

## **An expedited approach towards the rationale design of non-covalent SARS-CoV-2 main protease inhibitors with in vitro antiviral activity**

Naoya Kitamura,<sup>1</sup> Michael Dominic Sacco,<sup>2</sup> Chunlong Ma,<sup>1</sup> Yanmei Hu,<sup>1</sup> Julia Alma Townsend,<sup>3</sup> Xiangzhi Meng,<sup>4</sup> Fushun Zhang,<sup>4</sup> Xiujun Zhang,<sup>2</sup> Adis Kukuljac,<sup>2</sup> Michael Thomas Marty,<sup>3</sup> David Schultz,<sup>5</sup> Sara Cherry,<sup>6</sup> Yan Xiang,<sup>4</sup> Yu Chen,<sup>2,\*</sup> Jun Wang<sup>1,\*</sup>

<sup>1</sup>Department of Pharmacology and Toxicology, College of Pharmacy, The University of Arizona, Tucson, AZ, 85721, United States

<sup>2</sup>Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, 33612, United States

<sup>3</sup>Department of Chemistry and Biochemistry, The University of Arizona, Tucson, AZ, 85721, United States

<sup>4</sup>Department of Microbiology, Immunology and Molecular Genetics, University of Texas Health Science Center at San Antonio, San Antonio, TX, 78229, United States

<sup>5</sup>Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA, 19104, United States

<sup>6</sup>Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA, 19104, United States

\*Corresponding authors:

Jun Wang, Tel: 520-626-1366, Fax: 520-626-0749, email: [junwang@pharmacy.arizona.edu](mailto:junwang@pharmacy.arizona.edu)

Yu Chen, Tel: 813-974-7809, email: [ychen1@usf.edu](mailto:ychen1@usf.edu)

**Keywords:** SARS-CoV-2, COVID-19, main protease, 3CL protease, ML188, GC376

## Abstract

The main protease ( $M^{pro}$ ) of SARS-CoV-2 is a validated antiviral drug target. Several  $M^{pro}$  inhibitors have been reported with potent enzymatic inhibition and cellular antiviral activity, including GC376, boceprevir, calpain inhibitors II and XII, each containing a reactive warhead that covalently modifies the catalytic Cys145. In this study, we report an expedited drug discovery approach by coupling structure-based design and Ugi four-component (Ugi-4CR) reaction methodology to the design of non-covalent  $M^{pro}$  inhibitors. The most potent compound **23R** had cellular antiviral activity similar to covalent inhibitors such as GC376. Our designs were guided by overlaying the structure of SARS-CoV  $M^{pro}$  + ML188 (R), a non-covalent inhibitor derived from Ugi-4CR, with the X-ray crystal structures of SARS-CoV-2  $M^{pro}$  + calpain inhibitor XII/GC376/UAWJ247. Binding site analysis suggests a strategy of extending the P2 and P3 substitutions in ML188 (R) to achieve optimal shape complementary with SARS-CoV-2  $M^{pro}$ . Lead optimization led to the discovery of **23R**, which inhibits SARS-CoV-2  $M^{pro}$  and SARS-CoV-2 viral replication with an  $IC_{50}$  of 0.31  $\mu$ M and  $EC_{50}$  of 1.27  $\mu$ M, respectively. The binding and specificity of **23R** to SARS-CoV-2  $M^{pro}$  were confirmed in a thermal shift assay and native mass spectrometry assay. The co-crystal structure of SARS-CoV-2  $M^{pro}$  with **23R** revealed the P2 biphenyl fits snugly into the S2 pocket and the benzyl group in the  $\alpha$ -methylbenzyl faces towards the core of the enzyme, occupying a previously unexplored binding site located in between the S2 and S4 pockets. Overall, this study revealed the most potent non-covalent SARS-CoV-2  $M^{pro}$  inhibitors reported to date and a novel binding pocket that can be explored for  $M^{pro}$  inhibitor design.

The COVID-19 pandemic had a significant impact on global economy and public health, and there is an urgent need for therapeutic interventions. Encouraging progress has been made in developing mRNA vaccines including the Pfizer BNT162b2 and Moderna mRNA-1273. For small molecule antivirals, the viral polymerase inhibitor remdesivir gained FDA approval on Oct 22<sup>nd</sup> 2020. The combination therapy of remdesivir with a Janus kinase (JAK) inhibitor baricitinib also received the FDA emergency use authorization<sup>1</sup>. Among the other drug targets being explored at different stages of preclinical and clinical developments<sup>2</sup>, the viral main protease ( $M^{pro}$ ), also called 3-chymotrypsin-like protease ( $3CL^{pro}$ ), is one of the most extensively explored high profile antiviral drug targets<sup>3</sup>.  $M^{pro}$  is a cysteine protease encoded in the viral polyprotein as non-structural protein 5 (Nsp5) that cleaves the viral polyproteins pp1a and pp1ab at more than 11 sites. Despite its multiple proteolytic sites,  $M^{pro}$  was shown to have a high substrate specificity of glutamine at the P1 position<sup>4</sup>. As such, the majority of the reported  $M^{pro}$  inhibitors were designed to contain a 2-pyrrolidone at the P1 substitution as a mimetic of the glutamine in the substrate<sup>5</sup>. Most advanced  $M^{pro}$  inhibitors including PF-07304814<sup>6</sup>, GC376<sup>7,8</sup> and 6j<sup>9</sup> all belong to this category. PF-07304814, an  $\alpha$ -hydroxyl ketone prodrug, is being developed by Pfizer, which has optimal pharmacokinetic properties in humans and recently entered human clinical trials<sup>6</sup>. GC376 has *in vivo* antiviral efficacy in treating cats infected with lethal feline infectious peritonitis virus<sup>10,11</sup>. Recently, the GC376 analog 6j was shown to protect mice from MERS-CoV infection<sup>9</sup>. These promising results highlight the translational potential of  $M^{pro}$  inhibitors as potent SARS-CoV-2 antivirals and validate  $M^{pro}$  as an antiviral drug target for coronaviruses.

Drug discovery is a lengthy process, which involves iterative cycles of design, synthesis, and pharmacological characterization. In the event of COVID-19 pandemic, an expedited approach with a fast turnover of this development cycle is highly desired. Using SARS-CoV-2  $M^{pro}$  as a drug target, we report herein a fast-track drug discovery approach

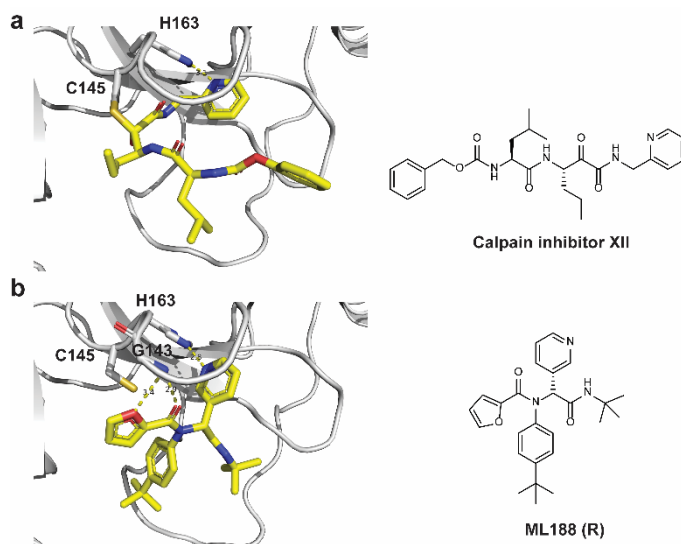
by coupling structure-based drug design and the Ugi four-component reaction (Ugi-4CR) methodology to the design of non-covalent M<sup>pro</sup> inhibitors. Specifically, through a screening of a focused library of protease inhibitors, we recently discovered several non-canonical SARS-CoV-2 M<sup>pro</sup> inhibitors including boceprevir, and calpain inhibitors II and XII<sup>7</sup>. These inhibitors differ from classic M<sup>pro</sup> inhibitors such as GC376 in that their P1 substitution does not contain a glutamine mimetic. The co-crystal structures of calpain inhibitors II and XII with SARS-CoV-2 M<sup>pro</sup> revealed a critical hydrogen bond between the methionine side chain from calpain inhibitor II and pyridinyl substitution from calpain inhibitor XII with the H163 side chain imidazole located at the S1 pocket<sup>8</sup>. Similarly, the carbonyl from the pyrrolidone in GC376 also forms a hydrogen bond with the H163 side chain imidazole<sup>7</sup>. Given the importance of this hydrogen bond with H163 for the high affinity binding of inhibitors to SARS-CoV-2 M<sup>pro</sup>, we hypothesize that non-covalent inhibitors without a reactive warhead targeting the Cys145, but retain the hydrogen bond capacity with H163 can be designed as potent SARS-CoV-2 M<sup>pro</sup> inhibitors. In this study, we report the structure-based design of non-covalent M<sup>pro</sup> inhibitors based on the overlaying structures of SARS-CoV or SARS-CoV-2 M<sup>pro</sup> in complex with existing inhibitors or the peptide substrate. The design was based on the scaffold of ML188 (R)<sup>12</sup>, a non-covalent SARS-CoV M<sup>pro</sup> inhibitor, which similarly contains a pyridinyl in the P1 substitution as calpain inhibitor XII. The overlaying structures revealed a strategy of extending the P2 and P3 substitutions in ML188 (R) to occupy the extra space in the S2 and S3/S4 pockets of SARS-CoV-2 M<sup>pro</sup> in order to increase the binding affinity. The most potent inhibitor from this study **23R** showed enzymatic inhibition and cellular antiviral activity similar to the covalent inhibitor GC376. Its mechanism of action was characterized in the thermal shift-binding assay, native mass spectrometry binding assay, and enzyme kinetic studies. An X-ray crystal structure of SARS-CoV-2 M<sup>pro</sup> in complex with **23R** was solved, revealing a previously unexplored binding site in between the S2 and S4 pockets. Overall, this study led to the identification of the most potent non-covalent M<sup>pro</sup> inhibitor

