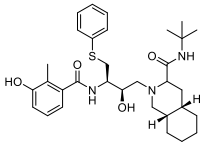
Lopinavir (**10**)



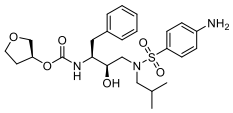
Atazanavir (**11**)



Darunavir (**12**)

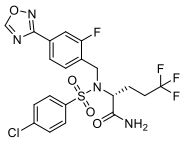


Nelfinavir (**13**)

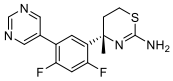


Amprenavir (**14**)

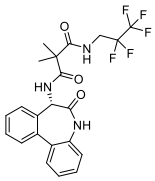
γ-secretase (aspartic protease) inhibitors



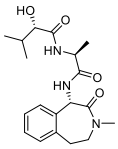
Avagacestat (**15**)



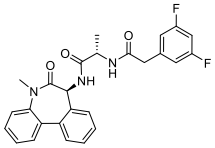
LY2811376 (**16**)



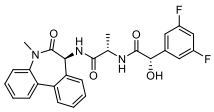
RO4929097 (**17**)



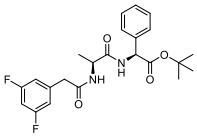
Semagacestat (LY450139) (**18**)



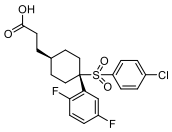
YO-01027 (**19**)



LY411575 (**20**)

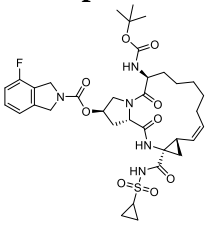


DAPT (GSI-IX) (**21**)

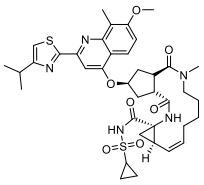


MK-0752 (**22**)

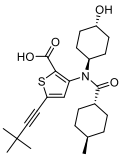
HCV protease (serine protease) inhibitors



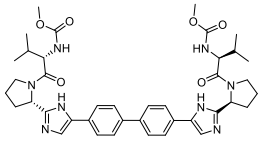
Danoprevir (**23**)



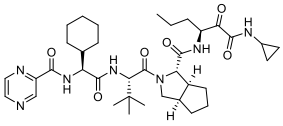
Simeprevir (**24**)



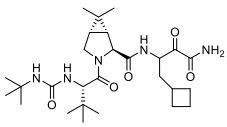
Lomibuvir (VX-222) (**25**)



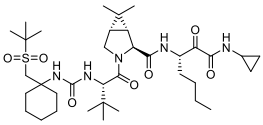
Daclatasvir (BMS-790052) (**26**)



Telaprevir (**27**)

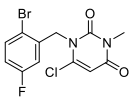


Boceprevir (**28**)

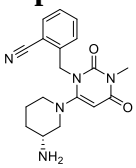


Narlaprevir (**29**)

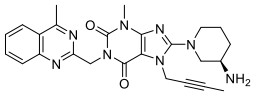
DPP-4 (serine protease) inhibitors



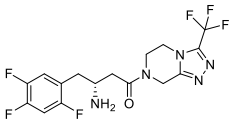
Trelagliptin (**30**)



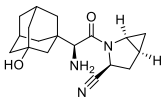
Alogliptin (**31**)



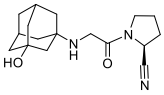
Linagliptin (**32**)



Sitagliptin (**33**)

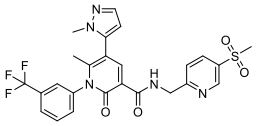


Saxagliptin (**34**)

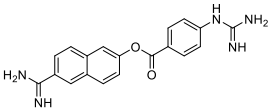


Vildagliptin (**35**)

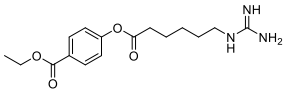
Miscellaneous serine protease inhibitors



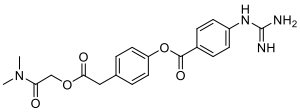
Alvelestat (**36**)



Nafamostat Mesylate (**37**)

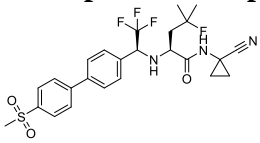


Gabexate (**38**)

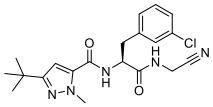


Camostat Mesilate (**39**)

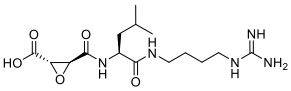
Cathepsin and calpain protease (cysteine protease) inhibitors



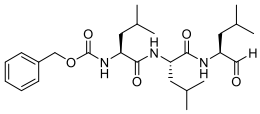
Odanacatib (MK-0822) (**40**)



Cathepsin Inhibitor 1 (**41**)

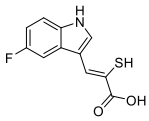


E-64 (**42**)

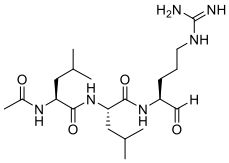


MG-132 (**43**)

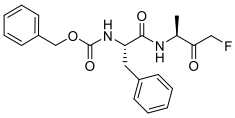
Miscellaneous cysteine protease inhibitors



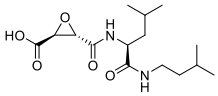
PD151746 (**44**)



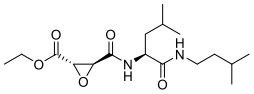
Leupeptin (**45**)



Z-FA-FMK (**46**)

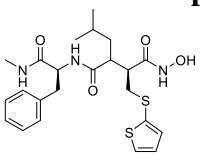


Loxistatin Acid (**47**)

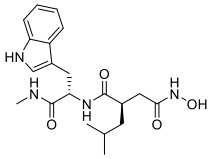


Aloxistatin (**48**)

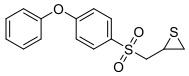
Matrix metallprotease inhibitors



Batimastat (BB-94) (**49**)

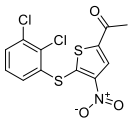


Ilomastat (**50**)

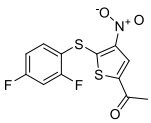


SB-3CT (**51**)

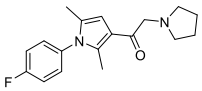
Miscellaneous protease inhibitors



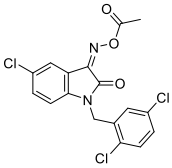
P5091 (P005091) (**52**)



P22077 (**53**)



IU1 S7134 (**54**)



LDN-57444 (**55**)

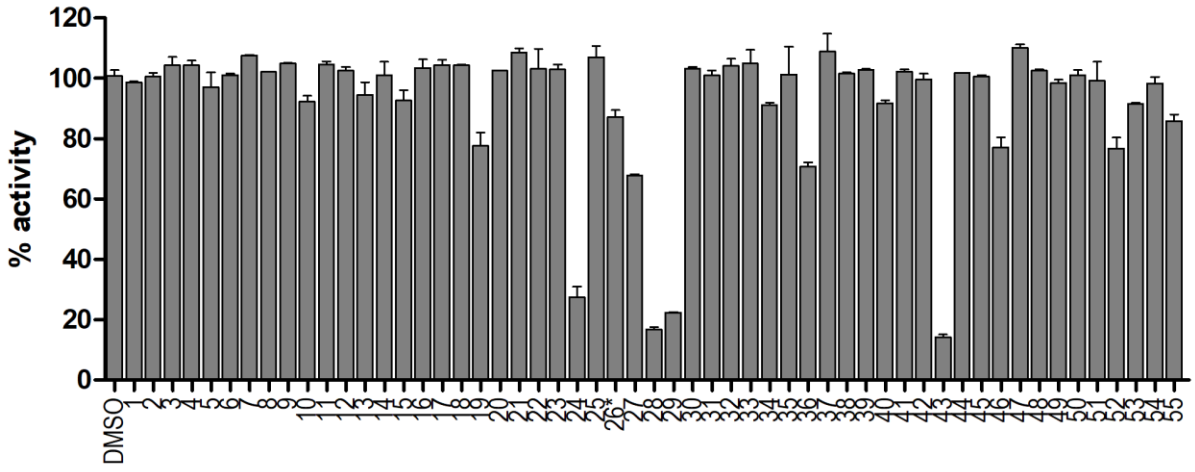


Figure 2: Screening of known protease inhibitors against SARS-CoV-2 M^{pro} using the FRET assay. 20 µM of compounds (**26** was tested at 2 µM) was pre-incubated with 100 nM of SARS-CoV-2 M^{pro} for 30 minutes at 30 °C, then 10 µM FRET substrate was added to reaction mixture to initiate the reaction. The reaction was monitored for 2 hours. The initial velocity was calculated by linear regression using the data points from the first 15 minutes of the reaction. The calculated initial velocity with each compound was normalized to DMSO control. The results are average ± standard deviation of two repeats.

Secondary screening of a focused library of calpain/cathepsin inhibitors and known viral $3CL^{pro}$ inhibitors

Given the encouraging results from the primary screening, we then further characterized the four hits (**24**, **28**, **29**, and **43**) in a consortium of assays including dose-response titration, thermal shift binding assay (TSA), and counter screening assays with two other viral cysteine proteases, the enterovirus A71 (EV-A71) 2A and 3C proteases, both of which are cysteine proteases (Table 2). The HCV NS3-4A protease inhibitors boceprevir (**28**) and narlaprevir (**29**) inhibited M^{pro} with IC_{50} values of 4.13 and 4.73 µM, respectively (Table 2), more potent than simeprevir (**24**) (IC_{50} =

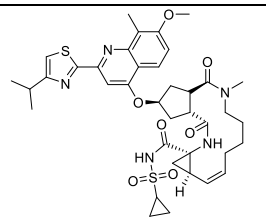
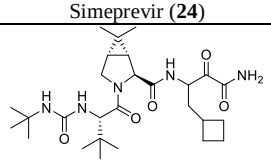
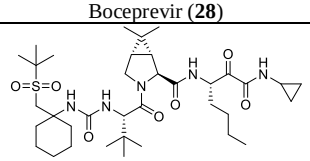
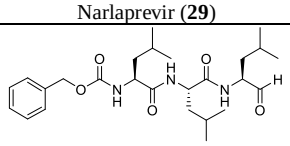
13.74 μM). Both compounds **28** and **29** also showed strong binding towards M^{pro} and shifted the melting temperature of the protein (ΔT_m) by 6.67 and 5.18 ^oC, respectively, at 40 μM. Despite their potent inhibition against the HCV NS3-4A serine protease and the SARS-CoV-2 cysteine M^{pro}, boceprevir (**28**) and narlaprevir (**29**) did not inhibit the EV-A71 2A and 3C proteases (IC_{50} > 20 μM), suggesting they are not non-specific cysteine protease inhibitors. The calpain inhibitor MG-132 (**43**) had an IC_{50} value of 3.90 μM against the M^{pro}, and was not active against the EV-A71 2A and 3C proteases (IC_{50} > 20 μM). The binding of MG-132 (**43**) to M^{pro} was also confirmed in the TSA assay with a ΔT_m of 4.02 ^oC.

