ARTICLE

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

Chunlong Ma [1], Michael Dominic Sacco[2], Brett Hurst[3,4], Julia Alma Townsend[5], Yanmei Hu[1], Tommy Szeto[1], Xiujun Zhang[2], Bart Tarbet[3,4], Michael Thomas Marty[5], Yu Chen[2] and Jun Wang[1]

A new coronavirus SARS-CoV-2, also called novel coronavirus 2019 (2019-nCoV), 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 6.35% as of May 26, 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 substrate-based peptidomimetic M^{pro} inhibitors. A complex crystal structure of SARS-CoV-2 M^{pro} with GC-376, determined at 2.15 Å resolution with three protomers 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 26, 2020, there are 343,562 deaths among 5,406,282 confirmed cases in 215 countries (https://www.who.int/emergencies/diseases/novel-coronavirus-2019). 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-CoV, the virus that causes severe acute respiratory symptoms in human and killed 774 people among 8098 infected worldwide in 2003.[1] SARS-CoV-2 shares ~82% of sequence identity with SARS-CoV and to a less extent with Middle East respiratory syndrome coronavirus (MERS-CoV) (~50%).[2,3] SARS-CoV-2 is an enveloped, positive-sense, single-stranded RNA virus that belongs to the β-lineage of the coronavirus.[4] The β-lineage also contains two other important human pathogens, the SARS-CoV and MERS-CoV. The mortality rate of COVID-19 is around 6.35% as of May 26, 2020, which is lower than that of SARS (~10%) and MERS (~34%).[1] However, current data indicate that COVID-19 is more contagious and has a larger R0 value than SARS and MERS[5], 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 any of these viral proteins might be potential antiviral drug candidates.[6,7] 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.[8–11] The M^{pro} has a 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 that 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

[1]Department of Pharmacology and Toxicology, College of Pharmacy, The University of Arizona, Tucson, AZ 85721, USA; [2]Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA; [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 and [5]Department of Chemistry and Biochemistry, The University of Arizona, Tucson, AZ 85721, USA
Correspondence: Yu Chen (ychen1@usf.edu) or Jun Wang (junwang@pharmacy.arizona.edu)
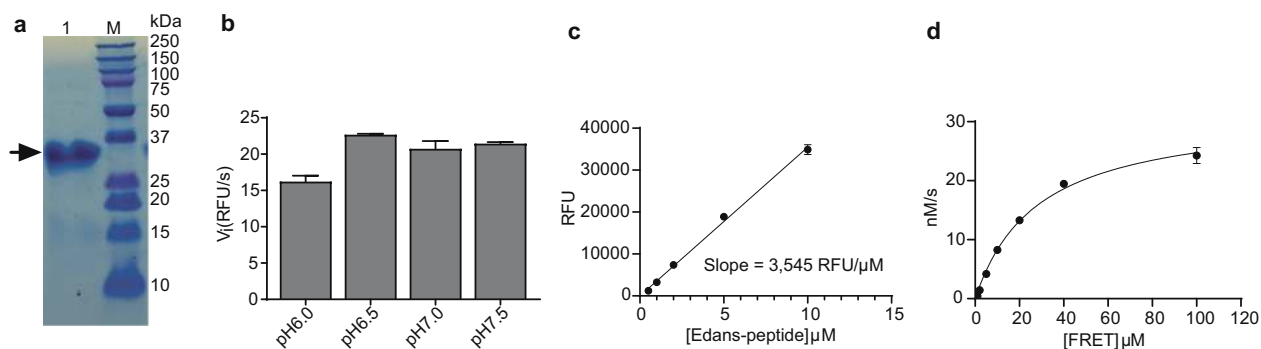
Fig. 1 SARS-CoV-2 M^{pro} expression and characterization. a SDS-PAGE of His-tagged M^{pro} (lane 1); Lane M, protein ladder; the calculated molecular weight of the His-tagged M^{pro} is 34,861 Da. **b** Reaction buffer optimization: 250 nM His-tagged M^{pro} was diluted into three reaction buffers with different pH values. **c** An Edans standard curve was generated to convert RFU to the amount of the cleaved substrate (nM). **d** Michaelis–Menten plot of 200 nM His-tagged M^{pro} with various concentrations of FRET substrate in pH 6.5 reaction buffer. The best-fit $V_{max} = 31.7 \pm 1.6$ nM/s; $K_m = 28.2 \pm 3.4$ µM, and the calculated $k_{cat}/K_m = 5,624$ s^{-1} M^{-1}.

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 some of the most potent and selective SARS-CoV-2 M^{pro} inhibitors so far with both enzymatic inhibition and cellular antiviral activity.[8,10,11] The X-ray crystal structure of SARS-CoV-2 M^{pro} with GC-376 showed that the compound can adopt 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 suggests that it might be feasible to design dual inhibitors against the viral M^{pro} and the host calpains/cathepsins, both of which are important for viral replication.

RESULTS
Establishing the FRET-based assay for the SARS-CoV-2 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) *Escherichia Coli.* with a His-tag in its C-terminus. The M^{pro} protein was purified with Ni-NTA column to high purity (Fig. 1a). The determined molecular weight of M^{pro} is 34,861 Da from native mass spectrometry, which matches the calculated molecular weight of M^{pro} without the N-terminal methionine (Supplementary information, Fig. S1). The first methionine might be cleaved by the methionine aminopeptidase as reported before.[12] 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).[8] We then tested the M^{pro} proteolytic activity in buffers with different pH. We found that M^{pro} displays the highest activity in pH 6.5 buffer (Fig. 1b), which contains 20 mM HEPES, 120 mM NaCl, 0.4 mM EDTA, 4 mM 1,4-dithiothreitol (DTT) and 20% glycerol. As such, all the following proteolytic assay was conducted using this pH 6.5 buffer. A standard curve was generated to convert the relative fluorescence unit (RFU) to the amount of the cleaved substrate (nM) such that k_{cat} can be calculated (Fig. 1c). Next, we characterized the enzymatic activity of this SARS-CoV-2 M^{pro} by measuring the K_m and V_{max} values. When 200 nM M^{pro} was mixed with various concentration of FRET substrate (0–100 µM), the initial velocity was measured and plotted against substrate concentration. Curve fitting with Michaelis–Menten equation gave the best-fit values of K_m and V_{max} as 28.2 ± 3.4 µM and 31.7 ± 1.6 RFU/s, respectively (Fig. 1d). The calculated k_{cat}/K_m was 5,624 s^{-1} M^{-1}, which is similar to the previously reported value of 3,426.1 s^{-1} M^{-1} by Hilgenfeld et al.[8],

suggesting that the C-terminal His-tag M^{pro} is as active as the native M^{pro}.