**23R** with potent enzymatic inhibition and *in vitro* cellular antiviral activity with a novel mechanism of action.

## RESULTS

### Rational design of non-covalent SARS-CoV-2 M<sup>pro</sup> inhibitors

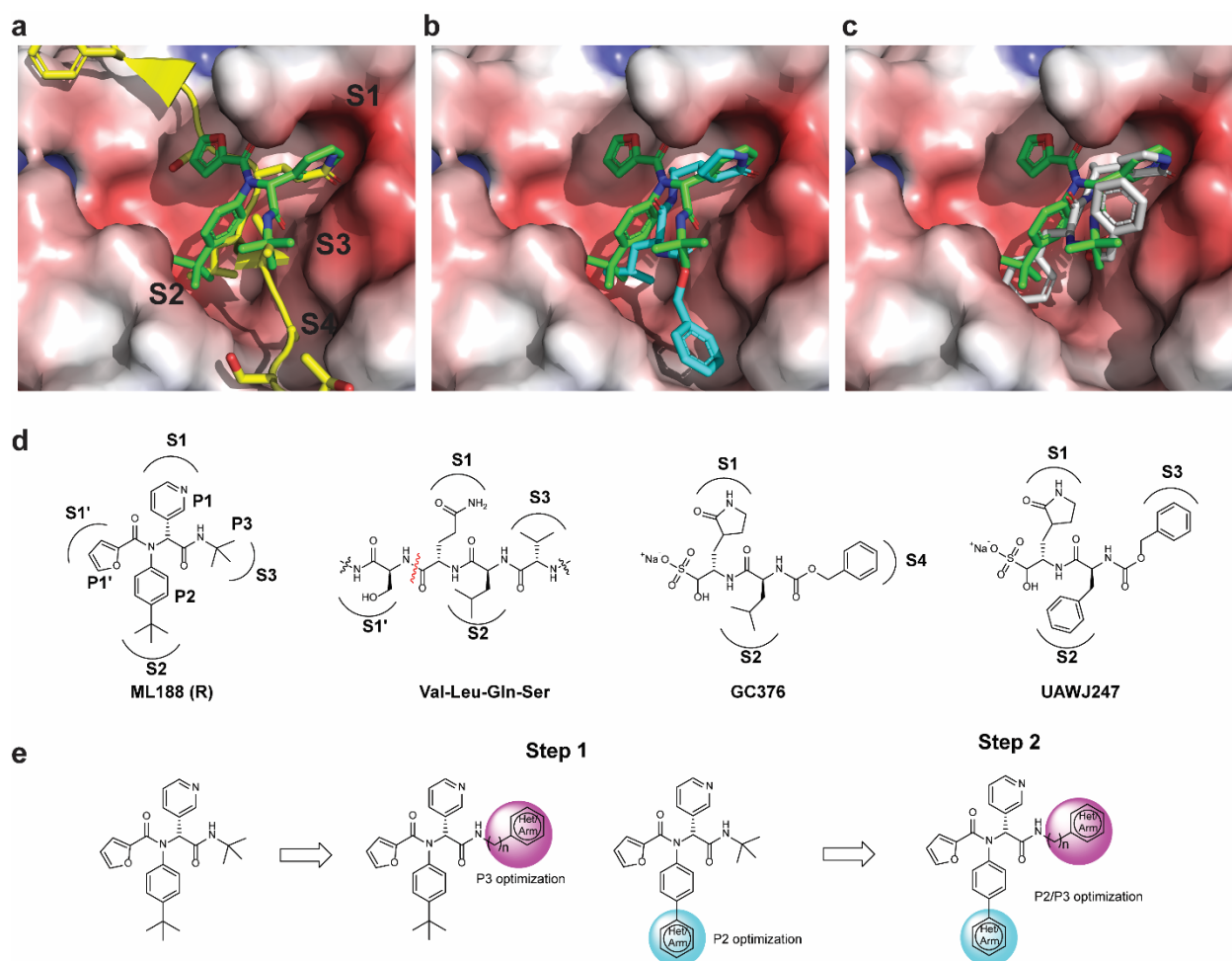
Among the non-canonical SARS-CoV-2 M<sup>pro</sup> inhibitors we recently discovered, calpain inhibitor XII has an unexpected binding mode showing an inverted conformation in the active site<sup>8</sup> (Fig. 1a). Instead of projecting the norvaline and leucine side chains into the S1 and S2 pockets as one would expect from its chemical structure, the pyridinyl substitution snugly fits in the S1 pocket and forms a hydrogen bond with the H163 imidazole (Fig. 1a). This hydrogen bond is essential, as replacing the pyridine with benzene led to an analog UAWJ257 with a significant loss of enzymatic inhibition<sup>8</sup>. Examining the X-ray crystal structures of SARS-CoV and SARS-CoV-2 M<sup>pro</sup> in the PDB database revealed another compound ML188 (R)<sup>12</sup>, which shares a similar binding mode with calpain inhibitor XII. ML188 (R) is a non-covalent SARS-CoV M<sup>pro</sup> inhibitor derived from a high-throughput screening hit<sup>12</sup>. The pyridinyl from ML188 (R) similarly fits in the S1 pocket and forms a hydrogen bond with the H163 side chain imidazole (Fig. 1b). In addition, the furyl oxygen and its amide oxygen both form a hydrogen bond with the G143 main chain amide amine. ML188 (R) was reported to inhibit the SARS-CoV M<sup>pro</sup> with an IC<sub>50</sub> value of 1.5 ± 0.3 μM and the SARS-CoV viral replication in Vero E6 cells with an EC<sub>50</sub> value of 12.9 μM<sup>12</sup>. Several follow up studies have been conducted to optimize the enzymatic inhibition and cellular antiviral activity of this series of compounds, however, no significant improvement has been made<sup>13,14</sup>.



**Fig. 1. Structure of M<sup>pro</sup> with inhibitors.** **a** X-ray crystal structure of SARS-CoV-2 M<sup>pro</sup> in complex with calpain inhibitor XII (PDB: 6XFN). **b** X-ray crystal structure of SARS-CoV M<sup>pro</sup> in complex with ML188 (R) (PDB: 3V3M). Hydrogen bonds are shown in dashed lines.

The similar binding mode of ML188 (R) with calpain inhibitor XII, coupled with the convenient synthesis through the one pot Ugi-4CR, inspired us to design non-covalent SARS-CoV-2 M<sup>pro</sup> inhibitors based on the ML188 (R) scaffold. Specifically, we leverage our understanding of the M<sup>pro</sup> inhibition mechanism based on the X-ray co-crystal structures of SARS-CoV-2 M<sup>pro</sup> with multiple inhibitors to guide the lead optimization (Figs. 2a-d)<sup>7,8</sup>. Overlaying the X-ray crystal structures of SARS-CoV M<sup>pro</sup> + ML188 (R) (PDB: 3V3M) and the SARS-CoV M<sup>pro</sup> H41A mutant + the peptide substrate (PDB: 2Q6G) revealed that the furyl, 4-*tert*-butylphenyl, pyridinyl, and *tert*-butyl of ML188 (R) fit in the S1', S2, S1, and S3 pockets respectively (Figs. 2a and 2d). Therefore, the furyl, 4-*tert*-butylphenyl, pyridinyl, and *tert*-butyl substitutions in ML188 (R) were defined as P1', P2, P1, and P3, respectively. Next, overlaying the structure of SARS-CoV M<sup>pro</sup> + ML188 (R) (PDB: 3V3M) and SARS-CoV-2 M<sup>pro</sup> + GC376 (PDB: 6WTT) suggested that the *tert*-butyl at the P3 substitution of ML188 (R) can be extended

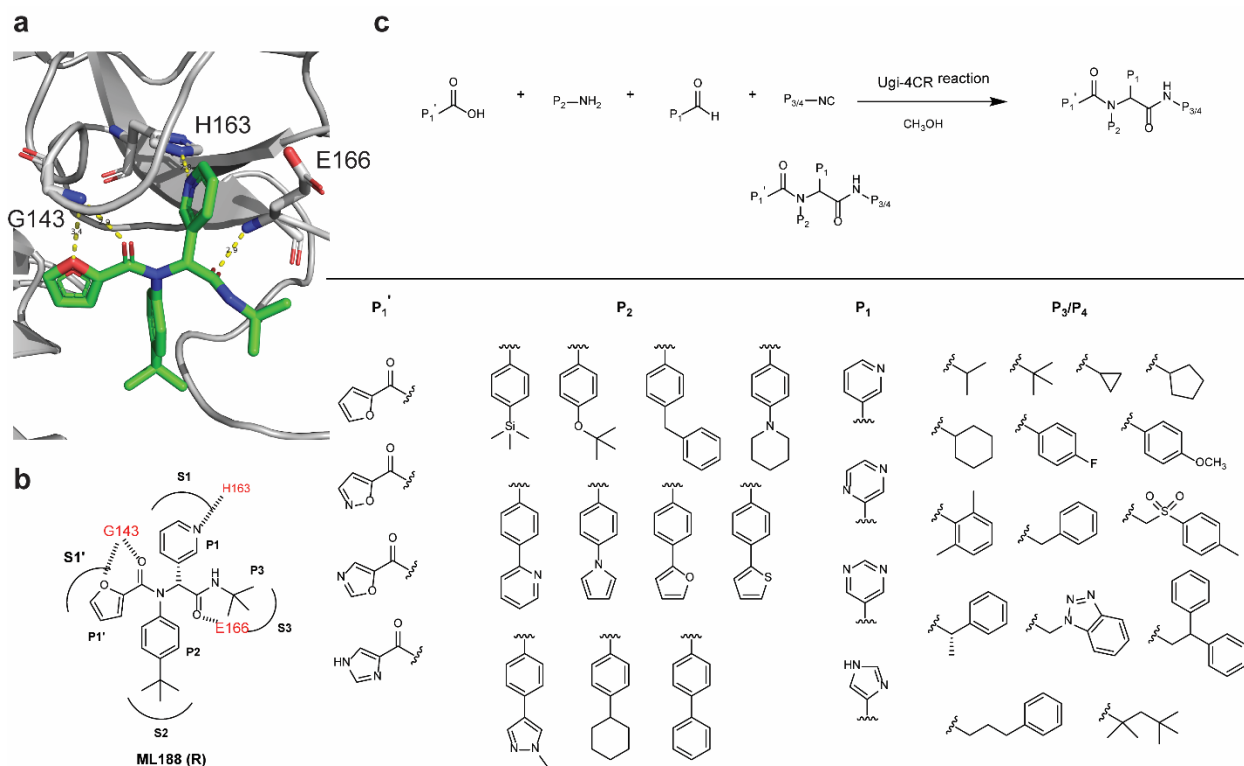
to fit in the S4 pocket (Figs. 2b and 2d). Previous structure-activity relationship studies of GC376 indicate that P4 substitution is important, while P3 substitution does not contribute significantly to the binding affinity, as it is solvent exposed<sup>3,8,9,15</sup>. Similarly, the overlaying structures of SARS-CoV M<sup>Pro</sup> + ML188 (R) (PDB: 3V3M) and SARS-CoV-2 M<sup>Pro</sup> + UAWJ247 (PDB: 6XBH) suggested that the 4-*tert*-butyl at the P2 substitution of ML188 (R) can be replaced by phenyl to occupy the extra space in the S2 pocket (Figs. 2c and 2d). Overall, the design mainly focuses on extending the P2 and P3 substitutions of ML188 (R) to achieve optimal shape complementarity with the SARS-CoV-2 M<sup>Pro</sup> (Fig. 2e). In practice, we adopted a stepwise optimization procedure in which the P3 and P2 substitutions were optimized individually in step 1, and then the optimal P2/P3 substitutions were combined in step 2 (Fig. 2e).