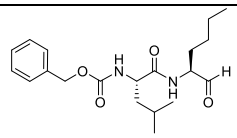
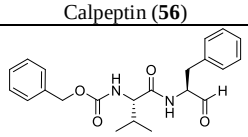
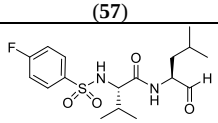
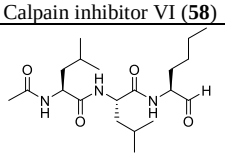
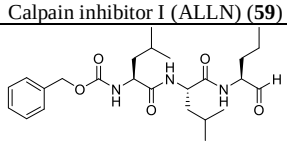
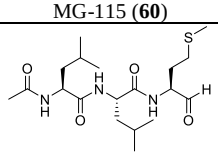
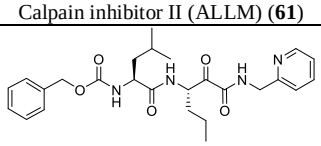
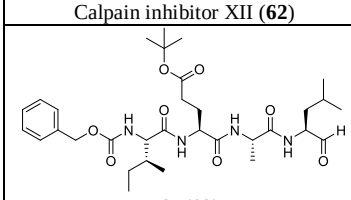
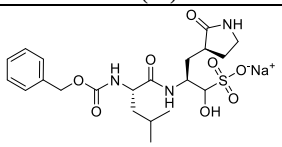
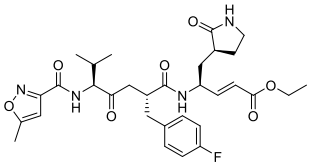
In light of the promising results of the calpain inhibitor MG-132 (**43**), we then pursued to testing other calpain and cathepsin inhibitors that are commercially available (**56-63**) (Table 2). These compounds were not included in the initial library because they have not been advanced to clinical studies. Among this series of analogs, calpain inhibitor II (**61**) and XII (**62**) are the most potent M^{pro} inhibitors with IC_{50} values of 0.97 and 0.45 μM, respectively. Binding of compounds **61** and **62** to M^{pro} shifted the melting curve of the protein by 6.65 and 7.86 ^oC, respectively. Encouragingly, both compounds **61** and **62** did not inhibit the EV-A71 2A and 3C proteases (IC_{50} > 20 μM). Calpain inhibitor I (**59**) and MG-115 (**60**) also showed potent inhibition against M^{pro} with IC_{50} values of 8.60 and 3.14 μM, respectively. Calpeptin (**56**) and PSI (**63**) had moderate activity against M^{pro} with IC_{50} values of 10.69 and 10.38 μM, respectively. In contrast, calpain inhibitors III (**57**) and VI (**58**) were not active (IC_{50} > 20 μM).

We also included two well-known viral 3CL protease inhibitors GC-376 (**64**) and rupintrivir (**65**) in the secondary screening. GC-376 (**64**) is an investigational veterinary drug that is being developed for feline infectious peritonitis (FIP).[13, 14] GC-376 (**64**) was designed to target the viral 3CL protease and had potent antiviral activity against multiple viruses including MERS, FIPV,

and norovirus.[13, 15] Rupintrivir (**65**) was developed as a rhinovirus antiviral by targeting the viral 3CL protease, but it was discontinued in clinical trials due to side effects.[16] In our study, we found that GC-376 (**64**) was the most potent M^{pro} inhibitor with an IC_{50} value of 0.03 µM. It shifted the melting curve of M^{pro} by 18.30 ^{o}C upon binding. In contrast, rupintrivir (**65**) was not active against M^{pro} (IC_{50} > 20 µM). Previous report also showed that rupintrivir was not active against the SARS-CoV $3CL^{pro}$ (M^{pro}) (IC_{50} > 100 µM).[17] Both compounds **64** and **65** were not active against the EV-A71 2A protease, but showed potent inhibition against the EV-A71 3C protease, which is consistent with previously reported results.[15, 18, 19]

Table 2: Characterization of HCV and calpain proteases inhibitors against SARS-CoV-2 M^{pro} using a consortium of secondary assays[a]

ID / Results	SARS-CoV-2 M^{pro} IC_{50} (µM)	SARS-CoV-2 M^{pro} TSA Tm/ΔTm (°C)	EV-A71 2A IC_{50} (µM)	EV-A71 3C IC_{50} (µM)	Development stage
DMSO	-------	55.74 ± 0.00	------	-------	
 Simeprevir (**24**)	13.74 ± 3.45	N.T.[b]	N.T.	N.T.	FDA-approved HCV drug
 Boceprevir (**28**)	4.13 ± 0.61	62.41 ± 0.21/6.67	>20	>20	FDA-approved HCV drug
 Narlaprevir (**29**)	5.73 ± 0.67	60.92 ± 0.14/5.18	>20	>20	FDA-approved HCV drug
 MG-132 (ApexBio) (**43**)	3.90 ± 1.01	59.76 ± 0.45/4.02	>20	>20	Preclinical; tested in mice[20]

 Calpeptin (**56**)	10.69 ± 2.77	56.84 ± 0.00/1.1	>20	>20	Preclinical; tested in mice and felines[21, 22]
 Calpain inhibitor III (MDL28170) (**57**)	>20	55.36 ± 0.14/-0.38	N.T.	N.T.	Preclinical; not tested in animal model
 Calpain inhibitor VI (**58**)	>20	55.46 ± 0.14/-0.28	N.T.	>20	Preclinical; tested in rats[23]
 Calpain inhibitor I (ALLN) (**59**)	8.60 ± 1.46	N.T.	>20	>20	Preclinical; tested in mice[24]
 MG-115 (**60**)	3.14 ± 0.97	60.51 ± 0.28/4.77	>20	>20	Preclinical; not tested in animal model
 Calpain inhibitor II (ALLM) (**61**)	0.97 ± 0.27	62.93 ± 0.14/6.65	>20	>20	Preclinical; not tested in animal model
 Calpain inhibitor XII (**62**)	0.45 ± 0.06	63.60 ± 0.01/7.86	>20	>20	Preclinical; not tested in animal model
 PSI (**63**)	10.38 ± 2.90[c]	N.T.	1.22	13.74 ± 3.86	Preclinical; tested in rats[25]
 GC376 (more reliable) (**64**)	0.030 ± 0.008	74.04 ± 0.07/18.30	>20	0.136 ± 0.025	Preclinical; tested in felines[13, 14]
 Rupintrivir (**65**)	> 20	N.T.	>20	0.042 ± 0.014	Dropped out of clinical trial

a: Value = mean ± S.E. from 3 independent experiments;

b: N.T. = not tested;

c: The IC_{50} of PSI (**64**) on SARS CoV-2 M^{pro} was calculated by end point reading of 1 hour digestion, instead of the initial velocity.

When plotting the IC_{50} values (log scale) of the inhibitors against M^{pro} from the FRET enzymatic assay with the melting temperature shifts (ΔT_m) from thermal shift binding assay (Fig. 3A), a linear correlation was observed, and the r^2 of the linear regression fitting is 0.94. This suggests that there is a direct correlation between the enzymatic inhibition and protein binding: a more potent enzyme inhibitor also binds to the protein with higher affinity. The stabilization of the M^{pro} against thermal denaturation was also compound concentration dependent (Fig. 3B).

The binding of four most potent inhibitors boceprevir (**28**), calpain inhibitors II (**61**), XII (**62**), and GC-376 (**64**) to SARS-CoV-2 M^{pro} was further characterized by native mass spectrometry (MS) (Figs. 3C-F). Native MS analysis showed that the M^{pro} formed a dimer complex with a mass of 69,722 Da, indicating a cleaved N-terminal methionine (Fig. S1). A small amount of monomer and dimer with a single C-terminal truncation of the His-tag was also observed, but the intact dimer was the predominant signal (Fig. S1). Addition of all four ligands tested, boceprevir (**28**), calpain inhibitors II (**61**), XII (**62**), and GC-376 (**64**), showed binding of up to two ligands per dimer (Fig. 3C-F), suggesting a binding stoichiometry of one drug per monomer.

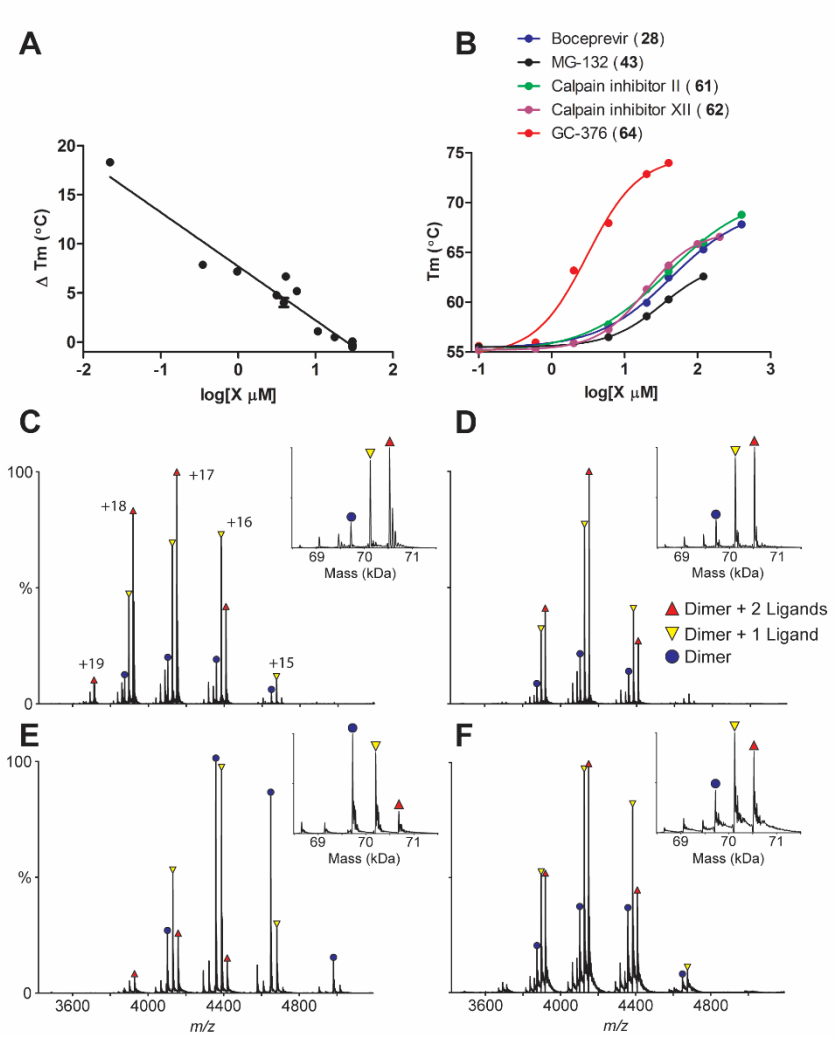


Figure 3: Binding of inhibitors to SARS-CoV-2 M^{pro} using thermal shift binding assay and native mass spectrometry. (A) Correlation of inhibition efficacy (IC_{50}) with ΔT_m from thermal shift binding assay. Data in Table 2 were used for the plot. The r^2 of fitting is 0.94. (B) Dose-dependent melting temperature (T_m) shift. Native MS reveals binding of SARS-CoV-2 M^{pro} to (C) GC-376 (**64**), (D) Calpain inhibitor II (**61**), (E) calpain inhibitor XII (**62**), and (F) Boceprevir (**28**). All ligand concentrations are 12.5 µM except E, which is 25 µM. Peak are annotated for dimer (*blue circle*), dimer with one bound ligand, (*yellow down triangle*), and dimer with two bound ligands (*red up triangle*). Other minor signals are truncated dimers, which bind ligands at

the same ratios. Charge states are annotated in C, and insets show the deconvolved zero-charge mass distribution.