Primary screening of a focused protease inhibitor 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 (Table 1). The protease inhibitors are grouped based on their targets and mechanism of action, including 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 min in the presence of 4 mM DTT before the addition of 10 µM FRET substrate. The addition of DTT was to quench non-specific thiol reactive compounds and to ensure that 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 (Fig. 2). 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.

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 combination 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 50% inhibitory concentration (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 °C and 5.18 °C, 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 that they are not non-specific cysteine protease inhibitors. The calpain inhibitor MG-132 (**43**) had an IC$_{50}$ value of 3.90 µM against M^{pro}, and was not active

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**)

Table 1. continued

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**)

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 °C.

In light of the promising results of the calpain inhibitor MG-132 (**43**), we then pursued testing for 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 °C, 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-CoV, 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 °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 inactive 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]

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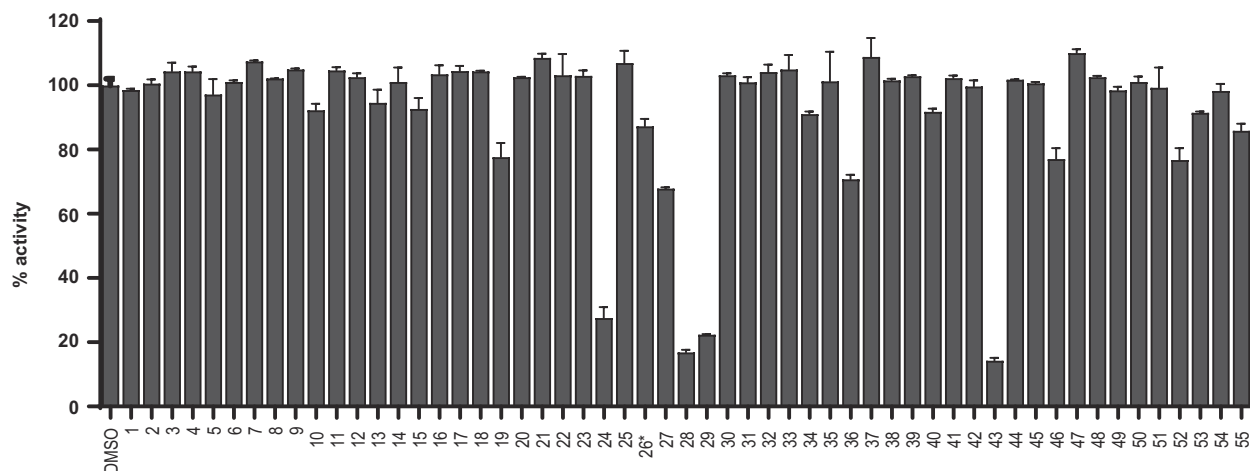
Fig. 2 Screening of known protease inhibitors against SARS-CoV-2 M^{pro} using the FRET assay. 20 µM of compounds (**26** was tested at 2 µM) was pre-incubated with 100 nM of SARS-CoV-2 M^{pro} for 30 min at 30 °C, and then 10 µM FRET substrate was added to reaction mixture to initiate the reaction.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

stabilization of the M^{pro} against thermal denaturation was also compound concentration dependent (Fig. 3b).

The binding of the 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) (Fig. 3c–f). Native MS analysis showed that M^{pro} formed a dimer complex with a mass of 69,722 Da, indicating that the extra N-terminal methionine in our M^{pro} construct was cleaved, yielding the native M^{pro} N-terminus with a serine residue (Supplementary information, Fig. S1). A small amount of protomer and dimer with a single C-terminal truncation of the His-tag was also observed, but the intact dimer was the predominant signal (Supplementary information, 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 protomer.

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 absence of inhibitor is typically a hallmark for a slow covalent binding inhibitor. In Fig. 4, the left column shows the progression curves up to 4 h. 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 min. We therefore chose the first 90 min of the progression curves for curve fitting (Fig. 4, middle column). We fit the progression curves in the presence of different concentrations of GC-376 (**64**) with the two-step Morrison equation (Eq. (3) in "Materials and 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 covalent bond is formed at a slower velocity between GC-376 (**64**) and M^{pro} with the second reaction rate constant (k$_2$) being 0.00245 ± 0.00047 s^{-1}, resulting in an overall k$_2$/K$_I$ value of 4.08 × 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 (Fig. 4b–e). 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.

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 (Supplementary information, 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 (Supplementary information, 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 50% effective concentration (EC$_{50}$) values ranging from 0.49 to 3.37 µM (Fig. 5a–d and Table 3). Their antiviral activity was further confirmed in the secondary viral yield reduction (VYR) assay (Fig. 5e–h and Table 3). The most potent compound was calpain inhibitor XII (**62**), which showed an EC$_{50}$ of 0.49 µM in the primary CPE assay and an EC$_{50}$ of 0.78 µM in the secondary VYR assay (Table 3; Supplementary information, Fig. S2). In comparison, remdesivir was reported to inhibit SARS-CoV-2 in the VYR assay with an EC$_{50}$ of 0.77 µM.[20] None of the compounds inhibited the unrelated influenza virus A/California/07/2009 (H1N1) virus (EC$_{50}$ > 20 µM) (Supplementary information, Table S1), suggesting that 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 some