**Fig. 2. Design rationale for the non-covalent SARS-CoV-2 M<sup>pro</sup> inhibitors.** **a** Overlaying X-ray crystal structures of SARS-CoV M<sup>pro</sup> + ML188 (R) (PDB: 3V3M, green) and SARS-CoV M<sup>pro</sup> H41A mutant + peptide substrate (PDB: 2Q6G, yellow with backbone shown as ribbon representation). **b** Overlaying X-ray crystal structures of SARS-CoV M<sup>pro</sup> + ML188 (R) (PDB: 3V3M) and SARS-CoV-2 M<sup>pro</sup> + GC376 (PDB: 6WTT). **c** Overlaying X-ray crystal structures of SARS-CoV M<sup>pro</sup> + ML188 (R) (PDB: 3V3M) and SARS-CoV-2 M<sup>pro</sup> + UAWJ247 (PDB: 6XBH). **d** Chemical structures of ML188 (R), peptide substrate VLQS, GC376, and UAWJ247. **f** Stepwise optimization of ML188 (R) towards potent non-covalent SARS-CoV-2 M<sup>pro</sup> inhibitor.

Guided by the design rationale elucidated above, a focused library of ML188 analogs were designed and synthesized (Fig. 3). As the P1' furyl and P1 pyridinyl both form a critical hydrogen bond with the M<sup>pro</sup> (Figs. 3a-b), the P1' and P1 substitutions were kept with minimal variations (Fig. 3c). All designed compounds were synthesized using the one pot Ugi four-component reaction and tested as enantiomer/diastereomer mixtures (Fig. 3c). To circumvent the need of relying on expansive chiral HPLC column for the separation of enantiomers, we strategically introduced the chiral isocyanide so that the diastereomer product mixture can be separated by convenient silica gel column or reverse phase HPLC column purification.





**Fig. 3. Design and synthesis of a focused library of non-covalent SARS-CoV-2 M<sup>pro</sup> inhibitors.** **a** X-ray crystal structure of SARS-CoV M<sup>pro</sup> + ML188 (R) (PDB: 3V3M). **b** Binding interactions of ML188 (R) with SARS-CoV M<sup>pro</sup>. **c** Synthesis of ML188 analogs using the Ugi four-component reaction.

### Structure-activity relationship studies of non-covalent SARS-CoV-2 M<sup>pro</sup> inhibitors

In total, 39 compounds were synthesized (Figs. 4a-4e) and all compounds were initially tested as a mixture of enantiomers or diastereomers in the FRET-based enzymatic assay against SARS-CoV-2 M<sup>pro</sup> at 20  $\mu$ M (Fig. 4f). Compounds showing more than 50% inhibition at 20  $\mu$ M were further titrated to determine the IC<sub>50</sub> values. Next, compounds with IC<sub>50</sub> values lower than 5  $\mu$ M were selected for cellular cytotoxicity profiling in Vero E6 cells, the cell line which was used for the SARS-CoV-2 antiviral assay. The purpose was to prioritize lead candidates for the *in vitro* cellular antiviral assay with infectious SARS-CoV-2. Compounds with

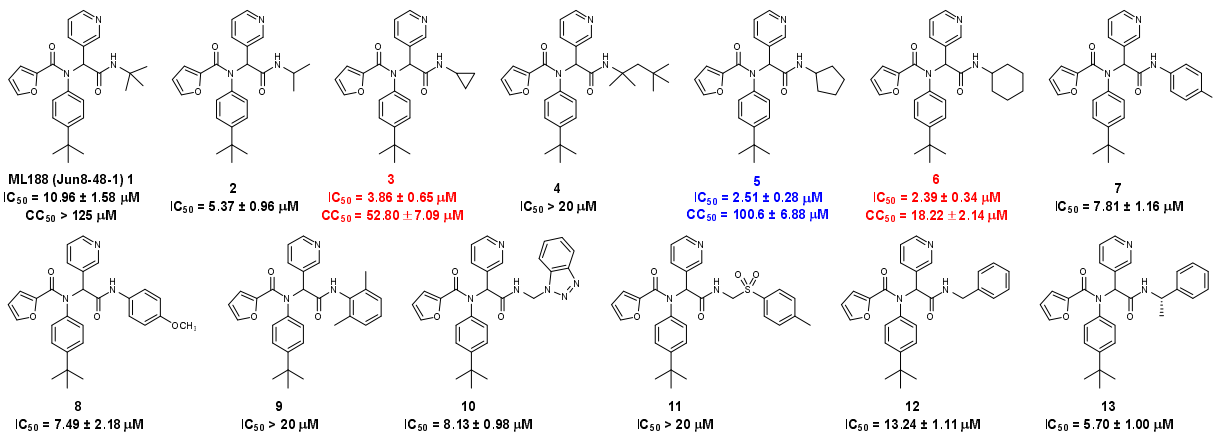
potent enzymatic inhibition ( $IC_{50} < 5 \mu M$ ) but moderate to high cellular cytotoxicity ( $CC_{50} < 100 \mu M$ ) were labeled in red. Compounds with both potent enzymatic inhibition ( $IC_{50} < 5 \mu M$ ) and low cellular cytotoxicity ( $CC_{50} > 100 \mu M$ ) were labeled in blue (Figs. 4a-4e). As shown in Fig. 4f, the majority of the designed compounds showed more than 50% inhibition when tested at  $20 \mu M$ . Specifically, Fig. 4a lists compounds with P4 variations. As a reference, ML188 (**1**) (racemic mixture) inhibits SARS-CoV-2  $M^{pro}$  with an  $IC_{50}$  value of  $10.96 \pm 1.58 \mu M$ . It was found that compounds **2**, **3**, **5**, **6**, **7**, **8**, **10**, and **13** had improved enzymatic inhibition compared to ML188 (**1**). These results suggest that: a) isopropyl (**2**), cyclopropyl (**3**), cyclopentyl (**5**), cyclohexyl (**6**), and phenyl (**7** and **8**) are the more favorable substitutions at the P3 position than *tert*-butyl; b) compound **13** with the (*S*)- $\alpha$ -methylbenzyl substitution at the P3 position had improved potency, which suggests that extending the substitutions to the S4 pocket might improve the enzymatic inhibition (Fig. 2b). Given the advantage of convenient separation of diastereomers over enantiomers, we therefore decided to fix the P3/P4 substitution as  $\alpha$ -methylbenzyl substitution during the P2 optimization (Fig. 4b). All compounds in Fig. 4b were designed to have extended substitutions at the 4-position of benzyl to occupy the extra space in the S2 pocket (Fig. 2c). Consistent with the design hypothesis, several compounds including **14**, **17**, **18**, **19**, **20**, **21**, and **23** had significantly improved enzymatic inhibition ( $IC_{50} < 3 \mu M$ ) compared to compound **13**. Replacing the *tert*-butyl in compound **13** with the bulkier trimethylsilyl led to compound **14** with a 2.9-fold increase in  $M^{pro}$  inhibition. Cyclohexyl (**17**), thienyl (**19**), pyrrolyl (**20**), pyridinyl (**21**), and phenyl (**23**) were found to be the most favorable substitutions at the S2 pocket. Compound **16** with piperidyl substitution had similar potency as compound **13**, while compound **15** with *O-tert* butyl was less active. Further extending the substitution to benzyl led to compound **22** that was inactive, suggesting biphenyl might be the largest substitution that can be accommodated at the S2 pocket.

The P1' and P1 substitutions (Figs. 4c and 4d) were chosen to retain the critical hydrogen bonds in ML188 (Fig. 3a). It was found that imidazole (**24**) was tolerated at the P1' position ( $IC_{50} = 0.96 \pm 0.09 \mu\text{M}$ ), followed by isoxazole (**25**) ( $IC_{50} = 2.47 \pm 0.27 \mu\text{M}$ ), and oxazole (**26**) ( $IC_{50} = 4.97 \pm 0.78 \mu\text{M}$ ). Pyrazine (**27**) was tolerated at the P1 position ( $IC_{50} = 4.93 \pm 0.79 \mu\text{M}$ ); however, pyrimidine (**29**) and imidazole (**30**) were not preferred ( $IC_{50} > 20 \mu\text{M}$ ).

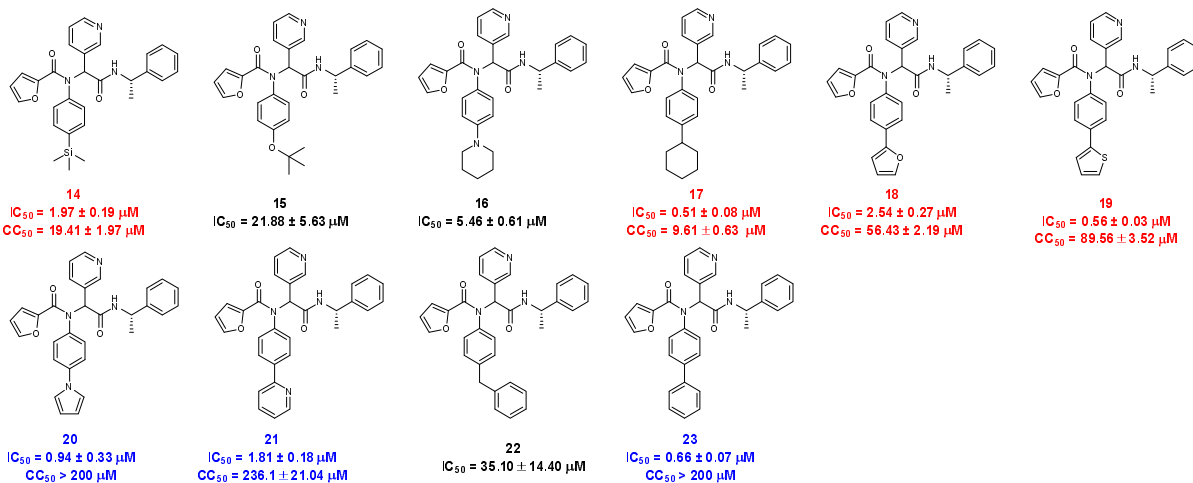
Next, the above identified favorable P1', P2, P1, and P3/P4 substitutions were combined and the designed compounds were shown in Fig. 4e. Compounds **36**, **37**, and **38** were the most potent leads with  $IC_{50}$  values of  $0.81 \pm 0.24$ ,  $0.67 \pm 0.15$ , and  $0.28 \pm 0.07 \mu\text{M}$ , respectively. Compound **39** and **40** were also highly active with  $IC_{50}$  values of  $1.48 \pm 0.56$  and  $1.05 \pm 0.09 \mu\text{M}$ , respectively.