Mechanism of action of hits

To elucidate the mechanism of action of hits against SARS-CoV-2 M^{pro}, we focus on five most potent compounds prioritized from the primary and secondary screenings including boceprevir (**28**), MG-132 (**43**), calpain inhibitor II (**61**), calpain inhibitor XII (**62**), and GC-376 (**64**). For this, we performed enzyme kinetic studies with different concentrations of inhibitors (Fig. 4). A biphasic enzymatic progression curve in the presence but not in the absence of inhibitor is typically a hallmark for a slow covalent binding inhibitor. In the Fig. 4, left column shows the progression curves up to 4 hours. Biphasic progression curves were observed for all 5 inhibitors at high drug concentrations. Significant substrate depletion was observed when the proteolytic reaction proceeded beyond 90 minutes, we therefore chose the first 90 minutes of the progression curves for curve fitting (Fig. 4 middle column). We fit the progression curves in the presence different concentrations of GC-376 (**64**) with the two-step Morrison equation (equation 3 in methods section). GC-376 (**64**) binds to SARS-CoV-2 M^{pro} with an equilibrium dissociation constant for the inhibitor (K_I) of 59.9 ± 21.7 nM in the first step. After initial binding, a slower covalent bond is formed between GC-376 (**64**) and M^{pro} with the second reaction rate constant (k_2) being 0.00245 ± 0.00047 s^{-1}, resulting an overall k_2/K_I value of 4.08 x 10^4 M^{-1} s^{-1} (Fig. 4A). However, when we tried to fit the proteolytic progression curves for boceprevir (**28**), MG-132 (**43**), calpain inhibitors II (**61**) and XII (**62**) using the same two-step reaction mechanism, we could not obtain accurate values for the second rate constant k_2. This is presumably due to significant substrate depletion before the equilibrium between EI and EI*, leading to very small values of k_2.

Accordingly, for these four inhibitors **28**, **43**, **61**, and **62**, only the dissociation constant K_I values from the first step were determined (Figs. 4B-4E). The inhibition constants (K_I) for boceprevir (**28**), MG-132 (**43**), calpain inhibitors II (**61**) and XII (**62**) are 1.18 ± 0.10 μM, 1.57 ± 0.13 μM, 0.40 ± 0.02 μM, and 0.13 ± 0.02 μM, respectively.

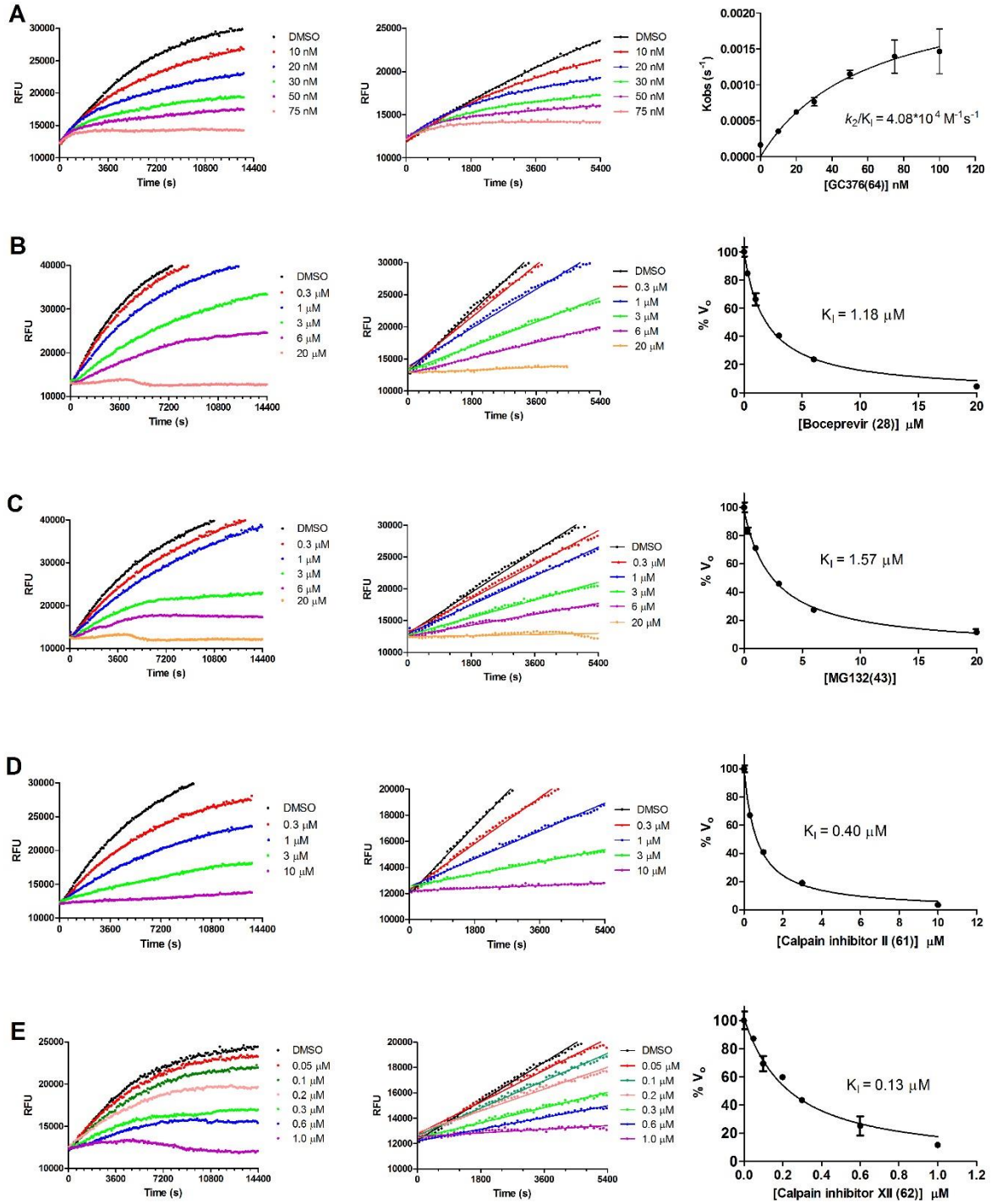


Figure 4: Proteolytic reaction progression curves of M^{pro} in the presence or the absence of compounds. In the kinetic studies, 5 nM M^{pro} was added to a solution containing various concentrations of protease inhibitors and 20 μM FRET substrate to initiate the reaction, the

reaction was then monitored for 4 hrs. Left column shows the reaction progression up to 4 hrs; middle column shows the progression curves for the first 90 minutes, which were used for curve fitting to generate the plot shown in the right column. Detailed methods were described in the Method section. (A) GC-376 (**64**); (B) Boceprevir (**28**); (C) MG-132 (**43**); (D) Calpian inhibitor II (**61**); (E) Calpain inhibitor XII (**62**).

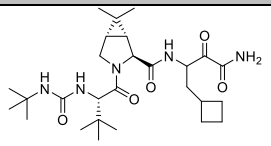
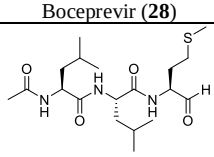
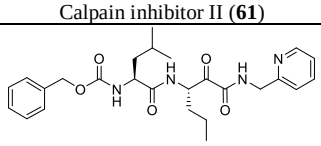
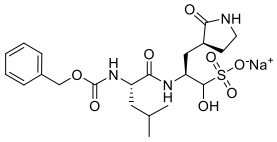
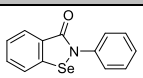
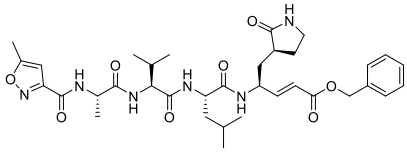
Cellular antiviral activity and cytotoxicity of hits

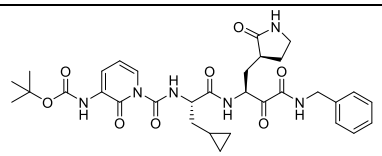
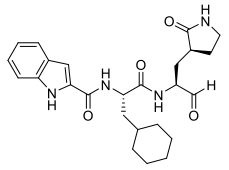
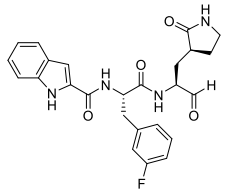
To test the hypothesis that inhibiting the enzymatic activity of M^{pro} will lead to the inhibition of SARS-CoV-2 viral replication, we performed cellular antiviral assays for the five promising hits **64**, **28**, **43**, **61**, and **62** against SARS-CoV-2. For this, we first tested the cellular cytotoxicity of these compounds in multiple cell lines (Table S1). GC-376 (**64**), boceprevir (**28**), and calpain inhibitor II (**61**) were well tolerated and had CC_{50} values of over 100 μM for all the cell lines tested. MG-132 (**43**) was cytotoxic to all the cells with CC_{50} values less than 1 μM except A549 cells. Calpain inhibitor XII (**62**) had acceptable cellular cytotoxicity with CC_{50} values above 50 μM for all the cell lines tested (Table S1).

Next, we chose four compounds boceprevir (**28**), calpain inhibitors II (**61**), XII (**62**), and GC-376 (**64**) for the antiviral assay with infectious SARS-CoV-2. MG-132 (**43**) was not included due to its cytotoxicity. Gratifyingly, all four compounds showed potent antiviral activity against SARS-CoV-2 in the primary viral cytopathic effect (CPE) assay with EC_{50} values ranging from 0.49 to 3.37 μM (Table 3). Their antiviral activity was further confirmed in the secondary viral yield reduction (VYR) assay. The most potent compound was calpain inhibitor XII (**62**), which showed EC_{50} of 0.49 μM in the primary CPE assay and EC_{90} of 0.45 μM in the secondary VYR assay. In comparison, remdesivir was reported to inhibit SARS-CoV-2 in the VYR assay with an

EC_{50} of 0.77 μM.[26] None of the compounds inhibited the unrelated influenza virus A/California/07/2009 (H1N1) virus (EC_{50} > 20 μM) (Table S1), suggesting the antiviral activity of the four compounds (boceprevir, calpain inhibitors II, XII, and GC-376) against SARS-CoV-2 is specific. In comparison with recently reported SARS-CoV-2 M^{pro} inhibitors (Table 3), the hits identified herein represent one of the most potent and selective drug candidates with broad chemical diversity.