Table 2. Characterization of HCV and calpain/cathepsin protease inhibitors against SARS-CoV-2 M^pro using a consortium of secondary assays[a].
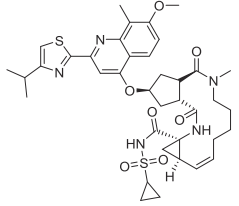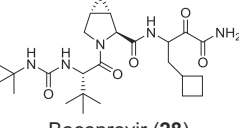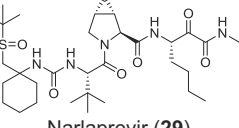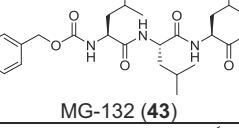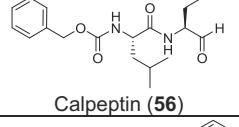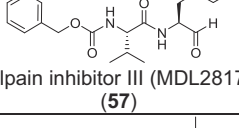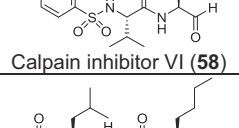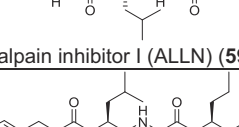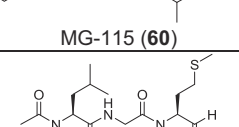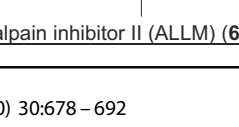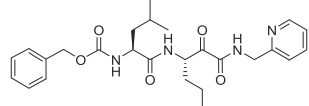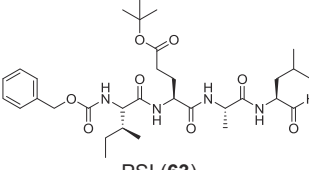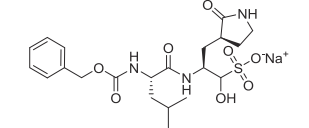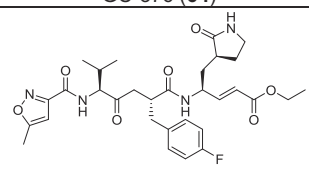
ID / Results	SARS-CoV-2 M^pro IC$_{50}$ (μM)	SARS-CoV-2 M^pro TSA Tm/ΔTm (°C)	EV-A71 2A IC$_{50}$ (μM)	EV-A71 3C IC$_{50}$ (μM)	Development stage
DMSO	-------	55.74 ± 0.00	------	-------	
Simeprevir (**24**)	13.74 ± 3.45	N.T.[b]	N.T.	N.T.	FDA-approved HCV drug
Boceprevir (**28**)	4.13 ± 0.61	62.41 ± 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 (**43**)	3.90 ± 1.01	59.76 ± 0.45/4.02	>20	>20	Preclinical; tested in mice[48]
Calpeptin (**56**)	10.69 ± 2.77	56.84 ± 0.00/1.1	>20	>20	Preclinical; tested in mice and felines[49,50]
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[51]
Calpain inhibitor I (ALLN) (**59**)	8.60 ± 1.46	N.T.	>20	>20	Preclinical; tested in mice[52]
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

Table 2. continued

Compound					
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[53]
GC-376 (**64**)	0.030 ± 0.008	74.04 ± 0.07/18.30	>20	0.136 ± 0.025	Preclinical; tested in felines[14, 15]
Rupintrivir (**65**)	> 20	N.T.	>20	0.042 ± 0.014	Dropped out of clinical trial

[a]Value = means ± SE from three 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 h digestion, instead of the initial velocity.

of the most potent and selective drug candidates with broad chemical diversity.

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$_2$21 space group at 2.15 Å resolution (PDB: 6WTT) (Supplementary information, Table S2). There are three protomers per asymmetric unit (ASU) with two constituting a biological dimer and the third forming a dimer with a crystallographic symmetry related neighboring protomers (Supplementary information, Fig. S3). The presence of three protomers in our crystal structure allowed us to capture different binding configurations of GC-376 (**64**) (Fig. 6), a unique feature that was not observed in previous X-ray crystal structures.[8,10,11] The pairwise r.m.s.d among the protomer backbone Cα atoms ranges from 0.435 to 0.564 Å. Previously, SARS-CoV M^{pro} and SARS-CoV-2 M^{pro} crystal structures have been solved most frequently as one protomer per ASU, and occasionally a dimer.[8,10,11,21,22] In its native state, M^{pro} requires dimerization to become catalytically active[23,24], which is supported by our native MS data (Supplementary information, Fig. S1). In our crystal structures, all three protomers appear catalytically competent, with the third protomer activated by the N-finger from an adjacent ASU. The well-defined electron density also clearly shows a serine at the protease N-terminus, again indicating that the first methionine in the protease construct was cleaved (Supplementary information, Fig. S3b).

GC-376 (**64**) forms an extensive network of hydrogen bonds with the active site while also exhibiting excellent geometric complementarity (Fig. 6). 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. 6e).[8,22,25,26] 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 (Fig. 6a, 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).[27,28] 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 (**64**) exhibits high complementarity with the S4 site, and the extensive non-polar interactions may contribute significantly to the potency of this compound (Fig. 6e).