Among the active compounds with  $IC_{50}$  value lower than  $5 \mu\text{M}$ , compounds **3**, **6**, **14**, **17**, **18**, **19**, **26**, **34**, and **36** had moderate to high cellular cytotoxicity in Vero E6 cells (Figs. 4a-4e red), while compounds **5**, **20**, **21**, **23**, **24**, **25**, **27**, **31**, **32**, **37**, **38**, **39**, and **40** were well tolerated and the  $CC_{50}$  values were greater than  $100 \mu\text{M}$ .

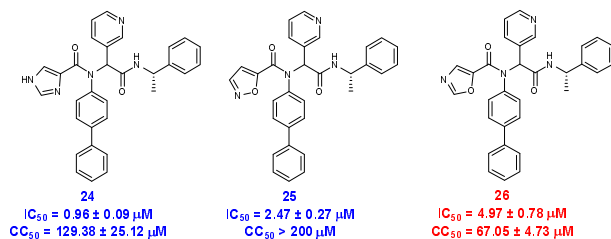
**a**



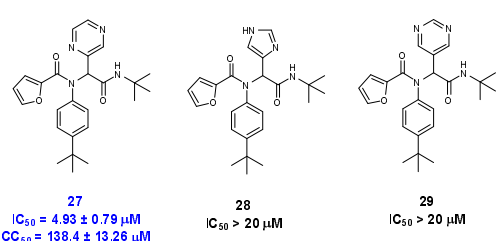
**b**



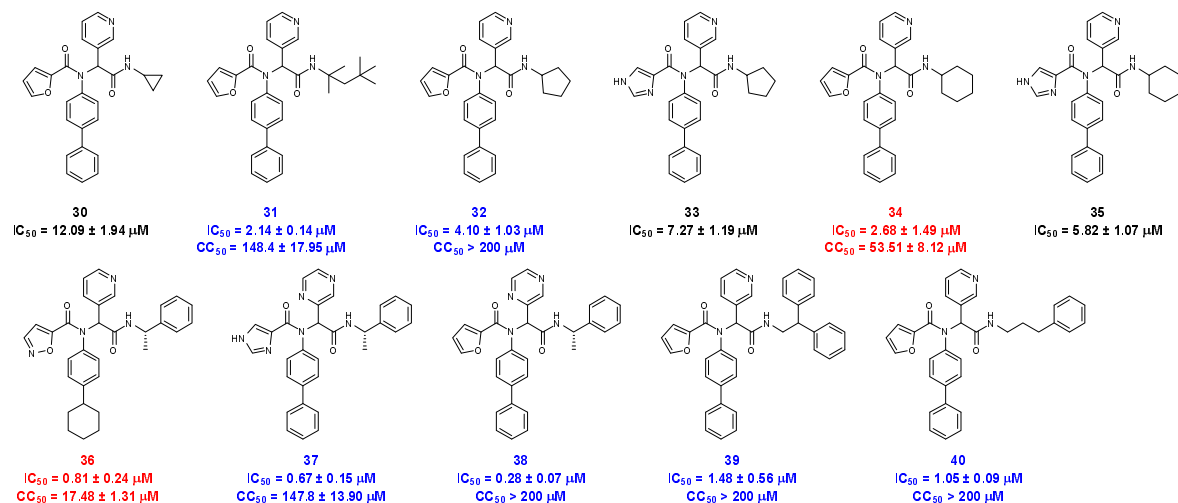
**c**

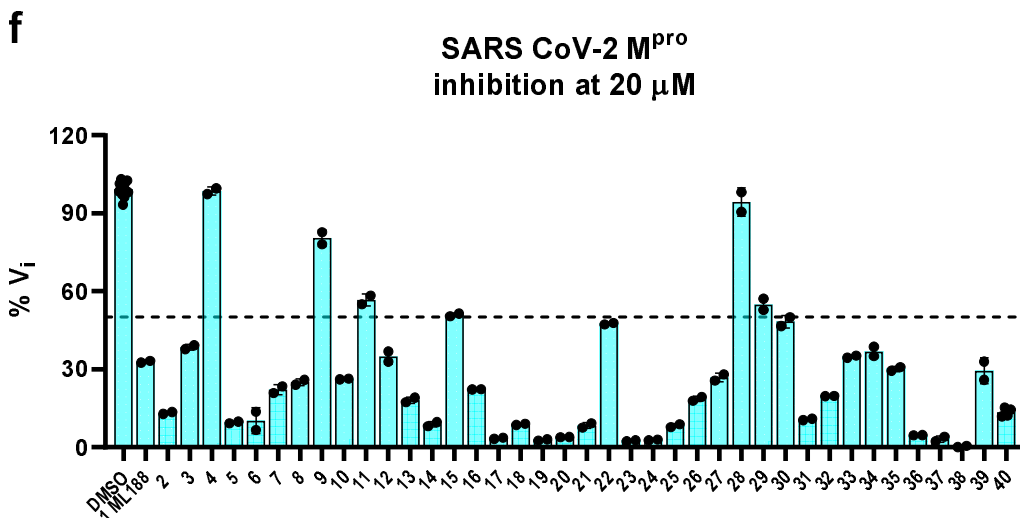


**d**



**e**





**Fig. 4. Structures of non-covalent SARS-CoV-2 M<sup>pro</sup> inhibitors and the enzymatic inhibition against SARS-CoV-2 M<sup>pro</sup>. a** Analogs with P3/P4 modifications. **b** Analogs with P2 modifications. **c** Analogs with P1' modifications. **d** Analogs with P1 modifications. **e** Analogs with combined P1', P1, P2, and P3/P4 modifications. Compounds with potent enzymatic inhibition ( $IC_{50} < 5 \mu M$ ) but moderate to high cellular cytotoxicity ( $CC_{50} < 100 \mu M$ ) were labeled in red. Compounds with both potent enzymatic inhibition ( $IC_{50} < 5 \mu M$ ) and low cellular cytotoxicity ( $CC_{50} > 100 \mu M$ ) were labeled in blue. **f** Percentage enzymatic inhibition of SARS-CoV-2 M<sup>pro</sup> by the designed compounds at 20  $\mu M$  compound concentration.

Next, compounds with potent enzymatic inhibition ( $IC_{50} \leq 1 \mu M$ ) and low cellular cytotoxicity ( $CC_{50} > 100 \mu M$ ) were prioritized for the cellular antiviral assay with infectious SARS-CoV-2 in Vero E6 cells using the immunofluorescence assay as the primary assay (Table 1). ML188 (**1**) was included as a control. It was found that ML188 (**1**) was inactive in the antiviral assay ( $EC_{50} > 20 \mu M$ ), probably due to its incomplete inhibition of the M<sup>pro</sup> in the cellular content. Gratifyingly, compounds **20**, **23**, **37**, **38**, and **40** all had potent cellular antiviral activity with  $EC_{50}$  values

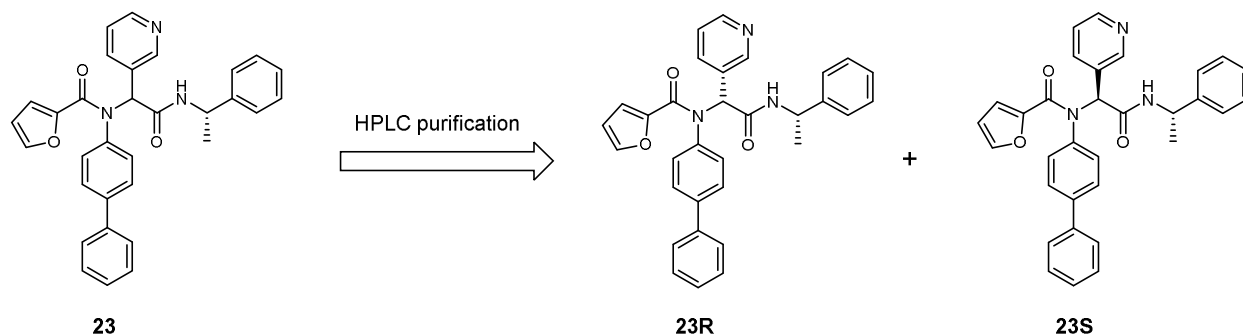
ranging from 0.82 to 4.54  $\mu\text{M}$ . Compound **24** was less active ( $\text{EC}_{50} = 13.06 \pm 2.30 \mu\text{M}$ ), possibly due to the poor cellular membrane permeability.