Table 3: Antiviral activity of hits against SARS-CoV-2 and the comparison with recently reported M^{pro} inhibitors.

Compounds	SARS-CoV-2 M^{pro} inhibition (μM)	SARS-CoV-2 Antiviral activity (μM) Primary CPE assay[a]	SARS-CoV-2 Antiviral activity (μM) Secondary viral yield reduction assay[a]	Development Stage
SARS-Cov-2 M^{pro} inhibitors identified in this study				
 Boceprevir (**28**)	IC_{50} = 4.13 ± 0.61 K_I = 1.18	EC_{50} = 1.90 CC_{50} > 100 SI_{50} > 52.6	N.T.	FDA-approved HCV drug
 Calpain inhibitor II (**61**)	IC_{50} = 0.97 ± 0.27 K_I = 0.40	EC_{50} = 2.07 ± 0.76 CC_{50} > 100 SI_{50} > 48.3	EC_{90} = 2.40 ± 1.01	Preclinical; not tested in animal model
 Calpain inhibitor XII (**62**)	IC_{50} = 0.45 ± 0.06 K_I = 0.13	EC_{50} = 0.49 ± 0.18 CC_{50} > 100 SI_{50} > 204	EC_{90} = 0.45 ± 0.17	Preclinical; not tested in animal model
 GC-376 (**64**)	IC_{50} = 0.030 ± 0.008 k_2/K_I = 40, 800 $M^{-1}S^{-1}$	EC_{50} = 3.37 ± 1.68 CC_{50} > 100 SI_{50} > 29.7	EC_{90} = 2.13 ± 1.05	Preclinical; tested in felines[13, 14]
Recently reported SARS-Cov-2 M^{pro} inhibitors[b]				
 **Ebselen**[11]	IC_{50} = 0.67 ± 0.09	4.67 ± 0.80	N. A.	In clinical trials
 **N3**[11]	k_{obs}/[I] = 11,300 ± 800 $M^{-1}S^{-1}$	16.77 ± 1.70	N. A.	Preclinical;[27] not tested in animal model

 **13b**[9]	IC_{50} = 0.67 ± 0.18	4 ~ 5	N. A.	Preclinical;[9] not tested in animal model
 **11a**[12]	IC_{50} = 0.053 ± 0.005	0.53 ± 0.01	N. A.	Preclinical;[12] favorable PK in rats and low toxicity in rats and dogs
 **11b**[12]	IC_{50} = 0.040 ± 0.002	0.72 ± 0.09	N. A.	Preclinical;[12] favorable PK in rats

[a]CPE EC_{50}, VYR EC_{90}, and cytotoxicity CC_{50} values are mean ± S.D. of 3 independent experiments. [b]Results were retrieved from recent publications.[9, 11, 12] N.T. = not tested. N. A. = not available

Complex crystal structure of SARS-CoV-2 M^{pro} with GC-376 (64)

The crystal structure of the SARS-CoV-2 M^{pro} in complex with GC-376 (**64**) was solved in the $p3_221$ space group at 2.15 Å resolution (Table S2). There are three monomers per asymmetric unit (ASU), with two constituting a biological dimer and the third forming a dimer with a crystallographic symmetry related neighboring protomer (Fig. S2). The presence of three monomers in our crystal structure allowed us to capture different binding configurations of GC-376 (**64**) (Fig. 5), a unique feature that was not observed in previous X-ray crystal structures [9, 11, 12]. The pairwise r.m.s.d among the monomer backbone Cα atoms ranges from 0.435 Å to 0.564 Å. Previously, SARS-CoV M^{pro} and SARS-CoV-2 M^{pro} crystal structures have been solved most frequently as a monomer per ASU, and occasionally a dimer [9, 11, 12, 28, 29]. In its native state, M^{pro} requires dimerization to become catalytically active [30, 31], which is supported by our native MS

data (Fig. S1). In our crystal structures, all three protomers appear catalytically competent, with the third protomer activated by the N-finger from an adjacent asymmetric unit (Fig. S2B).

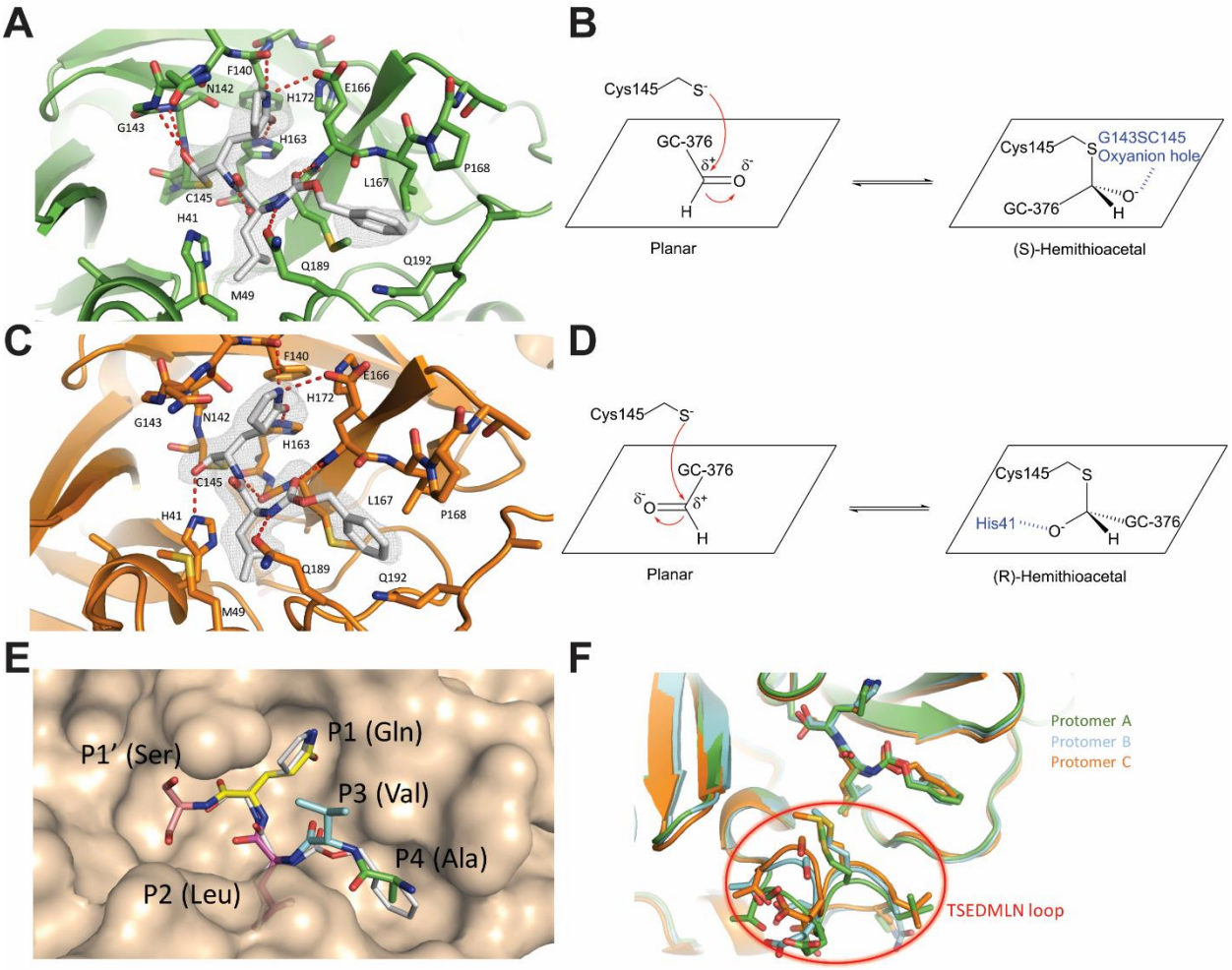


Figure 5: Molecular recognition of GC-376 (64) by SARS-CoV-2 M^{pro}. Complex of SARS-CoV-2 M^{pro} and GC-376 (**64**) with (A, B) protomer A and (C, D) protomer C. Unbiased F_o-F_c map, shown in grey, is contoured at 2 σ. Hydrogen bonds are shown as red dashed lines. (E) Surface representation of SARS-CoV-2 M^{pro} + GC-376 (**64**) (white) superimposed with the SARS-CoV M^{pro} natural, N-terminal substrate (PDB ID 2Q6G, with residues P1'-P4 in different colors). The SARS-CoV-2 M^{pro} cleaves between the P1' and P1 residues. (F) Superimposition of

the three protomers in the asymmetric subunit of SARS-CoV-2 M^{pro} with GC-376 (**64**). Significant conformational flexibility is observed, particularly in the TSEDMLN loop.

GC-376 (**64**) forms an extensive network of hydrogen bonds with the active site while also exhibiting excellent geometric complementarity (Fig. 5). These interactions are coupled with the thermodynamic payoff of covalent adduct formation between the aldehyde bisulfite warhead and Cys145, making GC-376 (**64**) one of the most potent SARS-CoV-2 M^{pro} inhibitors *in vitro* with an IC_{50} of 0.030 ± 0.008 µM (Table 3). Along with other known M^{pro} inhibitors **N3**, **13b**, **11a** and **11b** (Table 3), GC-376 (**64**) mimics the peptide substrate that is cleaved by this enzyme (Fig. 5E) [9, 27, 29, 32]. The glutamine surrogate γ-lactam ring is a cyclized derivative of the P1 glutamine side chain that normally occupies the S1 site; here it forms hydrogen bonds with the His163 and Glu166 side chains and the main chain of Phe140 (Figs. 5A & C). An amide bond connects the γ-lactam side chain to an isobutyl moiety that embeds itself in the hydrophobic S2 site formed by His41, Met49, and Met169. Normally, this S2 site in SARS-CoV-2 M^{pro} can accommodate a variety of hydrophobic substitutions such as isobutyl in GC-376 (**64**) and **N3**, cyclopropyl in **13b**, cyclohexyl in **11a**, and 3-fluorophenyl in **11b** (Table 3) [33, 34]. A carbamate bond in GC-376 (**64**), which forms hydrogen bonds with the main chain of Glu166 and the side chain of Gln189, connects the P2 isobutyl group to a phenylmethyl ester that interacts with the aliphatic S4 site. Compared with previous inhibitors, the phenylmethyl ester of GC-376 exhibits high complementarity with the S4 site, and the extensive non-polar interactions may contribute significantly to the potency of this compound (Fig. 5E).