Three copies of GC-376 (**64**) were found in the crystal structure, one in each protomer active site (Fig. 6a, c; Supplementary information, Fig. S4). The configurations of GC-376 (**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 (Fig. 6a; Supplementary information, Fig. S4), resulting in the (S)-configuration. It is noted that aldehydes (**11a** and **11b**) also bind the active site of SARS-CoV-2 M^{pro} in the (S)-configuration (PDB: 6M0K and 6LZE) (Supplementary information, Fig. S5a).[11] In protomer C, however, the same hydroxide group orients outwards from the oxyanion hole, forming hydrogen bonds with His41 (Fig. 6c), 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) (Supplementary information, Fig. S5b).[8] These
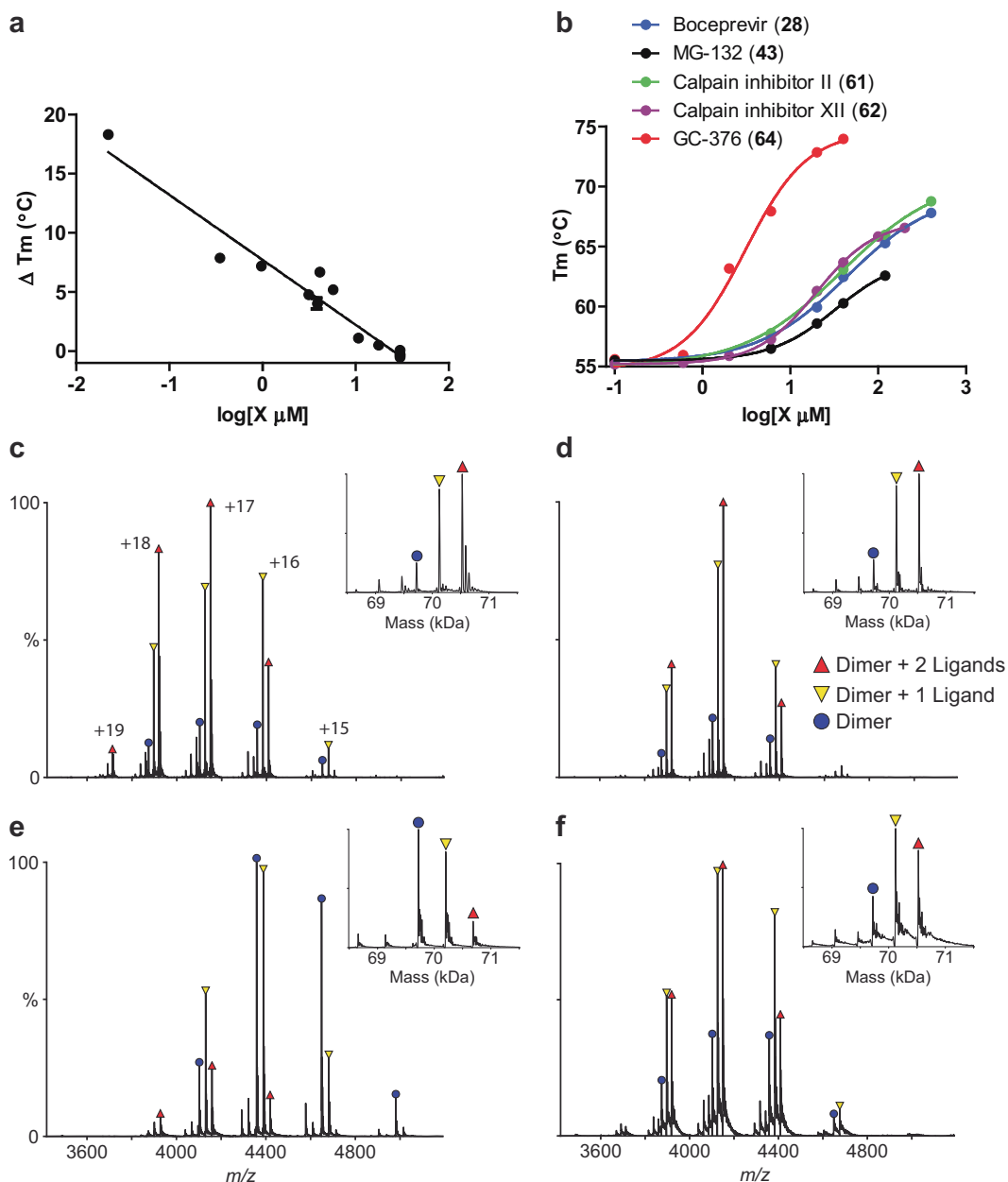
Fig. 3 Binding of inhibitors to SARS-CoV-2 M^{pro} using thermal shift binding assay and native mass spectrometry. a Correlation of inhibition efficacy (IC$_{50}$) with ΔT_m from thermal shift binding assay. Data in Table 2 were used for the plot. The r^2 of fitting is 0.94. **b** Dose-dependent melting temperature (T$_m$) shift. Native MS reveals binding of SARS-CoV-2 M^{pro} to GC-376 (**64**) (**c**), calpain inhibitor II (**61**) (**d**), calpain inhibitor XII (**62**) (**e**), and boceprevir (**28**) (**f**). All ligand concentrations are 12.5 μM except (**e**), which is 25 μM. Peaks 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.

two unique configurations R and S might be a result of the Cys145 thiol nucleophilic attacking the aldehyde of GC-376 (**64**) from two different faces (Fig. 6b, d). The fact that GC-376 (**64**) can adopt 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 protomers A, B and C is observed in the orientation of the phenylmethyl ester. In protomer C, the CH$_2$ of the phenylmethyl points toward the main chain of Leu167 in a "cis" conformation (Fig. 6c), whereas in protomers A and B this same CH$_2$ points downwards in a "trans" conformation (Fig. 6a; Supplementary information, Fig. S4). 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 loop that consists of residues 45–51 (TSEDMLN in SARS-CoV-2 M^{pro}) and forms the S2 and S3' subsites (Fig. 6f), 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 ASUs and resemble the conformations in previously determined structures.[10,11] 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
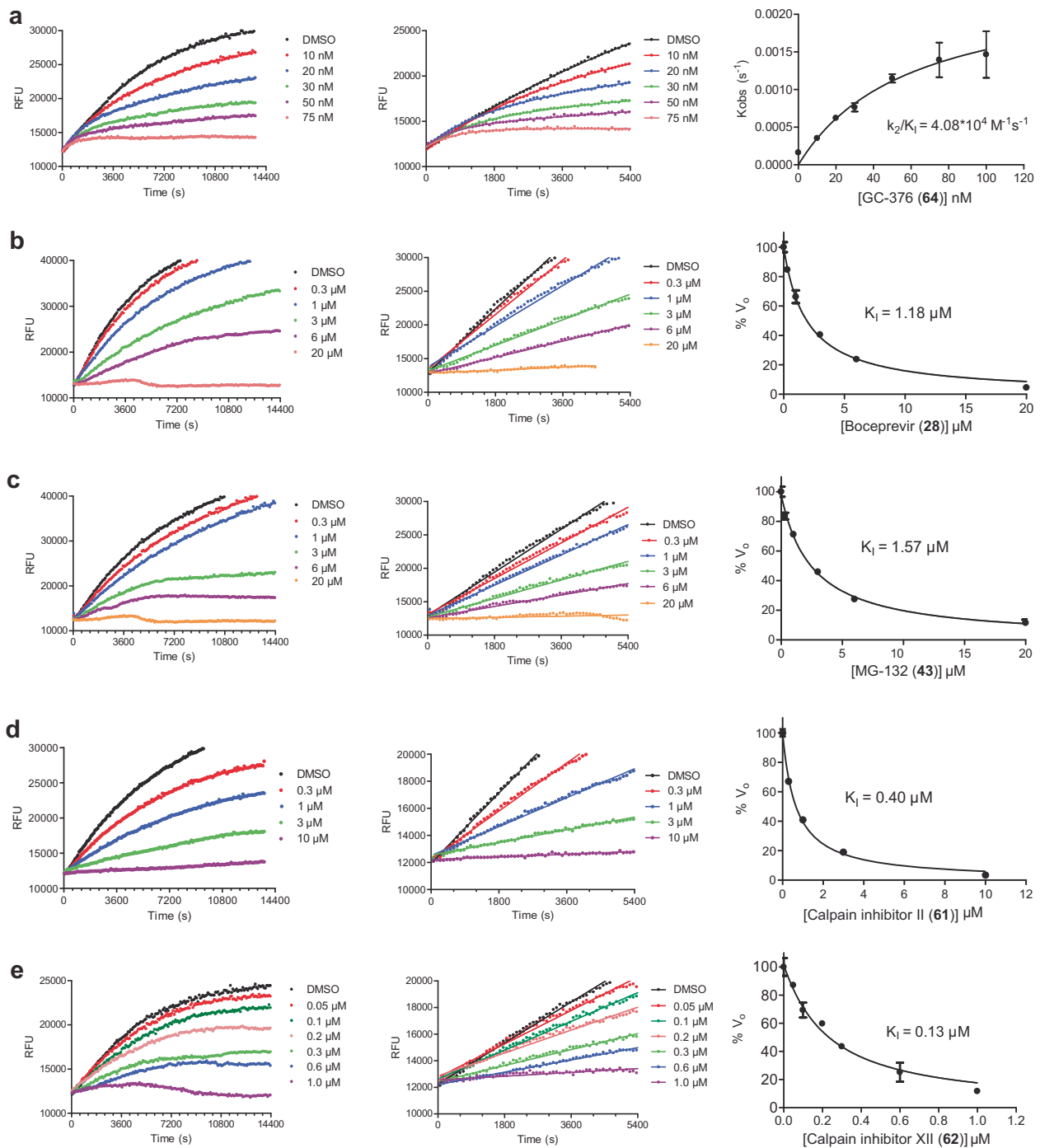
Fig. 4 Proteolytic reaction progression curves of M^pro in the presence or the absence of compounds. a–e In the kinetic studies, 5 nM M^pro was added to a solution containing various concentrations of protease inhibitors and 20 μM FRET substrate to initiate the reaction. The reaction was then monitored for 4 h. The left column shows the reaction progression up to 4 h; middle column shows the progression curves for the first 90 min, which were used for curve fitting to generate the plot shown in the right column. Detailed methods were described in "Materials and methods" section. GC-376 (**64**) (**a**); boceprevir (**28**) (**b**); MG-132 (**43**) (**c**); calpain inhibitor II (**61**) (**d**); calpain inhibitor XII (**62**) (**e**).