**Table 1 Antiviral activity and selectivity index of non-covalent SARS-CoV-2  $\text{M}^{\text{pro}}$  inhibitors.**

**(Selection criteria  $\text{IC}_{50} < 1 \mu\text{M}$ ,  $\text{CC}_{50} > 100 \mu\text{M}$ ).**

Compound ID	SARS-CoV-2 $\text{M}^{\text{pro}}$ $\text{IC}_{50}$ ( $\mu\text{M}$ )	Vero E6 Cytotoxicity $\text{CC}_{50}$ ( $\mu\text{M}$ )	SARS-CoV-2 antiviral assay $\text{EC}_{50}$ ( $\mu\text{M}$ )	Selectivity index SI $\text{CC}_{50}/\text{EC}_{50}$
<b>1 ML188</b>	$10.96 \pm 1.58$	$> 125$	$>20$	N.A.
<b>20</b>	$0.94 \pm 0.33$	$> 200$	$1.73 \pm 0.19$	$> 115.6$
<b>23</b>	$0.66 \pm 0.07$	$> 200$	$1.61 \pm 0.29$	$> 124.2$
<b>24</b>	$0.96 \pm 0.09$	$129.38 \pm 25.12$	$13.06 \pm 2.30$	9.9
<b>37</b>	$0.67 \pm 0.15$	$147.8 \pm 13.9$	$4.54 \pm 0.69$	32.6
<b>38</b>	$0.28 \pm 0.07$	$> 200$	$0.82 \pm 0.56$	$> 243.9$
<b>40</b>	$1.05 \pm 0.09$	$> 200$	$2.04 \pm 1.08$	$> 98.0$

Given the potent antiviral activity and a high selectivity index of these potent lead compounds, we then selected compound **23** for further characterization. The two diastereomers of **23** were separated by reverse phase HPLC (Fig. 5). Both diastereomers were tested in the FRET-based enzymatic assay. GC376 was included as a positive control. It was found that **23R** is the active diastereomer with an  $\text{IC}_{50}$  value of  $0.31 \pm 0.04 \mu\text{M}$ , while the **23S** diastereomer was more than 16-fold less active ( $\text{IC}_{50} = 5.61 \pm 0.71 \mu\text{M}$ ) (Table 2). The stereochemistry of **23R** was determined by the co-crystal structure with SARS-CoV-2  $\text{M}^{\text{pro}}$  as described in the following section. Compared with the parent compound ML188 (**1**), the optimized lead **23R** had more than a 35-fold increase in enzymatic inhibition against SARS-CoV-2  $\text{M}^{\text{pro}}$ . Compound **23R** also showed comparable potency against SARS-CoV  $\text{M}^{\text{pro}}$  with an  $\text{IC}_{50}$  value of  $0.27 \pm 0.03 \mu\text{M}$ . Neither ML188 (**1**) nor **23R** inhibited the SARS-CoV-2 papain-like protease ( $\text{PL}^{\text{pro}}$ ) ( $\text{IC}_{50} > 20 \mu\text{M}$ ) (Table 2), suggesting the inhibition of SARS-CoV-2  $\text{M}^{\text{pro}}$  by **23R** is specific.



**Fig. 5. Separation of the two diastereomers of compound 23.** The absolute stereochemistry of compound **23R** was determined in the co-crystal structure of this diastereomer with SARS-CoV-2 M<sup>pro</sup> (PDB: 7KX5).

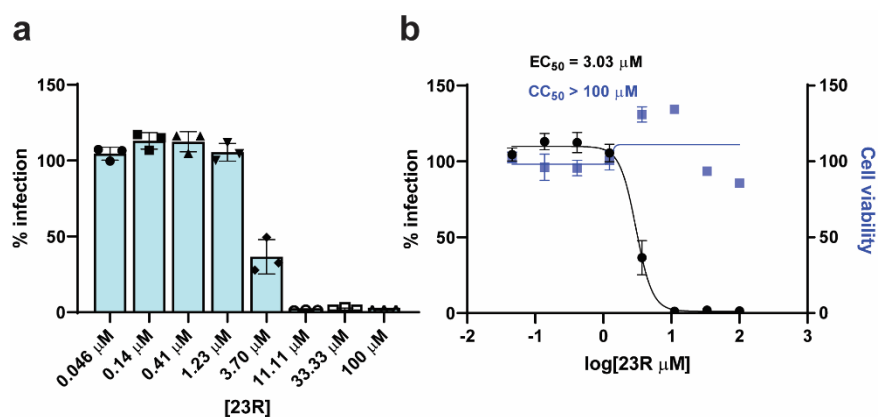
**Table 2 Enzymatic inhibition, antiviral activity and selectivity index of 23R.**

Compound ID	SARS-CoV-2 M <sup>pro</sup> IC <sub>50</sub> (μM)	SARS-CoV M <sup>pro</sup> IC <sub>50</sub> (μM)	SARS-CoV-2 PL <sup>pro</sup> IC <sub>50</sub> (μM)	Vero E6 Cytotoxicity CC <sub>50</sub> (μM)	SARS-CoV-2 antiviral assay EC <sub>50</sub> (μM)	Selectivity index SI CC <sub>50</sub> /EC <sub>50</sub>
<b>ML188 (1)</b>	10.96 ± 1.58	11.23 ± 1.61	> 60	> 125	>20	N.A.
<b>GC376</b>	0.033 ± 0.003	0.035 ± 0.002	> 60	> 125	1.51 ± 0.27	> 82.8
<b>23R</b>	0.31 ± 0.04	0.27 ± 0.03	> 60	> 100	1.27 ± 0.09	> 78.7
<b>23S</b>	5.61 ± 0.71	N.T.	N. T.	N.T.	N.T.	N.A.

N.T. = not tested. N.A. = not applicable.

Next, the antiviral activity of **23R** was tested against SARS-CoV-2 (USA-WA1/2020 isolate) in Vero E6 cells using the immunofluorescence assay. ML188 (**1**) and GC376 were included as controls. It was found that compound **23R** had an EC<sub>50</sub> value of 1.27 μM (Table 2), which was similar to the antiviral potency of the covalent inhibitor GC376 (EC<sub>50</sub> = 1.51 μM). Compound **23R** was also not cytotoxic to Vero E6 cells at up to 100 μM. In contrast, the parent compound ML188 (**1**) had no detectable antiviral activity when tested at up to 20 μM. To further confirm the

antiviral activity of compound **23R**, we performed secondary antiviral assay in the human lung epithelial Calu-3 cell line, which endogenously expresses TMPRSS2 and is widely used as a physiological relevant cell line for SARS-CoV-2 infection. It was found that compound **23R** inhibited SARS-CoV-2 (USA-WA1/2020 isolate) replication in Calu-3 cells with an  $EC_{50}$  value of 3.03  $\mu\text{M}$  and it was not cytotoxic at up to 100  $\mu\text{M}$  (Fig. 6). The 2.4-fold difference in antiviral potency between the Vero E6 and Calu-3 cell lines might due to differences in cell membrane permeability or metabolism.



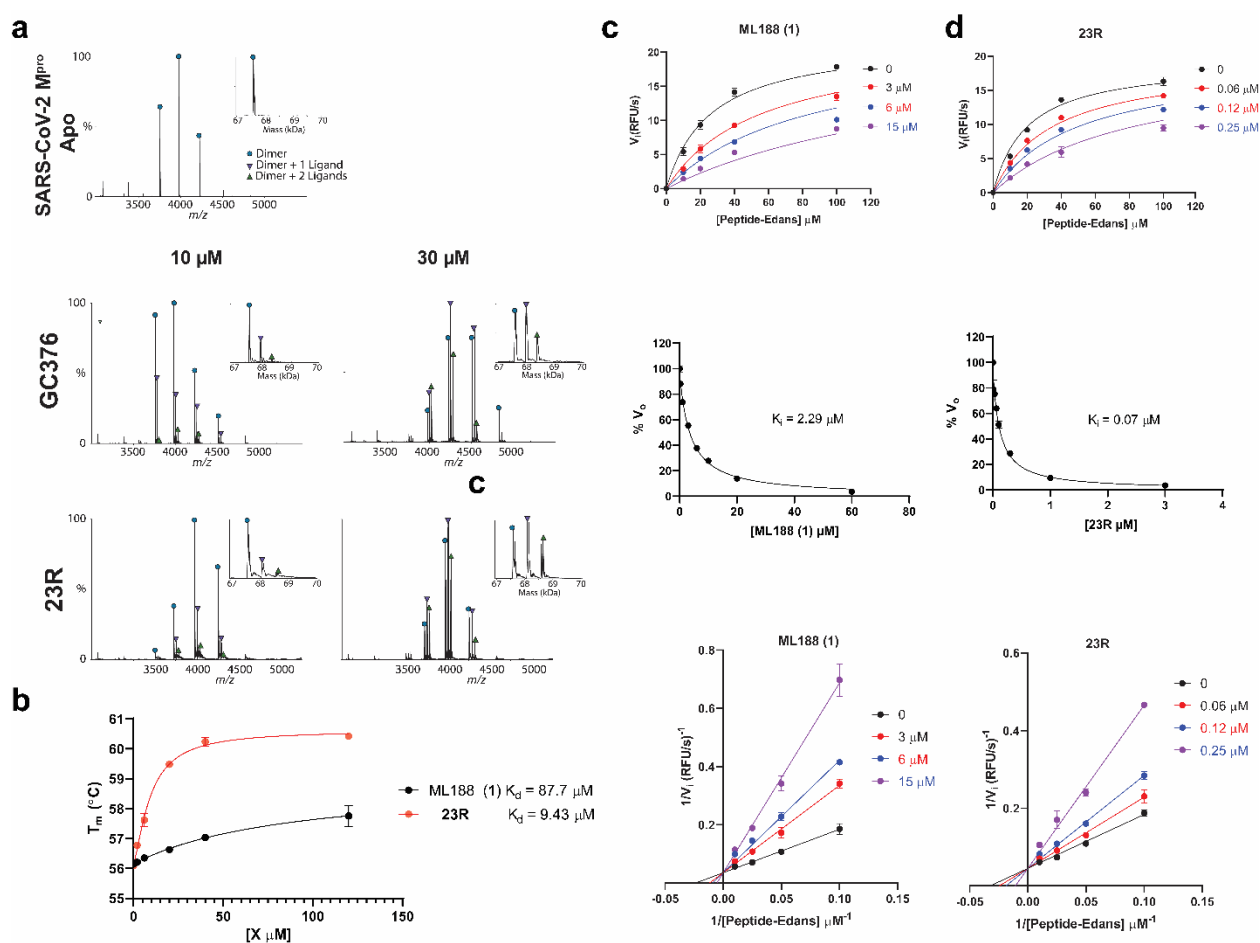
**Fig. 6. Antiviral activity of 23R against SARS-CoV-2 in Calu-3 cells.** **a** Raw data of the percentage of immunofluorescence positive cells with different concentrations of **23R**. **b** Antiviral potency and cytotoxicity plots.