Three copies of GC-376 (**64**) were found in the crystal structure, one in each protomer active site (Figs. 5A, 5C, S3). The configurations of G-C376 (**64**) were consistent in protomers A and B, where the thioacetal hydroxide is positioned in the "oxyanion hole" formed by the backbone amides of Gly143, Ser144, and Cys145 (Figs. 5A & S3), resulting in the (S)-configuration. It is noted that aldehydes **11a** and **11b** also bind in the active site of SARS-CoV-2 M^{pro} in the (S)-configuration (PDB: 6M0K and 6LZE) (Figs. S4A).[12] In protomer C, however, the same hydroxide group orients outwards from the oxyanion hole, forming hydrogen bonds with His41 (Fig. 5C), which gives the (R)-configuration. This (R)-configuration is consistent with the binding mode of α-ketoamide **13b** in the active site of SARS-CoV-2 M^{pro} (PDB: 6Y2F) (Fig. S4B).[9] These two unique configurations R and S might be a result of the Cys145 thiol nucleophilic attacking the aldehyde from two different faces (Figs. 5B & D). The fact that GC-376 (**64**) can adapt two different configurations R and S upon binding to the active site might explain its high binding affinity towards the target.

An additional difference between the configurations of GC-376 (**64**) in A, B and C is observed in the orientation of the phenylmethyl ester. In protomer C, the CH_2 of the phenylmethyl points towards the main chain of Leu167 in a 'cis' conformation (Fig. 5C), whereas in protomers A and B this same CH_2 points downwards in a 'trans' conformation (Figs. 5A & S3). Consequently, this influences the rotameric configuration of the Leu167 isobutyl moiety, where a rotational adjustment of 180˚ occurs at its "β" carbon. Furthermore, large rearrangements are observed in the flexible TSEDMLN loop consisting of residues 45-51 (TAEDMLN in SARS-CoV-2 M^{pro}) that form the S2 and S3' subsites (Fig. 5F), explaining the broad substrate scope in the P2 site (Table 3). The loop conformations in protomers B and C may be influenced by crystal packing interactions with protomers from adjacent asymmetric units and

resemble the conformations in previously determined structures [11, 12]. Meanwhile, the conformation of protomer A is less restrained and exhibits the most significant conformational divergence. The different loop conformations offer a glimpse of the protein plasticity that allows M^{pro} to accommodate peptides with differing amino acid composition, and underscores the importance of considering this flexibility when analyzing and modeling protein-ligand interactions for M^{pro} .

DISCUSSION

Coronaviruses have caused three epidemics/pandemics in the past twenty years including SARS, MERS, and COVID-19. With the ongoing pandemic of COVID-19, scientists and researchers around the globe are racing to find effective vaccines and antiviral drugs.[7] The viral polymerase inhibitor remdesivir holds the greatest promise and it is currently being evaluated in several clinical trials.[35, 36] The HIV drug combination lopinavir and ritonavir recently failed in a clinical trial for COVID-19 with no significant therapeutic efficacy was observed.[37] To address this unmet medical need, we initiated a drug repurposing screening to identify potent inhibitors against the SARS-CoV-2 M^{pro} from a collection of FDA-approved protease inhibitors. The M^{pro} has been shown to be a validated antiviral drug target for SARS and MERS.[38] As the SARS-CoV-2 M^{pro} shares a high sequence similarity with SARS and to a less extent with MERS, we reasoned that inhibiting the enzymatic activity of SARS-CoV-2 M^{pro} will similarly prevent viral replication.[9, 11]

Noticeable findings from our study include: 1) Boceprevir (**28**), an FDA-approved HCV drug, inhibits the enzymatic activity of M^{pro} with IC_{50} of 4.13 μM, and has an EC_{50} of 1.90 μM against the SARS-CoV-2 virus in the cellular viral cytopathic effect assay. The therapeutic

potential of boceprevir (**28**) should be further evaluated in relevant animal models and human clinic trials. Since boceprevir (**28**) is a FDA-approved drug, the dose, toxicity, formulation, and pharmacokinetic properties are already known, which will greatly speed up the design of follow up studies; 2) GC-376 (**64**), an investigational veterinary drug, showed promising antiviral activity against the SARS-CoV-2 virus (EC_{50} = 3.37 μM). It has the highest enzymatic inhibition against the M^{pro} with an IC_{50} value of 0.03 μM. This compound has promising in vivo efficacy in treating cats infected with FIP, and has favorable in vivo pharmacokinetic properties. Therefore, GC-376 (**64**) is ready to be tested in relevant animal models of SARS-CoV-2 when available. Importantly, the X-ray crystal structure of SARS-CoV-2 M^{pro} in complex with GC-376 (**64**) provides a molecular explanation of the high binding affinity of aldehyde-containing compounds as they can adapt two configurations R and S. The conformational flexibility at the TSEDMLN loop explains the broad substrate scope at the P2 position of M^{pro} inhibitors; 3) Three calpain/cathepsin inhibitors, MG-132 (**43**), calpain inhibitors II (**61**) and XII (**62**), are potent inhibitors of M^{pro} and inhibit SARS-CoV-2 with single-digit to submicromolar efficacy in the enzymatic assay. Calpain inhibitors II (**61**) and XII (**62**) also inhibit SARS-CoV-2 in the CPE assay with EC_{50} values of 2.07 and 0.49 μM, respectively. This result suggests that calpain/cathepsin inhibitors are rich sources of drug candidates for SARS-CoV-2. Indeed, previous studies have shown that calpain and cathepsin are required for the proteolytic processing of the coronavirus S protein, a step that is essential for the viral fusion and genome release during the early stage of viral replication.[39] Calpain and cathepsin inhibitors such as MDL28170 (calpain inhibitor III)[39], MG-132[40], calpain inhibitor VI[41] have been shown to inhibit SARS-CoV replication in cell culture. Other than the increased potency of targeting both M^{pro} and calpain/cathepsin, an additional benefit of such dual inhibitors might be their high genetic

barrier to drug resistance. A significant number of calpain/cathepsin inhibitors have been developed over the years for various diseases including cancer, neurodegeneration disease, kidney diseases, and ischemia/reperfusion injury.[42] Given our promising results of calpain inhibitors II (**61**) and XII (**62**) in inhibiting the SARS-CoV-2 M^{pro} and their potent antiviral activity in cell culture, it might be worthwhile to repurposing them as antivirals for SARS-CoV-2.

All potent SARS-CoV-2 M^{pro} inhibitors contain reactive warheads such as α-ketoamide (boceprevir (**28**), calpain inhibitor XII (**62**)) or aldehyde (MG-132 (**43**), calpain inhibitor II (**61**)) or aldehyde prodrug, the bisulfite (GC-376 (**64**)). This result suggests that reactive warheads might be essential for SARS-CoV-2 M^{pro} inhibition. The compounds identified in this study represent one of the most potent and selective hits reported so far, and are superior than recently reported SARS-CoV-2 M^{pro} inhibitors **ebselen**, **N3**, and **13b** (Table 3). Calpain inhibitor XII (**62**) had similar potency as the recently disclosed compounds **11a** and **11b** (Table 3).[12] Notably, calpain inhibitor II (**61**) and XII (**62**) have different chemical scaffolds as GC-376 (**64**), **N3**, **13b**, **11a**, and **11b**, therefore providing new opportunities for designing more potent and selective SARS-CoV-2 M^{pro} inhibitors.

Aside from the above positive results, we also showed that ritonavir (**9**) and lopinavir (**10**) failed to inhibit the SARS-CoV-2 M^{pro} (IC_{50} > 20 μM, Fig. 2), which might explain their lack efficacy in clinical trials for COVID-19.[37] Camostat (**39**) was recently proposed to inhibit SARS-CoV-2 entry through inhibiting the host TMPRSS2, a host serine protease that is important for viral S protein priming.[43] However, the antiviral activity of camostat has not been confirmed with infectious SARS-CoV-2 virus. In our study, we found camostat (**39**) has no inhibition against the SARS-CoV-2 M^{pro} (IC_{50} > 20 μM).

In summary, this study identified several potent SARS-CoV-2 M^{pro} inhibitors with potent enzymatic inhibition as well as cellular antiviral activity. Further development based on these hits might lead to clinically useful COVID-19 antivirals. They can be used either alone or in combination with polymerase inhibitors such as remdesivir as a means to achieve potential synergic antiviral effect as well as to suppress drug resistance.

MATERIALS AND METHODS

Details of materials and methods can be found in the supporting information.

DATA AVAILABILITY. The structure for SARS-CoV-2 Mpro has been deposited in the Protein Data Bank with accession number 6WTT.

ACKNOWLEDGEMENTS

This research was partially supported by the National Institutes of Health (NIH) (Grant AI147325) and the Arizona Biomedical Research Centre Young Investigator grant (ADHS18-198859) to J. W. B. H. and B. T. thanks for the support from the Respiratory Diseases Branch, National Institute of Allergy and Infectious Diseases, NIH, USA (Contract N01-AI-30048). J. A. T. and M. T. M. were funded by the National Institute of General Medical Sciences and National Institutes of Health (Grant R35 GM128624 to M. T. M.). We thank Michael Kemp for assistance with crystallization and X-ray diffraction data collection. We also thank the staff members of the Advanced Photon Source of Argonne National Laboratory, particularly those at the Structural Biology Center (SBC), with X-ray diffraction data collection. SBC-CAT is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

AUTHOR CONTRIBUTIONS

J. W. and C. M. conceived and designed the study; C. M. expressed the M^{pro} with the assistance of T. S.; C.M. performed the primary screening, secondary IC_{50} determination, thermal shift-binding assay, and enzymatic kinetic studies; M. S. carried out M^{pro} crystallization and structure determination with the assistance of X. Z, and analyzed the data with Y. C.; B. H. and B. T. performed the SARS-CoV-2 CPE and VYR assay; J .A. T. performed the native mass spectrometry experiments with the guidance from M. T. M.; Y. H. performed the plaque reduction assay with influenza A/California/07/2009 (H1N1) virus; J. W. and Y. C. secured funding and supervised the study; J. W., Y.C., and C. M. wrote the manuscript with the input from others.

ADDITIONAL INFORMATION

Supplementary information accompanies this paper at

Competing interests

J. W. and C. M. are inventors of a pending patent that claims the use of the identified compounds for COVID-19.

REFERENCES

1 https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

2 Mahase E. Coronavirus: covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate. *BMJ* 2020; **368**:m641.

3 Lu R, Zhao X, Li J *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 2020; **395**:565-574.

4 Wu A, Peng Y, Huang B *et al.* Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe* 2020; **27**:325-328.

5 Coronaviridae Study Group of the International Committee on Taxonomy of V. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* 2020; **5**:536-544.

6 Tang B, Bragazzi NL, Li Q, Tang S, Xiao Y, Wu J. An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov). *Infect Dis Model* 2020; **5**:248-255.