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 20 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.[6] The viral polymerase inhibitor remdesivir holds the greatest promise and it is currently being evaluated in several clinical trials.[29,30] Remdesivir was recently granted emergency use authorization by FDA. The HIV drug combination lopinavir and ritonavir recently failed in a clinical trial for COVID-19 as no significant therapeutic efficacy was observed.[31] To address this
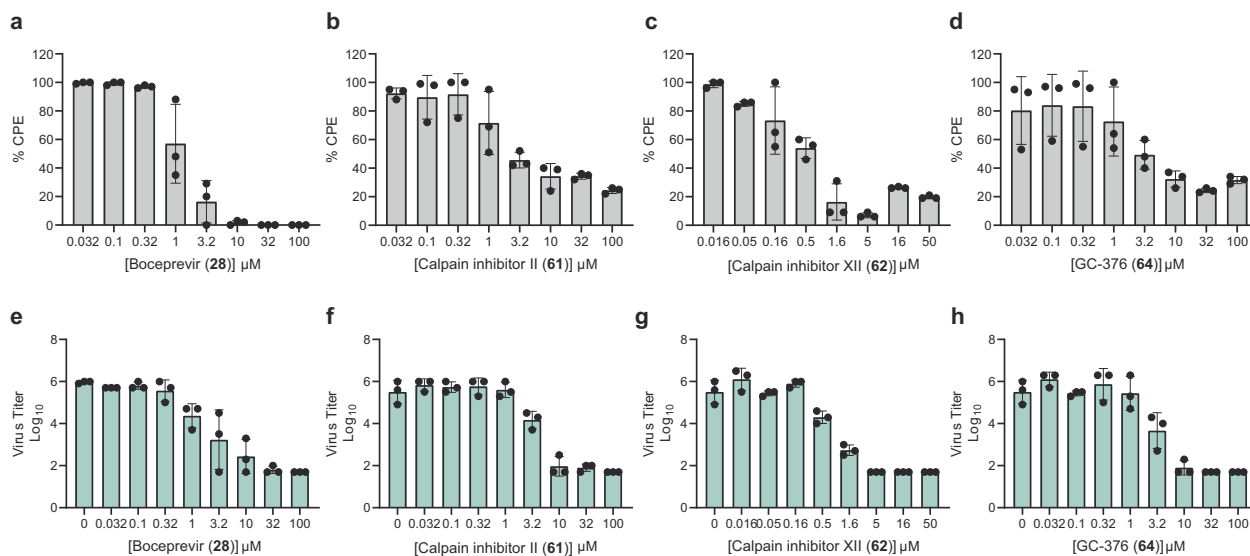
Fig. 5 Antiviral potency of boceprevir (**28**), calpain inhibitors II (**61**), XII (**62**), and GC-376 (**64**), in CPE and VYR assays. **a**–**d** Primary CPE assay results of boceprevir (**28**), calpain inhibitors II (**61**), XII (**62**), and GC-376 (**64**). **e**–**h** Secondary VYR assay results of boceprevir (**28**), calpain inhibitors II (**61**), XII (**62**), and GC-376 (**64**). Each dot represents the average of three repeats from one experiment, and each assay was repeated in three different experiments. The curve fittings were shown in Supplementary information, Fig. S2.

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-CoV and MERS-CoV.[32] As the SARS-CoV-2 M^{pro} shares a high sequence similarity with SARS-CoV and to a less extent with MERS-CoV, we reasoned that inhibiting the enzymatic activity of SARS-CoV-2 M^{pro} will similarly prevent viral replication.[8,10]

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 CPE assay. The therapeutic potential of boceprevir (**28**) should be further evaluated in relevant animal models and human clinical trials. Since boceprevir (**28**) is an 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 adopt 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} 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 II and cathepsin L 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
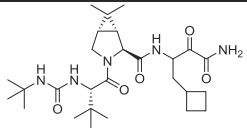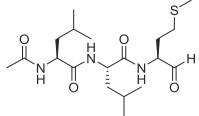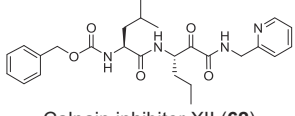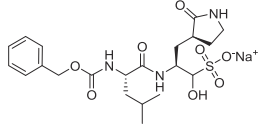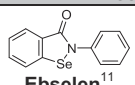
replication.[33] Calpain and cathepsin inhibitors such as MDL28170 (calpain inhibitor III),[33] MG-132,[34] calpain inhibitor VI[35] have been shown to inhibit SARS-CoV replication in cell culture. Other than the increased potency of targeting both M^{pro} and calpain/cathepsin, an additional benefit 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,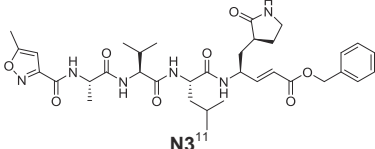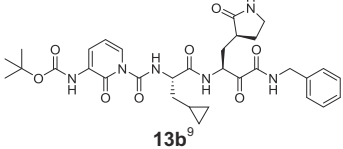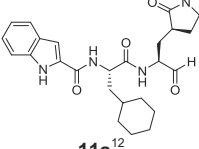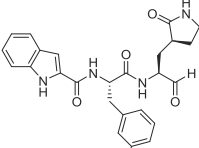 kidney diseases, and ischemia/reperfusion injury.[36] Given our promising results of calpain inhibitors II (**61**) and XII (**62**) in inhibiting the SARS-CoV-2 M^{pro} and their potent antiviral activity in cell culture, it might be worthwhile to repurpose them as antivirals for SARS-CoV-2.