### Mechanism of action of 23R in inhibiting SARS-CoV-2 $M^{\text{pro}}$

The mechanism of action of **23R** was characterized using the native mass spectrometry binding assay, the thermal shift binding assay, and the enzymatic kinetic studies (Fig. 7). In the native mass spectrometry binding assay, compound **23R** showed dose-dependent binding to SARS-CoV-2  $M^{\text{pro}}$ , similar to the positive control GC376, with a binding stoichiometry of one drug per monomer (Fig. 7a). Similarly, compound **23R** showed dose-dependent stabilization of



the SARS-CoV-2 M<sup>pro</sup> in the thermal shift binding assay with an apparent  $K_d$  value of 9.43  $\mu\text{M}$ , a 9.3-fold increase compared to ML188 (1) (Fig. 7b). In the enzymatic kinetic studies, **23R** was shown to be a reversible inhibitor with a  $K_i$  value of 0.07  $\mu\text{M}$  (Figs. 7c and 7d top and middle panels). In comparison, the  $K_i$  for the parent compound ML188 (1) is 2.29  $\mu\text{M}$ . The Lineweaver-Burk or double-reciprocal plot with different compound concentrations yielded an intercept at the Y-axis, suggesting that **23R** is a competitive inhibitor similar to ML188 (1) (Figs. 7c and 7d bottom panel).



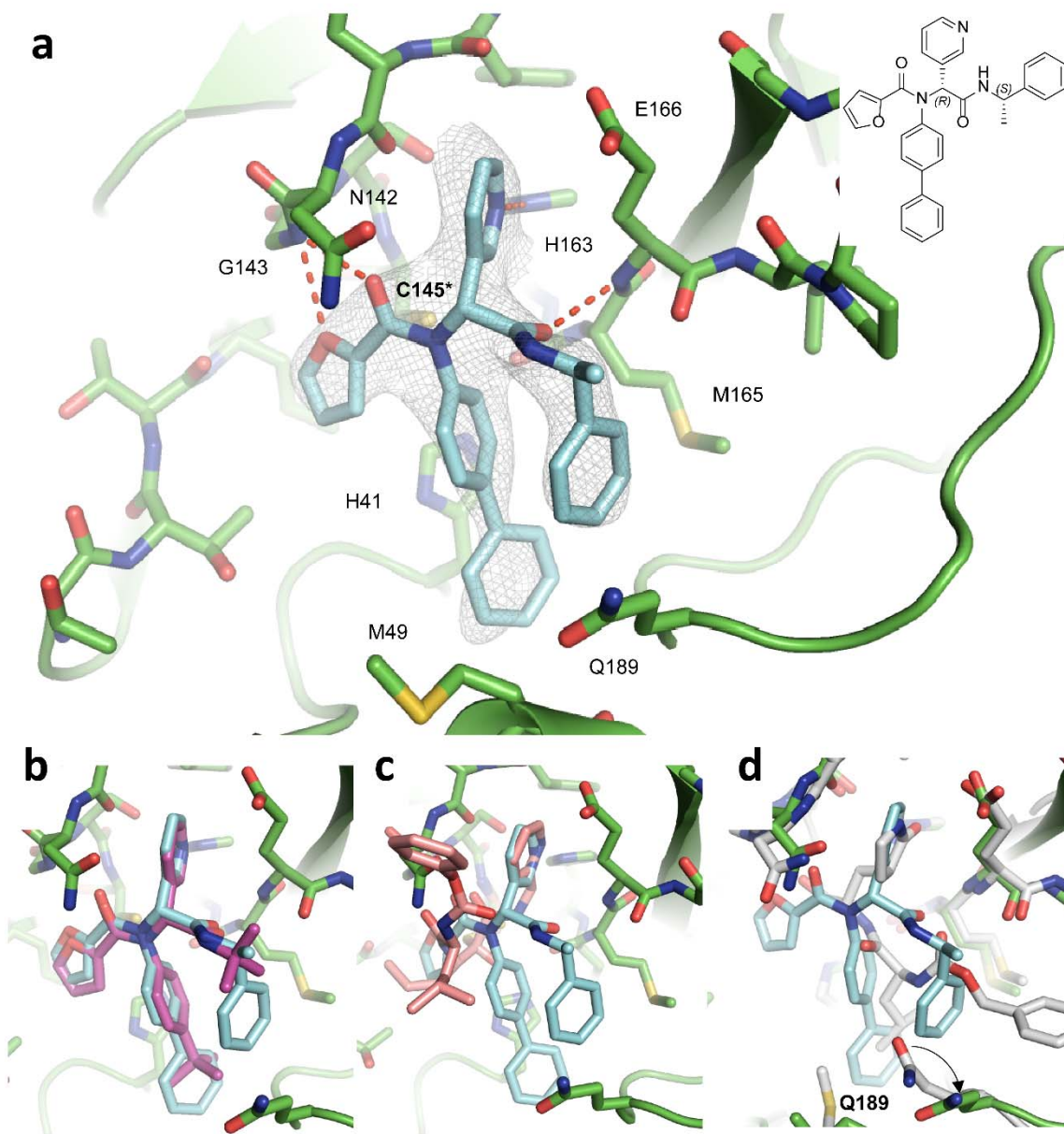
**Fig. 7. Characterization of the mechanism of action of 23R to SARS-CoV-2 M<sup>pro</sup> using native mass-spectrometry, thermal shift assay, and enzyme kinetic studies.** a Binding of **23R** to SARS-CoV-2 M<sup>pro</sup> using native mass spectrometry. Native mass spectra with the inset deconvolved spectra revealing ligand binding with 10  $\mu\text{M}$  or 30  $\mu\text{M}$  GC-376 added (middle

panel), and 10  $\mu\text{M}$  and 30  $\mu\text{M}$  **23R** (bottom panel) with 4 mM DTT added. The peaks are annotated with the blue circle as the dimer, green down triangle as the dimer with one ligand bound, and the purple up triangle as the dimer with two ligands bound. **b** Dose-dependent melting temperature ( $T_m$ ) shift in thermal shift assay. 3  $\mu\text{M}$  SARS-CoV-2  $M^{\text{pro}}$  protein was incubated with various concentrations of ML188 or **23R** in the presence of 4 mM DTT. Measured  $T_m$  was plotted against compound concentration with one-site binding function in Prism 8. **c, d** Enzymatic kinetic assay with ML188 and compound **23R**. Kinetic parameters in the presence of various concentrations of ML188 or **23R** were globally fit with Michaelis-Menten function in prism 8 (top panels); double reciprocal plots were shown in the bottom panels. The middle panels show the Morrison plots of compound ML188 and **23R** with 20  $\mu\text{M}$  FRET substrate was used.

### **X-ray crystal structure of SARS-CoV-2 $M^{\text{pro}}$ with 23R**

Using X-ray crystallography, we successfully determined the binding pose of **23R** with SARS-CoV-2  $M^{\text{pro}}$  at 2.6 Å resolution (Fig. 8a). Electron density reveals the body of **23R** extends throughout the substrate binding channel, with side chains occupying the S1', S1, S2, and S3 sub-pockets. The binding pose is similar to the previously solved structure of SARS-CoV  $M^{\text{pro}}$  with ML188 (R) (PDB: 3V3M)<sup>12</sup>, consistent with the similarities between the two compounds and between the two proteins (Fig. 8b). The furyl moiety of **23R** binds to a portion of the P1' site, which normally accommodates small hydrophobic residues. While the furylamide carbonyl group of **23R** does not insert into the oxyanion hole, it does form a bifurcated hydrogen bond with the apical residue of this oxyanion hole, Gly143. However, the furan ring oxygen is likely a weaker hydrogen bond acceptor than the amide oxygen and it lies outside of the plane of Gly143's amide NH. Directly attached to the furylamide moiety is a P2 biphenyl group and a P1 pyridinyl ring. The P2 biphenyl group projects directly into the S2 pocket, which prefers

hydrophobic residues such as leucine and phenylalanine. As expected, the P1 pyridinyl ring occupies the S1 pocket, which is known for its strict preference for glutamine. While most M<sup>pro</sup> inhibitors bear a pyrrolidinone glutamine mimetic at the P1 position, we determined that more hydrophobic residues can also bind to the S1 site, and that hydrogen bond formation with His163 is critical for inhibition<sup>8</sup>. In this instance, the pyridinyl ring of **23R** is nearly superimposable with the same moiety from calpain inhibitor XII (Fig. 8c) forming a close (2.9 Å) hydrogen bond with His163. An amide bond connecting the pyridinyl ring to the  $\alpha$ -methylbenzyl moiety forms a hydrogen bond with the main chain of Glu166. The benzyl ring of the  $\alpha$ -methylbenzyl moiety is partially positioned both in the S2 and S3 pockets, a novel binding pose that has not been observed with existing M<sup>pro</sup> inhibitors. Normally, a substituent at this position would be expected to flip away from the enzyme core towards the solvent-exposed S3 pocket, which explains why P3 substitutions have little to no influence on the enzymatic inhibition<sup>4</sup>. However, the hydrophobic nature of the benzyl ring in **23R** causes it to project towards the core near the S2 pocket, forcing Gln189 to rotate outwards (Fig. 8d). This conformation is reinforced by pi-stacking interactions with the first phenyl of the biphenyl substituent. Notably, the binding pose of **23R** features continuous intramolecular pi-stacking, where the phenyl is sandwiched by furan and benzyl groups, potentially contributing to its potent inhibition of M<sup>pro</sup>. Meanwhile, the S4 pocket remains largely unoccupied by **23R**, leaving room for further improvement. In summary, the X-ray crystal structure of SARS-CoV-2 M<sup>pro</sup> in complex with **23R** revealed two interesting structural features: 1) The P2 biphenyl is probably the largest substitution that can be accommodated in the S2 pocket, which is consistent with our design hypothesis. 2) The benzyl group from the terminal  $\alpha$ -methylbenzyl fits in a previously unexplored binding pocket located in between the S2 and S4 pockets. As such, the benzyl group faces towards the core of the enzyme instead of solvent-exposed as seen with other existing M<sup>pro</sup> inhibitors. Although this is unexpected from the design perspective, this novel binding mode suggests that the new binding pocket in between S2 and S4 that can be explored for inhibitor design.



**Fig. 8. X-ray crystal structure of SARS-CoV-2 M<sup>pro</sup> in complex with 23R.** **a** X-ray crystal structure of SARS-CoV-2 M<sup>pro</sup> in complex with 23R (PDB: 7KX5). **b** Overlaying structures of SARS-CoV-2 M<sup>pro</sup> + 23R (PDB: 7KX5) and SARS-CoV M<sup>pro</sup> + ML188 (R) (PDB: 3V3M). **c** Overlaying structures of SARS-CoV-2 M<sup>pro</sup> + 23R (PDB: 7KX5) and SARS-CoV-2 M<sup>pro</sup> + calpain inhibitor XII (PDB: 6XFN). **d** Overlaying structures of SARS-CoV-2 M<sup>pro</sup> + 23R (PDB: 7KX5) and SARS-CoV-2 M<sup>pro</sup> + GC376 (PDB: 6WTT).