7 Li G, De Clercq E. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). *Nat Rev Drug Discov* 2020; **19**:149-150.

8 Xia S, Liu MQ, Wang C *et al.* Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res* 2020; **30**:343-355.

9 Zhang L, Lin D, Sun X *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* 2020; **368**:409-412.

10 Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* 2003; **300**:1763-1767.

11 Jin Z, Du X, Xu Y *et al.* Structure of M(pro) from COVID-19 virus and discovery of its inhibitors. *Nature* 2020, asap.

12 Dai W, Zhang B, Su H *et al.* Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* 2020, asap.

13 Pedersen NC, Kim Y, Liu H *et al.* Efficacy of a 3C-like protease inhibitor in treating various forms of acquired feline infectious peritonitis. *J Feline Med Surg* 2018; **20**:378-392.

14 Kim Y, Liu H, Galasiti Kankanamalage AC *et al.* Reversal of the Progression of Fatal Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus Protease Inhibitor. *PLoS Pathog* 2016; **12**:e1005531.

15 Kim Y, Lovell S, Tiew KC *et al.* Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. *J Virol* 2012; **86**:11754-11762.

16 Hayden FG, Turner RB, Gwaltney JM *et al.* Phase II, randomized, double-blind, placebo-controlled studies of ruprintrivir nasal spray 2-percent suspension for prevention and treatment of experimentally induced rhinovirus colds in healthy volunteers. *Antimicrob Agents Chemother* 2003; **47**:3907-3916.

17 Shie JJ, Fang JM, Kuo TH *et al.* Inhibition of the severe acute respiratory syndrome 3CL protease by peptidomimetic alpha,beta-unsaturated esters. *Bioorg Med Chem* 2005; **13**:5240-5252.

18 Musharrafieh R, Ma C, Zhang J *et al.* Validating Enterovirus D68-2A(pro) as an Antiviral Drug Target and the Discovery of Telaprevir as a Potent D68-2A(pro) Inhibitor. *J Virol* 2019; **93**, e02221-18.

19 Kuo CJ, Shie JJ, Fang JM *et al.* Design, synthesis, and evaluation of 3C protease inhibitors as anti-enterovirus 71 agents. *Bioorg Med Chem* 2008; **16**:7388-7398.

20 Bonuccelli G, Sotgia F, Schubert W *et al.* Proteasome inhibitor (MG-132) treatment of mdx mice rescues the expression and membrane localization of dystrophin and dystrophin-associated proteins. *Am J Pathol* 2003; **163**:1663-1675.

21 Mani SK, Shiraishi H, Balasubramanian S *et al.* In vivo administration of calpeptin attenuates calpain activation and cardiomyocyte loss in pressure-overloaded feline myocardium. *Am J Physiol Heart Circ Physiol* 2008; **295**:H314-326.

22 Peng S, Kuang Z, Zhang Y, Xu H, Cheng Q. The protective effects and potential mechanism of Calpain inhibitor Calpeptin against focal cerebral ischemia-reperfusion injury in rats. *Mol Biol Rep* 2011; **38**:905-912.

23 Akdemir O, Ucankale M, Karaoglan A *et al.* Therapeutic efficacy of SJA6017, a calpain inhibitor, in rat spinal cord injury. *J Clin Neurosci* 2008; **15**:1130-1136.

24 Li SZ, Zhang HH, Zhang JN *et al.* ALLN hinders HCT116 tumor growth through Bax-dependent apoptosis. *Biochem Biophys Res Commun* 2013; **437**:325-330.

25 Delli Pizzi S, Rossi C, Di Matteo V *et al.* Morphological and metabolic changes in the nigro-striatal pathway of synthetic proteasome inhibitor (PSI)-treated rats: a MRI and MRS study. *PLoS One* 2013; **8**:e56501.

26 Wang M, Cao R, Zhang L *et al.* Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res* 2020; **30**:269-271.

27 Xue X, Yu H, Yang H *et al.* Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J Virol* 2008; **82**:2515-2527.

28 Yang H, Yang M, Ding Y *et al.* The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci U S A* 2003; **100**:13190-13195.

29 Muramatsu T, Takemoto C, Kim Y-T *et al.* SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proc Natl Acad Sci U S A* 2016; **113**:12997-13002.

30 Li C, Qi Y, Teng X *et al.* Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like proteinase. *J Biol Chem* 2010; **285**:28134-28140.

31 Anand K, Palm GJ, Mesters JR, Siddell SG, Ziebuhr J, Hilgenfeld R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J* 2002; **21**:3213-3224.

32 Zhang L, Lin D, Kusov Y *et al.* alpha-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J Med Chem* 2020, asap.

33 Chuck CP, Chong LT, Chen C, Chow HF, Wan DC, Wong KB. Profiling of substrate specificity of SARS-CoV 3CL. *PloS one* 2010; **5**:e13197.

34 Hegyi A, Ziebuhr J. Conservation of substrate specificities among coronavirus main proteases. *J Gen Virol* 2002; **83**:595-599.

35 Grein J, Ohmagari N, Shin D *et al.* Compassionate Use of Remdesivir for Patients with Severe Covid-19. *N Eng J Med* 2020, asap.

36 Holshue ML, DeBolt C, Lindquist S *et al.* First Case of 2019 Novel Coronavirus in the United States. *N Engl J Med* 2020; **382**:929-936.

37 Cao B, Wang Y, Wen D *et al.* A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19. *N Eng J Med* 2020, asap.

38 Zumla A, Chan JF, Azhar EI, Hui DS, Yuen KY. Coronaviruses - drug discovery and therapeutic options. *Nat Rev Drug Discov* 2016; **15**:327-347.

39 Simmons G, Gosalia DN, Rennekamp AJ, Reeves JD, Diamond SL, Bates P. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc Natl Acad Sci U S A* 2005; **102**:11876-11881.

40 Schneider M, Ackermann K, Stuart M *et al.* Severe acute respiratory syndrome coronavirus replication is severely impaired by MG132 due to proteasome-independent inhibition of M-calpain. *J Virol* 2012; **86**:10112-10122.

41 Barnard DL, Hubbard VD, Burton J *et al.* Inhibition of severe acute respiratory syndrome-associated coronavirus (SARSCoV) by calpain inhibitors and beta-D-N4-hydroxycytidine. *Antivir Chem Chemother* 2004; **15**:15-22.

42 Ono Y, Saido TC, Sorimachi H. Calpain research for drug discovery: challenges and potential. *Nat Rev Drug Discov* 2016; **15**:854-876.

43 Hoffmann M, Kleine-Weber H, Schroeder S *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020; **181**:271-280 e278.

Supporting information

Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease

Chunlong Ma,[1] Michael D. Sacco,[2] Brett Hurst,[3,4] Julia A. Townsend,[5] Yanmei Hu,[1] Tommy Szeto,[1] Xiujun Zhang,[2] Bart Tarbet, [3,4] Michael T. Marty,[5] Yu Chen,[2,*] Jun Wang[1,*]

[1]Department of Pharmacology and Toxicology, College of Pharmacy, The University of Arizona, Tucson, AZ, 85721, United States

[2]Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, 33612, United States

[3]Institute for Antiviral Research, Utah State University, Logan, UT, 84322, USA

[4]Department of Animal, Dairy and Veterinary Sciences, Utah State University, Logan, UT, 84322, USA

[5]Department of Chemistry and Biochemistry, The University of Arizona, Tucson, AZ, 85721, United States

*Corresponding author:

Jun Wang, Tel: 520-626-1366, Fax: 520-626-0749, email: junwang@pharmacy.arizona.edu

MATERIALS AND METHODS

Cell lines and viruses. Human rhabdomyosarcoma (RD); A549, MDCK, Caco-2, and Vero cells were maintained in Dulbecco's modified Eagle's medium (DMEM), BEAS2B and HCT-8 cells were maintained in RPMI 1640 medium. Both medium was supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin antibiotics. Cells were kept at 37°C in a 5% CO2 atmosphere. The USA_WA1/2020 strain of SARS-CoV-2 obtained from the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA).

Protein expression and purification. SARS CoV-2 main protease (M^{pro} or 3CL) gene from strain BetaCoV/Wuhan/WIV04/2019 was ordered from GenScript (Piscataway, NJ) in the pET29a(+) vector with E. coli codon optimization. pET29a(+) plasmids with SARS CoV-2 main protease was transformed into competent E. coli BL21(DE3) cells, and a single colony was picked and used to inoculate 10 ml of LB supplemented with 50 g/ml kanamycin at 37°C and 250 rpm. The 10-ml inoculum was added to 1 liter of LB with 50 g/ml kanamycin and grown to an optical density at 600 nm of 0.8, then induced using 1.0 mM IPTG. Induced cultures were incubated at 37 °C for an additional 3 h and then harvested, resuspended in lysis buffer (25 mM Tris [pH 7.5], 750 mM NaCl, 2 mM dithiothreitol [DTT] with 0.5 mg/ml lysozyme, 0.5 mM phenylmethylsulfonyl fluoride [PMSF], 0.02 mg/ml DNase I), and lysed with alternating sonication and French press cycles. The cell debris were removed by centrifugation at 12,000 g for 45 min (20% amplitude, 1 s on/1 s off). The supernatant was incubated with Ni-NTA resin for over 2 h at 4°C on a rotator. The Ni-NTA resin was thoroughly washed with 30 mM imidazole in wash buffer (50 mM Tris [pH 7.0], 150 mM NaCl, 2 mM DTT); and eluted with 100 mM imidazole in 50 mM Tris [pH 7.0], 150 mM NaCl, 2 mM DTT. The imidazole was removed via dialysis or on a 10,000-molecular-weight-

cutoff centrifugal concentrator spin column. The purity of the protein was confirmed with SDS-PAGE. The protein concentration was determined via 260nM absorbance with ε 32890. EV-A71 2Apro and 3Cpro were expressed in the pET28b(+) vector as previously described (*1-3*).

Peptide synthesis. The SARS-CoV-2 M^{pro} FRET substrate Dabcyl-KTSAVLQ/SGFRKME(Edans) was synthesized by solid-phase synthesis through iterative cycles of coupling and deprotection using the previously optimized procedure.(*4*) Specifically, chemmatrix rink-amide resin was used. Typical coupling condition was 5 equiv of amino acid, 5 equiv of HATU, and 10 equiv of DIEA in DMF for 5 minutes at 80 ^{o}C. For deprotection, 5% piperazine plus 0.1 M HOBt were used and the mixture was heated at 80°C for 5 minutes. The peptide was cleaved from the resin using 95% TFA, 2.5% Tris, 2.5% H_2O and the crude peptide was precipitated from ether after removal of TFA. The final peptide was purified by preparative HPLC. The purify and identify of the peptide were confirmed by analytical HPLC (> 98% purity) and mass spectrometry. $[M+3]^{3+}$ calculated 694.15, detected 694.90; $[M+4]^{4+}$ calculated 520.86, detected 521.35;

Native Mass Spectrometry. Prior to analysis, the protein was buffer exchanged into 0.2 M ammonium acetate (pH 6.8) and diluted to 10 μM. DTT was dissolved in water and prepared at a 400 mM stock. Each ligand was dissolved in ethanol and diluted to 10X stock concentrations. The final mixture was prepared by adding 4 μL protein, 0.5 μL DTT stock, and 0.5 μL ligand stock for final concentration of 4 mM DTT and 8 μM protein. Final ligand concentrations were used as annotated. The mixtures were then incubated for 10 minutes at room temperature prior to analysis. Each sample was mixed and analyzed in triplicate.