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 some of the most potent and selective hits reported so far, and have similar activity to the recently reported SARS-CoV-2 M^{pro} inhibitors (**ebselen**, **N3**, and **13b**) (Table 3). Calpain inhibitor XII (**62**) had similar potency to the recently disclosed compounds (**11a** and **11b**) (Table 3).[11] Notably, boceprevir (**28**), calpain inhibitor II (**61**) and XII (**62**) have different chemical scaffolds than the substrate mimetics 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 why they lack efficacy in clinical trials for COVID-19.[31] 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.[37] In our study, we found that 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

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} = 1.31 ± 0.58 CC_{50} > 100 SI_{50} > 76.3	EC_{50} = 1.95 ± 1.62	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_{50} = 3.70 ± 0.69	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_{50} = 0.78 ± 0.37	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_{50} = 3.14 ± 1.01	Preclinical; tested in felines[14, 15]
Recently reported SARS-Cov-2 M^{pro} inhibitors[b]				
Ebselen[11]	IC_{50} = 0.67 ± 0.09	4.67 ± 0.80[c]	N. A.	In clinical trials
N3[11]	$k_{obs}/[I]$ = 11,300 ± 800 $M^{-1}S^{-1}$	16.77 ± 1.70	N. A.	Preclinical;[28] 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

N.A. not available.
[a]CPE EC_{50}, VYR EC_{90}, and cytotoxicity CC_{50} values are means ± SD of 3 independent experiments.
[b]Results were retrieved from recent publications.[8,10,11]
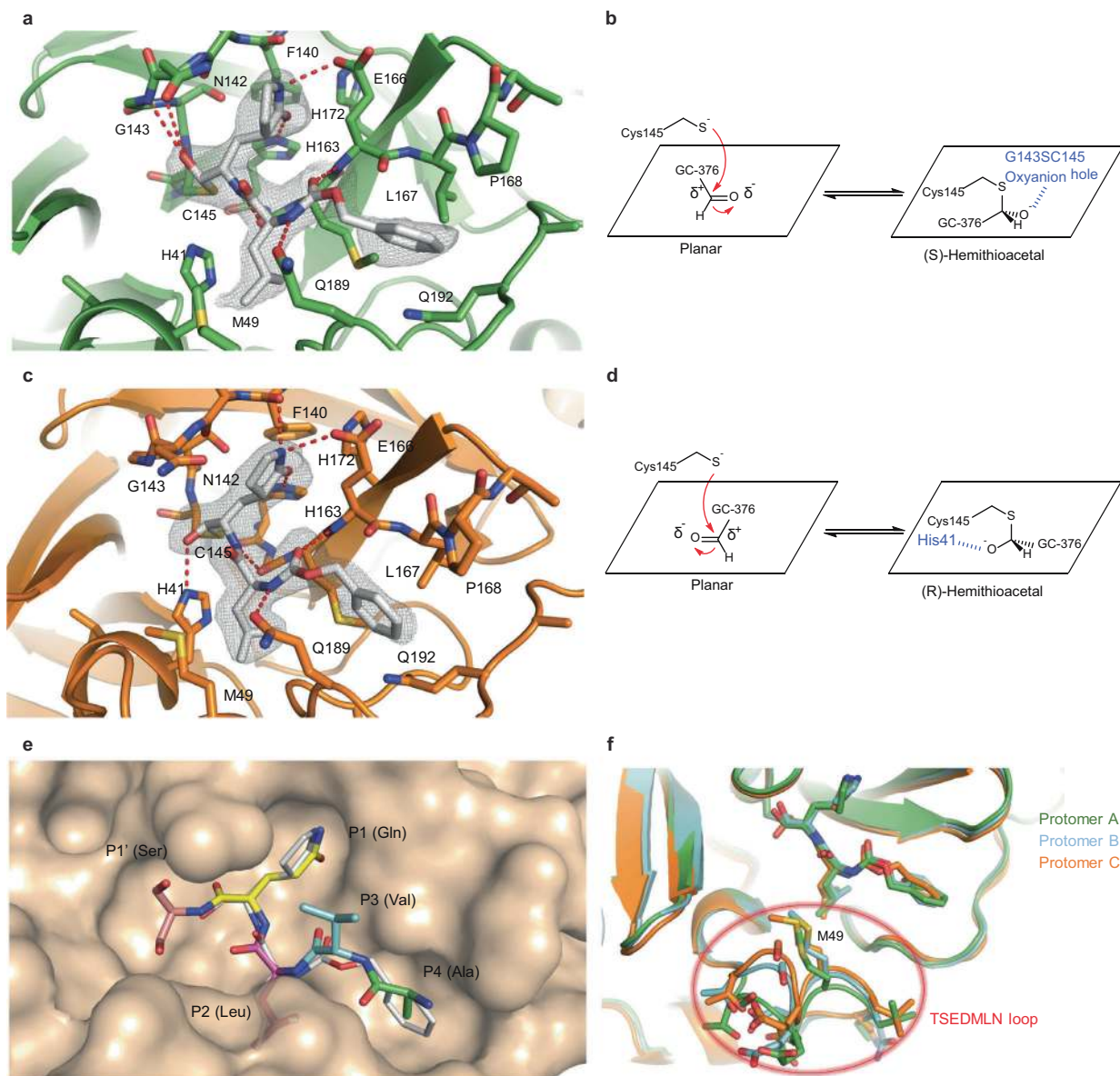[c]The antiviral activity of ebselen was determined in plaque assay.

Fig. 6 Molecular recognition of GC-376 (64) by SARS-CoV-2 M^pro. Complex of SARS-CoV-2 M^pro and GC-376 (**64**) with protomer A (**a**, **b**) and protomer C (**c**, **d**). Unbiased F$_o$–F$_c$ map, shown in gray, is contoured at 2σ. 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.

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 synergistic antiviral effect as well as to suppress drug resistance.