## Discussion

Given the social and economic impact of the ongoing COVID-19 pandemic, the need for effective therapeutics is apparent. Researchers around the globe are racing to come up with intervention strategies. The viral M<sup>pro</sup> is a high profile antiviral drug target and several M<sup>pro</sup> inhibitors are now in animal model studies and human clinical trial<sup>6</sup>. Among the known M<sup>pro</sup> inhibitors, the majority of them are covalent inhibitors such as GC376 analogs that contain a pyrrolidone in the P1 position as a glutamine mimetic. Several structurally distinct compounds including ebselen, disulfiram, carmofur, PX-12, tideglusib, and shiknonin were claimed as M<sup>pro</sup> inhibitors<sup>16,17</sup>, but were later invalidated as promiscuous non-specific cysteine protease inhibitors<sup>18,19</sup>. In addition, non-covalent inhibitors such as ML188 (R) were developed and validated as SARS-CoV M<sup>pro</sup> inhibitors<sup>5,12</sup>. Several follow up studies have been conducted to optimize the enzymatic potency of this series of compounds against SARS-CoV M<sup>pro</sup> and the SARS-CoV-2 M<sup>pro</sup>. However, no significant improvement has been made and ML188 (R) remains the only non-covalent inhibitor with moderate antiviral activity against SARS-CoV ( $EC_{50} = 12.9 \pm 0.7 \mu\text{M}$ )<sup>13,14</sup>. Nevertheless, given the sequence and structure similarities between SARS-CoV and SARS-CoV-2 M<sup>pro</sup>, and the common binding mode of ML188 (R) and calpain inhibitor XII, we hypothesize that ML188 (R) is a promising scaffold for lead optimization. In this study, we developed an expedited drug discovery approach to the design of non-covalent inhibitors of SARS-CoV-2 M<sup>pro</sup>. The design was based on the binding pose of calpain inhibitor XII with SARS-CoV-2 M<sup>pro</sup> (PDB: 6XFN). The highlights of this study include: 1) The overlaying X-ray crystal structures with multiple M<sup>pro</sup> inhibitors revealed the chemical space that can be explored for drug design. 2) All designed compounds were synthesized by the one-pot Ugi-4CR, which greatly facilitated the lead optimization. Indeed, we were able to improve the enzymatic inhibition potency by 35-fold from a focused library of 39 compounds. This is a significant advantage compared to covalent inhibitors such as GC376, which involves at least a five-step

synthesis. 3) By introducing the chiral isocyanide, the diastereomer product can be conveniently separated by either silica gel column or reverse phase HPLC column, bypassing the need for an expensive chiral HPLC column. This greatly speeds up the co-crystallization. 4) The X-ray crystal structure of SARS-CoV-2 M<sup>pro</sup> in complex with **23R** reveals a new binding pocket in between S2 and S4 sites that can be explored for drug design. Overall, using the expedited drug discovery approach, this study revealed a promising non-covalent M<sup>pro</sup> inhibitor **23R** with a confirmed mechanism of action and potent cellular antiviral activity for further development.

## MATERIALS

**Cell lines and viruses.** VERO E6 cells (ATCC, CRL-1586) were cultured in Dulbecco's modified Eagle's medium (DMEM), supplemented with 5% heat inactivated FBS in a 37°C incubator with 5% CO<sub>2</sub>. SARS-CoV-2, isolate USA-WA1/2020 (NR-52281), was obtained through BEI Resources and propagated once on VERO E6 cells before it was used for this study. Studies involving the SARS-CoV-2 were performed at the UTHSCSA biosafety level-3 laboratory by personnel wearing powered air purifying respirators.

**Protein expression and purification.** SARS CoV-2 main protease (M<sup>pro</sup> or 3CL) gene from strain BetaCoV/Wuhan/WIV04/2019 and SARS-CoV main protease from strain CDC#200301157 in the pET29a(+) vector with *E. coli* codon optimization were ordered from GenScript (Piscataway, NJ). The M<sup>pro</sup> gene was then subcloned into pE-SUMO vector as described previously<sup>7,8</sup>. The expression and purification of SARS-CoV and SARS-CoV-2 M<sup>pro</sup> with unmodified N- and C-termini was detailed in our previous publication<sup>8</sup>.

The expression and purification of SARS CoV-2 papain-like protease (PL<sup>pro</sup>) was also described in our previous publications<sup>7,8,18</sup>.

**Peptide synthesis.** The SARS-CoV-2 M<sup>pro</sup> FRET substrate Dabcyl-KTSAVLQ/SGFRKME(Edans) was synthesized as described before.<sup>7</sup> The SARS-CoV-2 PL<sup>pro</sup>

FRET substrate Dabcyl-FTLRGG/APTKV(Edans) was synthesized by solid-phase synthesis through iterative cycles of coupling and deprotection using the previously optimized procedure.<sup>20</sup>

**Compound synthesis and characterization.** Details for the synthesis procedure and characterization for compounds can be found in the supplementary information.

**Native Mass Spectrometry.** Prior to analysis, the protein was buffer exchanged into 0.2 M ammonium acetate (pH 6.8) and diluted to 10  $\mu$ M. 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  $\mu$ L protein, 0.5  $\mu$ L DTT stock, and 0.5  $\mu$ L ligand stock for final concentration of 4 mM DTT and 70  $\mu$ M protein. Final ligand concentrations were 10  $\mu$ M and 30  $\mu$ M. 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.<sup>21</sup>

**Enzymatic assays.** The main protease (M<sup>Pro</sup>) enzymatic assays were carried out in M<sup>Pro</sup> reaction buffer containing 20 mM HEPES pH 6.5, 120 mM NaCl, 0.4 mM EDTA, 20% glycerol and 4 mM DTT, and the SARS-CoV-2 papain-like protease (PL<sup>Pro</sup>) enzymatic assays were carried out in PL<sup>Pro</sup> reaction buffer containing 50 mM HEPES, pH7.5, 0.01% triton-100 and 5 mM DTT. The percentage of inhibition and enzymatic IC<sub>50</sub> values were calculated as previously described<sup>7,8</sup>. Briefly, the assay was performed in 96-well plates with 100  $\mu$ l of 100 nM M<sup>Pro</sup> protein or 200 nM PL<sup>Pro</sup> protein in their respective reaction buffers. Then 1  $\mu$ l testing compound

at various concentrations was added to each well and incubated at 30 °C for 30 min. The enzymatic reaction was initiated by adding 1 µl of 1 mM corresponding FRET substrate (the final substrate concentration is 10 µM). The reaction was monitored in a Cytation 5 image reader with filters for excitation at 360/40 nm and emission at 460/40 nm at 30 °C for 1 hr. The initial velocity of the enzymatic reaction with and without testing compounds was calculated by linear regression for the first 15 min of the kinetic progress curve.

For the Morrison plot, 10 µl 100 nM SARS-CoV-2 M<sup>pro</sup> protein was added to 190 µl of M<sup>pro</sup> reaction buffer containing testing compound and the FRET substrate, and the reaction was monitored for 2 hr. The final FRET substrate concentration in this assay is 20 µM. K<sub>i</sub> was determined with Morrison equation in Prism 8 as described previously<sup>7,8</sup>.

For Michealis-Menten and Lineweaver-Burk plots, assay was carried as follows: 50 µl of 50 µM M<sup>pro</sup> protein was added to 50 µl reaction buffer containing testing compound and various concentrations of FRET substrate to initiate the enzyme reaction. The initial velocity of the enzymatic reaction with and without testing compounds was calculated by linear regression for the first 15 min of the kinetic progress curve, the plotted against substrate concentrations in Prism 8 with Michaelis-Menten equation.

**Differential scanning fluorimetry (DSF).** The thermal shift binding assay (TSA) was carried out using a Thermal Fisher QuantStudio™ 5 Real-Time PCR System as described previously<sup>7,8</sup>. Briefly, 3 µM SARS-CoV-2 M<sup>pro</sup> protein in M<sup>pro</sup> reaction buffer was incubated with various concentrations of compound ML188 or **23R** at 30 °C for 30 min. 1X SYPRO orange dye was added and fluorescence of the well was monitored under a temperature gradient range from 20 °C to 90 °C with 0.05 °C/s incremental step. Measured T<sub>m</sub> was plotted against compound concentration with one-site binding function in Prism 8.



**Cytotoxicity measurement.** Evaluation of the cytotoxicity of compounds were carried out using the neutral red uptake assay<sup>22,23</sup>. Briefly, 80,000 cells/mL of the tested cell lines were dispensed into 96-well cell culture plates at 100  $\mu$ L/well. Twenty-four hours later, the growth medium was removed and washed with 150  $\mu$ L PBS buffer. 200  $\mu$ L fresh serum-free medium containing serial diluted compounds was added to each well. After incubating for 5 days at 37 °C, the medium was removed and replaced with 100  $\mu$ L DMEM medium containing 40  $\mu$ g/mL neutral red and incubated for 2-4 h at 37 °C. The amount of neutral red taken up was determined by measuring the absorbance at 540 nm using a Multiskan FC Microplate Photometer (Fisher Scientific). The CC<sub>50</sub> values were calculated from best-fit dose response curves with variable slope in Prism 8.

**Immunofluorescence assay.** Antiviral immunofluorescence assay was carried out as previously described<sup>8</sup>. Briefly, Vero E6 cells in 96-well plates (Corning) were infected with SARS-CoV-2 (USA-WA1/2020 isolate) at MOI of 0.05 in DMEM supplemented with 1% FBS. Immediately before the viral inoculation, the tested compounds in a three-fold dilution concentration series were also added to the wells in triplicate. The infection proceeded for 48 h without the removal of the viruses or the compounds. The cells were then fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton-100, blocked with DMEM containing 10% FBS, and stained with a rabbit monoclonal antibody against SARS-CoV-2 NP (GeneTex, GTX635679) and an Alexa Fluor 488-conjugated goat anti-mouse secondary antibody (ThermoFisher Scientific). Hoechst 33342 was added in the final step to counterstain the nuclei. Fluorescence images of approximately ten thousand cells were acquired per well with a 10x objective in a Cytation 5 (BioTek). The total number of cells, as indicated by the nuclei staining, and the fraction of the infected cells, as indicated by the NP staining, were quantified with the cellular analysis module of the Gen5 software (BioTek).