Native mass spectrometry (MS) was performed using a Q-Exactive HF quadrupole-Orbitrap mass spectrometer with the Ultra-High Mass Range research modifications (Thermo Fisher

Scientific). Samples were ionized using nano-electrospray ionization in positive ion mode using 1.0 kV capillary voltage at a 150 °C capillary temperature. The samples were all analyzed with a 1,000–25,000 m/z range, the resolution set to 30,000, and a trapping gas pressure set to 3. Between 10 and 50 V of source fragmentation was applied to all samples to aid in desolvation. Data were deconvolved and analyzed with UniDec.(*5*)

Enzymatic assays. For reaction condition optimization, 200 µM SARS CoV-2 Main protease was used. pH6.0 buffer contains 20 mM MES pH6.0, 120 mM NaCl, 0.4 mM EDTA, 4 mM DTT and 20% glycerol; pH6.5 buffer contains 20 mM HEPES pH6.5, 120 mM NaCl, 0.4 mM EDTA, 4 mM DTT and 20% glycerol, pH7.0 buffer contains 20 mM HEPES pH7.0, 120 mM NaCl, 0.4 mM EDTA, 4 mM DTT and 20% glycerol. Upon addition of 20 µM FRET substrate, the reaction progress was monitored for 1 hr. The first 15 min of reaction was used to calculate initial velocity (V_i) via linear regression in prism 5. Main protease displays highest proteolytic activity in pH6.5 buffer. All the following enzymatic assays were carried in pH6.5 buffer.

For the measurements of K_m/V_{max}, screening the protease inhibitor library, as well as IC_{50} measurements, proteolytic reaction with 100 nM Main protease in 100 µl pH6.5 reaction buffer was carried out at 30 °C in a Cytation 5 imaging reader (Thermo Fisher Scientific) with filters for excitation at 360/40 nm and emission at 460/40 nm. Reactions were monitored every 90 s. For K_m/V_{max} measurements, a FRET substrate concentration ranging from 0 to 200 µM was applied. The initial velocity of the proteolytic activity was calculated by linear regression for the first 15 min of the kinetic progress curves. The initial velocity was plotted against the FRET concentration with the classic Michaelis-Menten equation in Prism 5 software. For the screening protease inhibitor library and IC_{50} measurements, 100 nM Main protease was incubated with protease inhibitor at 30°C for 30 min in reaction buffer, and then the reaction was initiated by adding 10

μM FRET substrate, the reaction was monitored for 1 h, and the initial velocity was calculated for the first 15 min by linear regression. The IC_{50} was calculated by plotting the initial velocity against various concentrations of protease inhibitors by use of a dose-response curve in Prism 5 software. Proteolytic reaction progress curve kinetics measurements with GC376, MG132, Boceprevir, Calpain inhibitor II, and Calpain inhibitor XII used for curve fitting, were carried out as follows: 5 nM Main protease protein was added to 20 μM FRET substrate with various concentrations of testing inhibitor in 200 μl of reaction buffer at 30 °C to initiate the proteolytic reaction. The reaction was monitored for 4 hrs. The progress curves were fit to a slow binding Morrison equation (equation 3) as described previously (*1*, *6*):

$$E + I \underset{k_{-1}}{\overset{k_1}{\Leftrightarrow}} EI \underset{k_{-2}}{\overset{k_2}{\Leftrightarrow}} EI* \qquad (1)$$

$$K_I = k_{-1}/k_1 \qquad (2)$$

$$P(t) = P_0 + V_s t - (Vs - V_0)(1 - e^{-kt})/k \qquad (3)$$

$$k = k_2[I]/(K_I + [I]) \qquad (4)$$

where P(t) is the fluorescence signal at time t, P_0 is the background signal at time zero, V_0, V_s, and and k represent, respectively, the initial velocity, the final steady-state velocity and the apparent first-order rate constant for the establishment of the equilibrium between EI and EI* (*6*). k_2/K_I is commonly used to evaluate the efficacy for covalent inhibitor. We observed substrate depletion when proteolytic reactions progress longer than 90 min, therefore only first 90 min of the progress curves were used in the curve fitting (Figure 6 middle column). In this study, we could not accurately determine the k_2 for the protease inhibitors: Calpain inhibitor II, MG132, Boceprevir, and Calpain inhibitor XII, due to the very slow k_2 in these case: significant substrate depletion before the establishment of the equilibrium between EI and EI*. In these

cases, K_I was determined with Morrison equation in Prism 5.

Differential scanning fluorimetry (DSF). The binding of protease inhibitors on Main protease protein was monitored by differential scanning fluorimetry (DSF) using a Thermal Fisher QuantStudio™ 5 Real-Time PCR System. TSA plates were prepared by mixing Main protease protein (final concentration of 3 μM) with inhibitors, and incubated at 30 °C for 30 min. 1× SYPRO orange (Thermal Fisher) were added and the fluorescence of the plates were taken under a temperature gradient ranging from 20 to 90 °C (incremental steps of 0.05 °C/s). The melting temperature (T_m) was calculated as the mid-log of the transition phase from the native to the denatured protein using a Boltzmann model (Protein Thermal Shift Software v1.3). Thermal shift which was represented as ΔT_m was calculated by subtracting reference melting temperature of proteins in DMSO from the T_m in the presence of compound.

Cytotoxicity measurement. A549, MDCK, HCT-8, Caco-2, Vero, and BEAS2B cells for cytotoxicity CPE assays were seeded and grown overnight at 37 °C in a 5% CO2 atmosphere to ~90% confluence on the next day. Cells were washed with PBS buffer and 200 μl DMEM with 2% FBS and 1% penicillin−streptomycin, and various concentration of protease inhibitors was added to each well. 48 hrs after addition the protease inhibitors, cells were stained with 66 μg/ mL neutral red for 2 h, and neutral red uptake was measured at an absorbance at 540 nm using a Multiskan FC microplate photometer (Thermo Fisher Scientific). The CC_{50} values were calculated from best-fit dose−response curves using GraphPad Prism 5 software.

SARS-CoV-2 CPE assay. Antiviral activities of test compounds were determined in nearly confluent cultures of Vero 76 cells. The assays were performed in 96-well Corning microplates. Cells were infected with approximately 60 cell culture infectious doses ($CCID_{50}$) of SARS-CoV-2 and 50% effective concentrations (EC_{50}) were calculated based on virus-induced cytopathic

effects (CPE) quantified by neutral red dye uptake after 5 days of incubation. Three microwells at each concentration of compound were infected. Two uninfected microwells served as toxicity controls. Cells were stained for viability for 2 h with neutral red (0.11% final concentration). Excess dye was rinsed from the cells with phosphate-buffered saline (PBS). The absorbed dye was eluted from the cells with 0.1 ml of 50% Sörensen's citrate buffer (pH 4.2)-50% ethanol. Plates were read for optical density determination at 540 nm. Readings were converted to the percentage of the results for the uninfected control using an Excel spreadsheet developed for this purpose. EC_{50} values were determined by plotting percent CPE versus $\log_{10}$ inhibitor concentration. Toxicity at each concentration was determined in uninfected wells in the same microplates by measuring dye uptake.

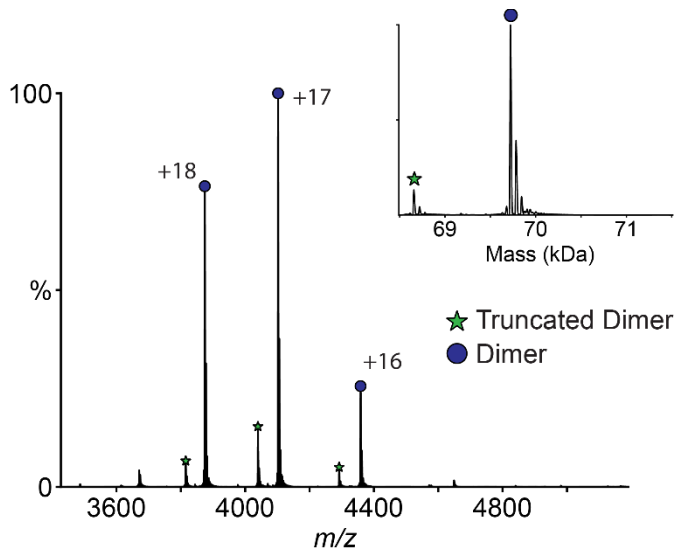
SARS-CoV-2 VYR assay. Virus yield reduction (VYR) assays were conducted by first replicating the viruses in the presence of test compound. Supernatant was harvested 3 days post-infection from each concentration of test compound and the virus yield was determined by endpoint dilution method. Briefly, supernatant virus was serially diluted in $\log_{10}$ increments then plated onto quadruplicate wells of 96-well plates seeded with Vero 76 cells. The presence or absence of CPE for determining a viral endpoint was evaluated by microscopic examination of cells 6 days after infection. From these data, 90% virus inhibitory concentrations (EC_{90}) were determined by regression analysis.

Influenza A virus A/California/07/2009 (H1N1) plaque reduction assay. The plaque assay was performed according to previously published procedures.(*7*)

M^{pro} crystallization and structure determination. 10 mg / mL of SARS-CoV-2 Mpro was incubated with 2 mM GC376 at 4˚ C O/N. The protein was diluted to 2.5 mg / mL the following

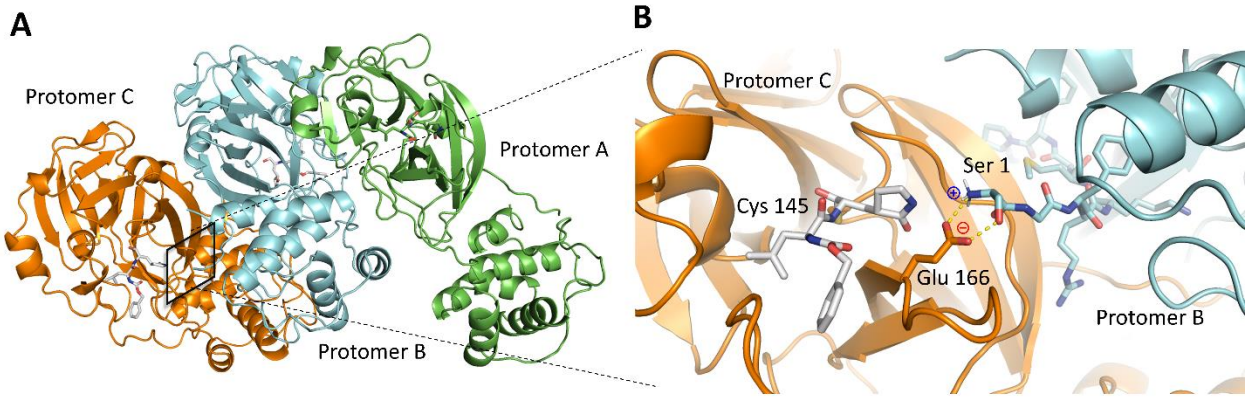
day in protein buffer (50 mM Tris pH 7.0, 150 mM NaCl, 4 mM DTT). Since GC3760 is water soluble, no precipitation was observed, and centrifugation was not necessary. Crystals were grown by mixing 2 uL of the protein solution with 1 ul of the precipitant solution (15 % PEG 2K, 10% 1,6-hexanediol, and 0.2 M NaCl) in a hanging-drop vapor-diffusion apparatus. Crystals were cryoprotected by transferring to a cryoprotectant solution (20% PEG 2K, 10% 1,6-hexanediol, 20% glycerol) and flash-frozen in liquid nitrogen.