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% CO$_2$ atmosphere. The USA-WA1/2020 strain of SARS-CoV-2 was 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, USA) in the pET29a(+) vector with *E. coli* codon optimization. pET29a(+) plasmids with SARS-CoV-2 M^pro was transformed into competent *E. coli* BL21(DE3) cells, and a single colony was picked to inoculate 10 mL of LB supplemented with 50 g/mL kanamycin and grown at 37 °C with shaking at 250 rpm. The 10-mL inoculum was added to 1 L of LB with 50 g/mL kanamycin and grown to an optical density at 600 nm of 0.8, and 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 DTT with 0.5 mg/mL lysozyme, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 0.02 mg/mL DNase I), and lysed by 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 elution buffer (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 260 nm absorbance with ε 32890. EV-A71 2A^{pro} and 3C^{pro} were expressed in the pET28b(+) vector as previously described.[18,38,39]

The protein sequence of the SARS-CoV-2 M^{pro} used in the enzymatic assay and X-ray crystallization is listed below:

SGFRKMAFPS GKVEGCMVQV TCGTTTLNGL WLDDVVYCPR HVICT-SEDML NPNYEDLLIR KSNHNFLVQA GNVQLRVIGH SMQNCVLKLK VDTANPKTPK YKFVRIQPGQ TFSVLACYNG SPSGVYQCAM RPNF-TIKGSF LNGSCGSVGF NIDYDCVSFC YMHHMELPTG VHAGTDLEGN FYGPFVDRQT AQAAGTDTTI TVNVLAWLYA AVINGDRWFL NRFTT-TLNDF NLVAMKYNYE PLTQDHVDIL GPLSAQTGIA VLDMCASLKE LLQNGMNGRT ILGSALLEDE FTPFDVVRQC SGVTFQ LEHHHHHH.

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.[40] 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 min at 80 °C. For deprotection, 5% piperazine plus 0.1 M HOBt were used and the mixture was heated at 80 °C for 5 min. The peptide was cleaved from the resin using 95% TFA, 2.5% Tris and 2.5% H$_2$O, and the crude peptide was precipitated from ether after removal of TFA. The final peptide was purified by preparative HPLC. The purify and identity of the peptide were confirmed by analytical HPLC (> 98% purity) and mass spectrometry. [M + 3]$^{3+}$ was calculated as 694.15 and detected as 694.90; [M + 4]$^{4+}$ was calculated as 520.86 and detected as 521.35.

Edans standard curve

Edans standard curve was generated as described below: 200 nM SARS-CoV-2 M^{pro} was incubated with varying concentrations of FRET substrate (0.5–40 μM) and the reaction progress was monitored until the fluorescence signals reached plateau, at which point we deemed all the FRET substrate was digested by M^{pro}. The endpoint fluorescence signal was plotted against FRET substrate concentration with a linear regression function in Prism 5.

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 10× 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 min at room temperature prior to analysis. Each sample was mixed and analyzed in triplicate.

Native 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 1000–25,000 m/z range, the resolution set to 30,000, and a trapping gas pressure set to 3. Source fragmentation (10–50 V) was applied to all samples to

aid in desolvation. Data were deconvolved and analyzed with UniDec.[41]

Enzymatic assays

For reaction condition optimization, 200 μM SARS-CoV-2 M^{pro} was used. pH 6.0 buffer contains 20 mM MES, pH 6.0, 120 mM NaCl, 0.4 mM EDTA, 4 mM DTT and 20% glycerol; pH 6.5 buffer contains 20 mM HEPES, pH 6.5, 120 mM NaCl, 0.4 mM EDTA, 4 mM DTT and 20% glycerol; pH 7.0 buffer contains 20 mM HEPES, pH 7.0, 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 h. The first 15 min of reaction was used to calculate initial velocity via linear regression in Prism 5. M^{pro} displays the highest proteolytic activity in pH 6.5 buffer. All the following enzymatic assays were carried out in pH 6.5 buffer.

For the measurements of K_m/V_{max}, screening of the protease inhibitor library, as well as IC$_{50}$ measurements, proteolytic reaction with 100 nM M^{pro} in 100 μL of pH 6.5 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 of protease inhibitor library and IC$_{50}$ measurements, 100 nM M^{pro} was incubated with protease inhibitor at 30 °C for 30 min in reaction buffer, 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 M^{pro} 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 h. The progress curves were fit to a slow binding Morrison equation (Eq. (3)) as described previously[18,42]:

$$E + I \underset{k_1}{\overset{K_2}{\rightleftharpoons}} EI \underset{k_1}{\overset{K_2}{\rightleftharpoons}} EI^* \tag{1}$$

$$K_I = k_{-1}/k_1 \tag{2}$$

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

$$k = k_2[I]/(K_I + [I]) \tag{4}$$

where P(t) is the fluorescence signal at time t, P$_0$ is the background signal at time 0, V$_0$, V$_s$, 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*.[42] 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 (Fig. 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 cases: 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

The binding of protease inhibitors to M^{pro} was monitored by differential scanning fluorimetry (DSF) using a Thermo Fisher QuantStudio^{TM} 5 Real-Time PCR System. TSA plates were prepared by mixing M^{pro} protein (final concentration of 3 μM) with inhibitors, and incubated at 30 °C for 30 min. 1× SYPRO orange (Thermo Fisher Scientific) 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 that 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% CO$_2$ atmosphere to ~90% confluence on the next day. Cells were washed with phosphate-buffered saline (PBS) and 200 μL DMEM containing 2% FBS and 1% penicillin–streptomycin, and various concentrations of protease inhibitors were added to each well. 48 h after addition of 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 Graph-Pad 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-fold of 50% cell culture infectious doses (CCID$_{50}$) of SARS-CoV-2 and 50% effective concentrations (EC$_{50}$) were calculated based on virus-induced CPE quantified by neutral red dye uptake after 5 days of incubation. USA-WA1/2020 strain of SARS-CoV-2 obtained from the WRCEVA was used for both CPE and VYR assays. Three microwells at each concentration of compound were infected. Two uninfected microwells served as toxicity controls. Cells were stained for viability with neutral red (0.11% final concentration) for 2 h. Excess dye was rinsed from the cells with 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

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 and 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 post 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.[43]

M^{pro} crystallization and structure determination

10 mg/mL of SARS-CoV-2 M^{pro} was incubated with 2 mM GC376 at 4 °C overnight. 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 μL of the protein solution with 1 μL of the precipitant solution (15% PEG 2 K, 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 2 K, 10% 1,6-hexanediol, 20% glycerol) and flash-frozen in liquid nitrogen.