**Antiviral assay in Calu-3 cells.** Calu-3 cells (ATCC, HTB-55) grown in Minimal Eagles Medium supplemented with 0.1% non-essential amino acids, 0.1% penicillin/streptomycin, and 10% FBS are plated in 384 well plates. The next day, 50 nL of drug suspended in DMSO is added as an 8-pt dose response with three-fold dilutions between test concentrations in triplicate, starting at 40  $\mu$ M final concentration. The negative control (DMSO, n=32) and positive control (10  $\mu$ M Remdesivir, n=32) are included on each assay plate. Calu3 cells are pretreated with controls and test drugs (in triplicate) for 2 hours prior to infection. In BSL3 containment, SARS-CoV-2 (isolate USA WA1/2020) diluted in serum free growth medium is added to plates to achieve an MOI=0.5. Cells are incubated continuously with drugs and SARS-CoV-2 for 48 hours. Cells are fixed and then immunostained with anti-dsRNA (J2) and nuclei are counterstained with Hoechst 33342 for automated microscopy. Automated image analysis quantifies the number of cells per well (toxicity) and the percentage of infected cells (dsRNA+ cells/cell number) per well. SARS-CoV-2 infection at each drug concentration was normalized to aggregated DMSO plate control wells and expressed as percentage-of-control ( $\text{POC} = \% \text{ Infection}_{\text{sample}} / \text{Avg } \% \text{ Infection}_{\text{DMSO cont}}$ ). A non-linear regression curve fit analysis (GraphPad Prism 8) of POC Infection and cell viability versus the  $\log_{10}$  transformed concentration values to calculate  $\text{IC}_{50}$  values for Infection and  $\text{CC}_{50}$  values for cell viability. Selectivity index (SI) was calculated as a ratio of drug's  $\text{CC}_{50}$  and  $\text{IC}_{50}$  values ( $\text{SI} = \text{CC}_{50} / \text{IC}_{50}$ ).

**M<sup>pro</sup> crystallization and structure determination.** **23R** was added to 20 mg/mL SARS-CoV-2 M<sup>pro</sup> to a final concentration of 1.75 mM and incubated overnight at 4°C. This mixture was then diluted four-fold with protein stock buffer (20 mM Tris pH 7.5, 200 mM NaCl, 1 mM DTT) then spun down at 13,000  $\times$  g for 1 min to remove precipitate. Crystals were grown by mixing the protein-inhibitor sample with an equal volume of crystallization buffer (20% PEG 3350, 0.2 M NaF) in a vapor diffusion, hanging drop apparatus. Crystals were then transferred to a drop with crystallization buffer containing 5 mM **23R** for 1 h, followed by a brief soaking in a cryoprotectant

solution of 30% PEG 3350 and 15% glycerol with 2 mM **23R**. Crystals were then flash frozen in liquid nitrogen for X-ray diffraction.

X-ray diffraction data for the SARS-CoV-2 M<sup>Pro</sup> structures were collected on the SBC 19-ID beamline at the Advanced Photon Source (APS) in Argonne, IL, and processed with the HKL3000 software suite. The CCP4 versions of MOLREP<sup>24</sup> was used for molecular replacement using a previously solved SARS-CoV-2 M<sup>Pro</sup> structure, 6YB7. Structural refinement was performed using REFMAC5<sup>25</sup> and COOT<sup>26</sup>. The crystallographic statistics is shown in Table S1.

### **Data availability**

The complex structure for SARS-CoV-2 M<sup>Pro</sup> with **23R** has been deposited in the Protein Data Bank with the accession ID of 7KX5 (SARS-CoV-2 M<sup>Pro</sup> + Jun8-76-3A).

### **Acknowledgements**

This research was partially supported by the National Institutes of Health (NIH) (Grants AI147325 and AI157046) and the Arizona Biomedical Research Centre Young Investigator grant (ADHS18-198859) to J. W. The antiviral assay in Calu-3 cells was conducted through the NIAID preclinical service under a non-clinical evaluation agreement. 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. The SARS-CoV-2 experiments were supported by a COVID-19

pilot grant from UTHSCSA and NIH grant AI151638 to Y.X. SARS-Related Coronavirus 2, Isolate USA-WA1/2020 (NR-52281) was deposited by the Centers for Disease Control and Prevention and obtained through BEI Resources, NIAID, NIH.

## **AUTHOR CONTRIBUTIONS**

J. W. conceived and designed the study; N. K. synthesized and characterized the compounds; M. S. carried out M<sup>pro</sup> crystallization and structure determination with the assistance of X. Z, and analyzed the data with Y. C.; C. M. expressed the M<sup>pro</sup> and PL<sup>pro</sup>, and performed the IC<sub>50</sub> determination, thermal shift-binding assay, and enzymatic kinetic studies; Y. H. performed the cytotoxicity assay; X. M. and F. Z. performed the SARS-CoV-2 immunofluorescence assay in Vero E6 cells under the guidance of Y. X. D. S. performed the SARS-CoV-2 immunofluorescence assay in Calu-3 cells under the guidance of S. C. J. T. performed the native mass spectrometry experiments with the guidance from M. T. M.; J. W. and Y. C. secured funding and supervised the study; J. W., Y.C., and M.S. wrote the manuscript with the input from others.

## **Competing interests**

J. W., N. K., and C. M. are inventors of a patent claiming the use of compounds **23R** and related compounds as potential SARS-CoV-2 antivirals.

## **Additional Information**

Supplementary information accompanies this paper at

## **References**

- 1 Kalil, A. C. *et al.* Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19. *N. Engl. J. Med.* doi:10.1056/NEJMoa2031994 (2020).
- 2 Gil, C. *et al.* COVID-19: Drug Targets and Potential Treatments. *J. Med. Chem.* **63**, 12359-12386, doi:10.1021/acs.jmedchem.0c00606 (2020).
- 3 Ullrich, S. & Nitsche, C. The SARS-CoV-2 main protease as drug target. *Bioorg. Med. Chem. Lett.* **30**, 127377, doi:10.1016/j.bmcl.2020.127377 (2020).
- 4 Rut, W. *et al.* SARS-CoV-2 Mpro inhibitors and activity-based probes for patient-sample imaging. *Nature Chem. Biol.* doi:10.1038/s41589-020-00689-z (2020).
- 5 Ghosh, A. K., Brindisi, M., Shahabi, D., Chapman, M. E. & Mesecar, A. D. Drug Development and Medicinal Chemistry Efforts toward SARS-Coronavirus and Covid-19 Therapeutics. *ChemMedChem* **15**, 907-932, doi:10.1002/cmdc.202000223 (2020).
- 6 Boras, B. *et al.* Discovery of a Novel Inhibitor of Coronavirus 3CL Protease as a Clinical Candidate for the Potential Treatment of COVID-19. *bioRxiv*, 2020.2009.2012.293498, doi:10.1101/2020.09.12.293498 (2020).
- 7 Ma, C. L. *et al.* Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res.* **30**, 678-692, doi:10.1038/s41422-020-0356-z (2020).
- 8 Sacco, M. D. *et al.* Structure and inhibition of the SARS-CoV-2 main protease reveals strategy for developing dual inhibitors against Mpro and cathepsin L. *Sci. Adv.* eabe0751, doi:10.1126/sciadv.abe0751 (2020).
- 9 Rathnayake, A. D. *et al.* 3C-like protease inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV-infected mice. *Sci. Transl. Med.* **12**, eabc5332, doi:10.1126/scitranslmed.abc5332 (2020).
- 10 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, doi:10.1177/1098612X17729626 (2018).

- 11 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, doi:10.1371/journal.ppat.1005531 (2016).
- 12 Jacobs, J. *et al.* Discovery, synthesis, and structure-based optimization of a series of N-(tert-butyl)-2-(N-arylamido)-2-(pyridin-3-yl) acetamides (ML188) as potent noncovalent small molecule inhibitors of the severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL protease. *J. Med. Chem.* **56**, 534-546, doi:10.1021/jm301580n (2013).
- 13 Turlington, M. *et al.* Discovery of N-(benzo[1,2,3]triazol-1-yl)-N-(benzyl)acetamido)phenyl) carboxamides as severe acute respiratory syndrome coronavirus (SARS-CoV) 3CLpro inhibitors: Identification of ML300 and noncovalent nanomolar inhibitors with an induced-fit binding. *Bioorg. Med. Chem. Lett.* **23**, 6172-6177, doi:<https://doi.org/10.1016/j.bmcl.2013.08.112> (2013).
- 14 Zaidman, D. *et al.* An automatic pipeline for the design of irreversible derivatives identifies a potent SARS-CoV-2 Mpro inhibitor. *bioRxiv*, 2020.2009.2021.299776, doi:10.1101/2020.09.21.299776 (2020).
- 15 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, doi:10.1126/science.abb3405 (2020).
- 16 Jin, Z. *et al.* Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289-293, doi:10.1038/s41586-020-2223-y (2020).
- 17 Li, J. *et al.* Crystal structure of SARS-CoV-2 main protease in complex with the natural product inhibitor shikonin illuminates a unique binding mode. *Sci Bull (Beijing)*, doi:10.1016/j.scib.2020.10.018 (2020).
- 18 Ma, C. *et al.* Ebselen, Disulfiram, Carmofur, PX-12, Tideglusib, and Shikonin Are Nonspecific Promiscuous SARS-CoV-2 Main Protease Inhibitors. *ACS Pharmacol. Transl. Sci.* **3**, 1265-1277. doi:10.1021/acspsci.0c00130 (2020).

- 19 Gurard-Levin, Z. A. *et al.* Evaluation of SARS-CoV-2 3C-like protease inhibitors using self-assembled monolayer desorption ionization mass spectrometry. *Antiviral Res.* **182**, 104924, doi:10.1016/j.antiviral.2020.104924 (2020).
- 20 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, doi:10.1021/ja102581n (2011).
- 21 Marty, M. T. *et al.* Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370-4376, doi:10.1021/acs.analchem.5b00140 (2015).
- 22 Repetto, G., del Peso, A. & Zurita, J. L. Neutral red uptake assay for the estimation of cell viability/cytotoxicity. *Nat. Protoc.* **3**, 1125-1131, doi:10.1038/nprot.2008.75 (2008).
- 23 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, doi:10.1124/mol.116.105346 (2016).
- 24 Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022-1025, doi:Doi 10.1107/S0021889897006766 (1997).
- 25 Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta. Crystallogr. D* **67**, 355-367, doi:10.1107/S0907444911001314 (2011).
- 26 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D* **60**, 2126-2132, doi:10.1107/S0907444904019158 (2004).