X-ray diffraction data for the SARS-CoV2-M^{pro} GC376 complex structure was collected on the SBC 19-ID beamline at the Advanced Photon Source (APS) in Argonne, IL, and processed with the HKL3000 software suite(*8*). The CCP4 versions of MOLREP were used for molecular replacement using a previously solved SARS-CoV-2 M^{pro} (PDB ID 5RGG) as a reference model(*9*). Rigid and restrained refinements were performed using REFMAC and model building with COOT(*10, 11*). Protein structure figures were made using PyMOL (Schrödinger, LLC).



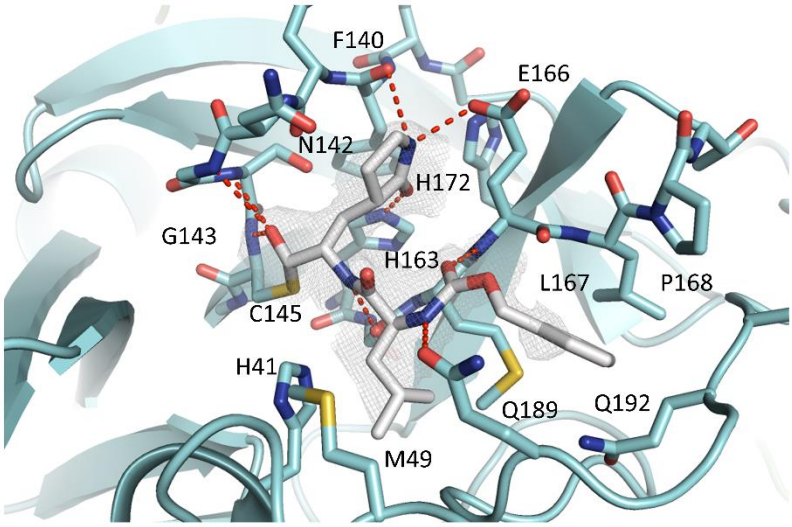
Supplementary Figure 1. Native mass spectrum of SARS-CoV-2 M^{pro}.

Native mass spectrum of SARS-CoV-2 M^{pro} with 4 mM DTT shows a dimer (*blue circle*) with a small amount of truncated dimer where one subunit has lost the C-terminal His tag (*green star*). The primary charge states are labeled, and the inset shows the deconvolved zero-charge mass distribution.

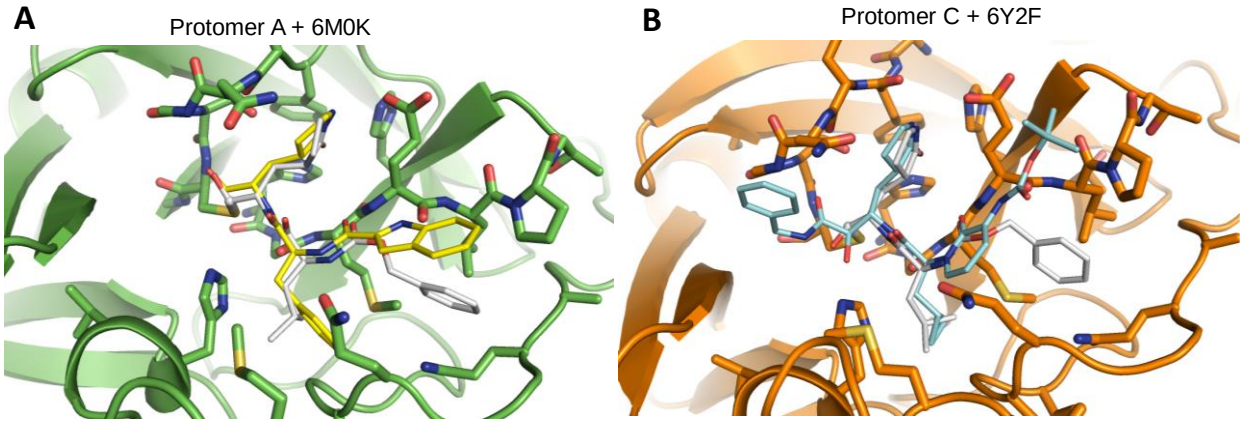


Supplementary Figure 2. Overall structure of SARS-CoV-2 M^{pro}.

(A) The three protomers in the asymmetric unit. Protomers B and C form a biological dimer. Protomer A dimerizes with a protomer from an adjacent asymmetric unit (not depicted). (B) The N-finger, or the N-terminal eight residues interact with Glu166 of the adjacent protomer, an important feature for catalytic activity.



Supplementary Figure 3. Complex structure of SARS-CoV-2 M^{pro} protomer B with GC-376 (64). Unbiased F_o-F_c map, shown in grey, is contoured at 2 σ. Hydrogen bonds are shown as red dashed lines.



Supplementary Figure 4. Overlay structures of current X-ray crystal structure with previously solved structures. (A) Overlay structures of protomer A with compound **13b** at the active site (PDB: 6M0K). (B) Overlay structures of protomer C with compound **N3** at the active site (PDB: 6Y2F). Unbiased F_o-F_c map, shown in grey, is contoured at 2 σ. Hydrogen bonds are shown as red dashed lines.

Supplementary Table 1: Cytotoxicity of SARS-CoV-2 M^{pro} inhibitors on various cell lines[a] and the counter screening against influenza virus.

	GC-376 (**64**)	Boceprevir (**28**)	MG-132 (**43**)	Calpain inhibitor II (**61**)	Calpain inhibitor XII (**62**)
MDCK	>100	>100	0.34 ± 0.02	>100	60.36 ± 2.28
Vero	>100	>100	0.45 ± 0.02	>100	>100
HCT-8	>100	>100	0.47 ± 0.02	>100	73.29 ± 11.80
A549	>100	>100	10.71 ± 3.50	>100	>100
Caco-2	>100	>100	<0.15	>100	82.02 ± 0.37
BEAS2B	>100	>100	0.14 ± 0.03	>100	78.91 ± 13.70
A/California/07/2009 (H1N1) antiviral activity[b] (μM)	> 20	> 20	N.T.	> 20	> 20

[a]Cytotoxicity was evaluated by measuring CC_{50} values (50% cytotoxic concentration) with CPE assay described in the method section. CC_{50} = mean ± S.E. of 3 independent experiments. [b]Antiviral activity against influenza virus was tested in plaque assay.

Supplementary Table 2: Table of Crystallization Statistics

Data Collection	**PDB ID 6WTT**
Structure	SARS-CoV-2 M^{pro} + GC-376
Space Group	P $3_2$21
Cell Dimension	
a, b, c (Å)	101.83, 101.83, 160.02
α, β, γ (°)	90.00, 90.00, 120.00
Resolution (Å)	50.00 - 2.15
	(2.19 - 2.15)
R_{merge} (%)	0.107 (0.885)
<I>/σ<I>	5.9 (2.16)
Completeness (%)	100 (99.7)
Redundancy	9.12 (7.1)
Refinement	
Resolution (Å)	45.64 - 2.15
	(2.27 - 2.15)
No. reflections/free	52836 / 2711
R_{work}/R_{free}	0.227 /0.299
No. Heavy Atoms	7429
Protein	6968
Ligand/Ion	92
Water	369
B-Factors ($Å^2$)	
Protein	35.60
Ligand/Ion	33.07
Solvent	32.69
RMS Deviations	
Bond Lengths (Å)	0.015
Bond Angles (°)	1.84
Ramachandran Favored (%)	94.22
Ramachandran Allowed (%)	5.78
Ramachandran Outliers (%)	0.00
Rotameric Outliers (%)	2.08

***** Numbers in parentheses represent the highest resolution shell.

References:

1. R. Musharrafieh *et al.*, Validating Enterovirus D68-2A(pro) as an Antiviral Drug Target and the Discovery of Telaprevir as a Potent D68-2A(pro) Inhibitor. *Journal of virology* **93**, (2019).
2. L. Shang *et al.*, Biochemical characterization of recombinant Enterovirus 71 3C protease with fluorogenic model peptide substrates and development of a biochemical assay. *Antimicrob Agents Chemother* **59**, 1827-1836 (2015).
3. Q. Cai *et al.*, Conformational plasticity of the 2A proteinase from enterovirus 71. *Journal of virology* **87**, 7348-7356 (2013).
4. S. D. Cady, J. Wang, Y. Wu, W. F. DeGrado, M. Hong, Specific binding of adamantane drugs and direction of their polar amines in the pore of the influenza M2 transmembrane domain in lipid bilayers and dodecylphosphocholine micelles determined by NMR spectroscopy. *Journal of the American Chemical Society* **133**, 4274-4284 (2011).
5. M. T. Marty *et al.*, Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Analytical chemistry* **87**, 4370-4376 (2015).
6. J. F. Morrison, C. T. Walsh, The behavior and significance of slow-binding enzyme inhibitors. *Adv Enzymol Relat Areas Mol Biol* **61**, 201-301 (1988).
7. C. Ma, J. Zhang, J. Wang, Pharmacological Characterization of the Spectrum of Antiviral Activity and Genetic Barrier to Drug Resistance of M2-S31N Channel Blockers. *Mol Pharmacol* **90**, 188-198 (2016).
8. W. Minor, M. Cymborowski, Z. Otwinowski, M. J. A. C. S. D. B. C. Chruszcz, HKL-3000: the integration of data reduction and structure solution–from diffraction images to an initial model in minutes. **62**, 859-866 (2006).

9. A. Vagin, A. Teplyakov, Molecular replacement with MOLREP. *Acta crystallographica. Section D, Biological crystallography* **66**, 22-25 (2010).

10. G. N. Murshudov *et al.*, REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography* **67**, 355-367 (2011).

11. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126-2132 (2004).