X-ray diffraction data for the SARS-CoV-2 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.[44] The CCP4 versions of MOLREP was used for molecular replacement using a previously solved SARS-CoV-2 M^{pro} structure (PDB ID: 5RGG) as a reference model.[45] Rigid and restrained refinements were performed using REFMAC and model building was performed with COOT.[46,47] Protein structure figures were made using PyMOL (Schrödinger, LLC).

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.D.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 assays; 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 paper with the input from others.

ADDITIONAL INFORMATION

Supplementary information accompanies this paper at https://doi.org/10.1038/s41422-020-0356-z.

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. Mahase, E. Coronavirus: covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate. *BMJ* **368**, m641 (2020).
2. Lu, R. et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).
3. Wu, A. et al. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* **27**, 325–328 (2020).

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

5. Tang, B. et al. An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov). *Infect. Dis. Model.* **5**, 248–255 (2020).

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

7. Xia, S. 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.* **30**, 343–355 (2020).

8. Zhang, L. et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* **368**, 409–412 (2020).

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

10. Jin, Z. et al. Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).

11. Dai, W. et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* https://doi.org/10.1126/science.abb4489 (2020).

12. Wingfield, P. T. N-terminal methionine processing. *Curr. Protoc. Protein Sci.* **88**, 6.14.1–6.14.3 (2017).

13. Pedersen, N. C. et al. Efficacy of a 3C-like protease inhibitor in treating various forms of acquired feline infectious peritonitis. *J. Feline Med. Surg.* **20**, 378–392 (2018).

14. Kim, Y. et al. Reversal of the progression of fatal coronavirus infection in cats by a broad-spectrum coronavirus protease inhibitor. *PLoS Pathog.* **12**, e1005531 (2016).

15. Kim, Y. et al. Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. *J. Virol.* **86**, 11754–11762 (2012).

16. Hayden, F. G. 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.* **47**, 3907–3916 (2003).

17. Shie, J. J. et al. Inhibition of the severe acute respiratory syndrome 3CL protease by peptidomimetic alpha,beta-unsaturated esters. *Bioorg. Med. Chem.* **13**, 5240–5252 (2005).

18. Musharrafieh, R. 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.* **93**, e0221-18 (2019).

19. Kuo, C. J. et al. Design, synthesis, and evaluation of 3C protease inhibitors as anti-enterovirus 71 agents. *Bioorg. Med. Chem.* **16**, 7388–7398 (2008).

20. Wang, M. et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **30**, 269–271 (2020).

21. Yang, H. et al. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad. Sci. USA* **100**, 13190–13195 (2003).

22. Muramatsu, T. et al. SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proc. Natl Acad. Sci. USA* **113**, 12997–13002 (2016).

23. Li, C. et al. Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like proteinase. *J. Biol. Chem.* **285**, 28134–28140 (2010).

24. Anand, K. et al. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J.* **21**, 3213–3224 (2002).

25. Xue, X. et al. Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J. Virol.* **82**, 2515–2527 (2008).

26. Zhang, L. et al. Alpha-ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication: structure-based design, synthesis, and activity assessment. *J. Med. Chem.* **63**, 4562–4578 (2020).

27. Chuck, C. P. et al. Profiling of substrate specificity of SARS-CoV 3CL. *PLoS ONE* **5**, e13197 (2010).

28. Hegyi, A. & Ziebuhr, J. Conservation of substrate specificities among coronavirus main proteases. *J. Gen. Virol.* **83**, 595–599 (2002).

29. Grein, J. et al. Compassionate use of remdesivir for patients with severe Covid-19. *N. Engl. J. Med.* **382**, 2327–2336 (2020).

30. Holshue, M. L. et al. First case of 2019 novel coronavirus in the United States. *N. Engl. J. Med.* **382**, 929–936 (2020).

31. Cao, B. et al. A trial of lopinavir–ritonavir in adults hospitalized with severe Covid-19. *N. Engl. J. Med.* **382**, 1787–1799 (2020).

32. Zumla, A., Chan, J. F., Azhar, E. I., Hui, D. S. & Yuen, K. Y. Coronaviruses—drug discovery and therapeutic options. *Nat. Rev. Drug Discov.* **15**, 327–347 (2016).

33. Simmons, G. et al. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc. Natl Acad. Sci. USA* **102**, 11876–11881 (2005).

34. Schneider, M. et al. Severe acute respiratory syndrome coronavirus replication is severely impaired by MG132 due to proteasome-independent inhibition of M-calpain. *J. Virol.* **86**, 10112–10122 (2012).

35. Barnard, D. L. et al. Inhibition of severe acute respiratory syndrome-associated coronavirus (SARSCoV) by calpain inhibitors and beta-D-N4-hydroxycytidine. *Antivir. Chem. Chemother.* **15**, 15–22 (2004).

36. Ono, Y., Saido, T. C. & Sorimachi, H. Calpain research for drug discovery: challenges and potential. *Nat. Rev. Drug Discov.* **15**, 854–876 (2016).

37. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280 (2020).

38. Shang, L. 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).

39. Cai, Q. et al. Conformational plasticity of the 2A proteinase from enterovirus 71. *J. Virol.* **87**, 7348–7356 (2013).

40. Cady, S. D., Wang, J., Wu, Y., DeGrado, W. F. & Hong, M. 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. *J. Am. Chem. Soc.* **133**, 4274–4284 (2011).

41. Marty, M. T. et al. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).

42. Morrison, J. F. & Walsh, C. T. The behavior and significance of slow-binding enzyme inhibitors. *Adv. Enzymol. Relat. Areas Mol. Biol.* **61**, 201–301 (1988).

43. Ma, C., Zhang, J. & Wang, J. 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).

44. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 859–866 (2006).

45. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 (2010).

46. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).

47. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

48. Bonuccelli, G. 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.* **163**, 1663–1675 (2003).

49. Mani, S. K. et al. In vivo administration of calpeptin attenuates calpain activation and cardiomyocyte loss in pressure-overloaded feline myocardium. *Am. J. Physiol. Heart Circ. Physiol.* **295**, H314–326 (2008).

50. 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.* **38**, 905–912 (2011).

51. Akdemir, O. et al. Therapeutic efficacy of SJA6017, a calpain inhibitor, in rat spinal cord injury. *J. Clin. Neurosci.* **15**, 1130–1136 (2008).

52. Li, S. Z. et al. ALLN hinders HCT116 tumor growth through Bax-dependent apoptosis. *Biochem. Biophys. Res. Commun.* **437**, 325–330 (2013).

53. Delli Pizzi, S. 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* **8**, e56501 (2013).