The P3 *O-Tert*-Butyl-Threonine is Key to High Cellular and Antiviral Potency for Aldehyde-Based SARS-CoV-2 Main Protease Inhibitors

Yuying Ma[a,†], Kai S. Yang[a,†], Zhi Zachary Geng[a,†], Yugendar R. Alugubelli[a, †], Namir Shaabani,[b, †] Erol C. Vatansever[a,†], Xinyu R. Ma[a, †], Chia-Chuan Cho[a, †], Kaustav Khatua[a, †], Lauren Blankenship[a], Ge Yu[a], Banumathi Sankaran[c], Pingwei Li[d], Robert Allen[b], Henry Ji[b,*], Shiqing Xu[a,*], and Wenshe Ray Liu[a,d,e,f,*]

[a]Texas A&M Drug Discovery Laboratory, Department of Chemistry, Texas A&M University, College Station, TX 77843, USA

[b]Sorrento Therapeutics, Inc. San Diego, CA 92121, USA

[c]Molecular Biophysics and Integrated Bioimaging, Berkeley Center for Structural Biology, Laurence Berkeley National Laboratory, Berkeley, CA 94720, USA

[d]Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA

[e]Institute of Biosciences and Technology and Department of Translational Medical Sciences, College of Medicine, Texas A&M University, Houston, TX 77030, USA

[f]Department of Molecular and Cellular Medicine, College of Medicine, Texas A&M University, College Station, TX 77843, USA

[†]Contributed equally to the paper.

*Correspondence should be addressed to Henry Ji: hji@sorrentotherapeutics.com, Shiqing Xu: shiqing.xu@tamu.edu and Wenshe Ray Liu: wsliu2007@tamu.edu

Supporting Information Placeholder

ABSTRACT: As an essential enzyme to SARS-CoV-2, main protease (M^{Pro}) is a viable target to develop antivirals for the treatment of COVID-19. By varying chemical compositions at both P2 and P3 sites and the *N*-terminal protection group, we synthesized a series of M^{Pro} inhibitors that contain β-(S-2-oxopyrrolidin-3-yl)-alaninal at the P1 site. These inhibitors have a large variation of determined IC_{50} values that range from 4.8 to 650 nM. The determined IC_{50} values reveal that relatively small side chains at both P2 and P3 sites are favorable for achieving high *in vitro* M^{Pro} inhibition potency, the P3 site is tolerable toward unnatural amino acids with two alkyl substituents on the α-carbon, and the inhibition potency is sensitive toward the *N*-terminal protection group. X-ray crystal structures of M^{Pro} bound with 16 inhibitors were determined. All structures show similar binding patterns of inhibitors at the M^{Pro} active site. A covalent interaction between the active site cysteine and a bound inhibitor was observed in all structures. In M^{Pro}, large structural variations were observed on residues N142 and Q189. All inhibitors were also characterized on their inhibition of M^{Pro} in 293T cells, which revealed their *in cellulo* potency that is drastically different from their *in vitro* enzyme inhibition potency. Inhibitors that showed high *in cellulo* potency all contain *O-tert*-butyl-threonine at the P3 site.

Based on the current and a previous study, we conclude that *O-tert*-butyl-threonine at the P3 site is a key component to achieve high cellular and antiviral potency for peptidyl aldehyde inhibitors of M^{Pro}. This finding will be critical to the development of novel antivirals to address the current global emergency of concerning the COVID-19 pandemic.

■ INTRODUCTION

COVID-19 is the prevailing pandemic that has ravaged much of the world. As the COVID-19 pathogen, SARS-CoV-2 uses its membrane Spike protein to recognize the human receptor ACE2 for infection.[1-2] Current COVID-19 vaccines all target this process for the neutralization of the virus. However, the continuous emergence of new viral strains that evade vaccines demands other antivirals to be developed as well. SARS-CoV-2 has a very large open reading frame ORF1ab that is translated to two large polypeptides pp1a and pp1b in the human cell host.[3-4] The processing of pp1a and pp1b to 16 nonstructural proteins (nsps) that are functionally critical to the viral replication relies on proteolytic activities of two internal nsp fragments, nsp3 and nsp5.[5] Nsp5 is also called 3C-like

protease and more recently main protease (M^{Pro}). Since M^{Pro} hydrolyzes 13 out of the total of 16 nsps, it has been considered as a viable target for the development of antivirals. In the past year, a number of papers have been published on the development of peptidyl aldehydes that contain β-(S-2-oxopyrrolidin-3-yl)-alaninal (opal) at the P1 position as potent M^{Pro} inhibitors.[6-14] Other inhibitors were developed as well.[15-28] However, systematic structure-activity relationship (SAR) studies of opal-based M^{Pro} inhibitors are needed. In the current study, we explore variations at the P2 and P3 sites and the *N*-terminal protection group in opal-based peptidyl M^{Pro} inhibitors. These inhibitors were characterized in their *in vitro* and *in cellulo* inhibition of M^{Pro} and their interactions with M^{Pro} by X-ray protein crystallography. Our study reveals that *O-tert*-butyl-threonine at the P3 site in an opal-based M^{Pro} inhibitor is critical to achieve high cellular and antiviral potency.

■ RESULTS

The Design and Synthesis of MPI11-28. In previous studies, we developed MPI1-10 and characterized both their M^{Pro} inhibition potency and their antiviral potency.[13-14] As a tripeptidyl aldehyde, MPI8 shows the highest antiviral potency with an EC_{50} value of 30 nM to neutralize SARS-CoV-2 (USA-WA1/2020) in Vero E6 cells.[29] MPI3, another tripeptidyl aldehyde doesn't have high antiviral potency. However, it has the highest *in vitro* M^{Pro} inhibition potency with an IC_{50} value as 8.5 nM.[13] To explore how substituents at different positions in a tripeptidyl opal-based M^{Pro} inhibitor (shown on the top of Figure 1) influence its potency, we decided to carry out a systematic SAR study. We maintained opal at the P1 position due to its established preferential binding to the M^{Pro} P1 binding pocket and its covalent adduct formation with C145, the M^{Pro} catalytic cysteine. M^{Pro} tolerates leucine and phenylalanine at the P2 site of a substrate. Therefore, we chose to vary this site in our inhibitor design with β-alkyl alanines with a size between or around leucine and phenylalanine. Chosen alkyl substituents are isopropyl, phenyl, cyclohexyl, t-butyl, isopropenyl, cyclopropyl, 2-furyl, and 2-thienyl. M^{Pro} doesn't have a binding
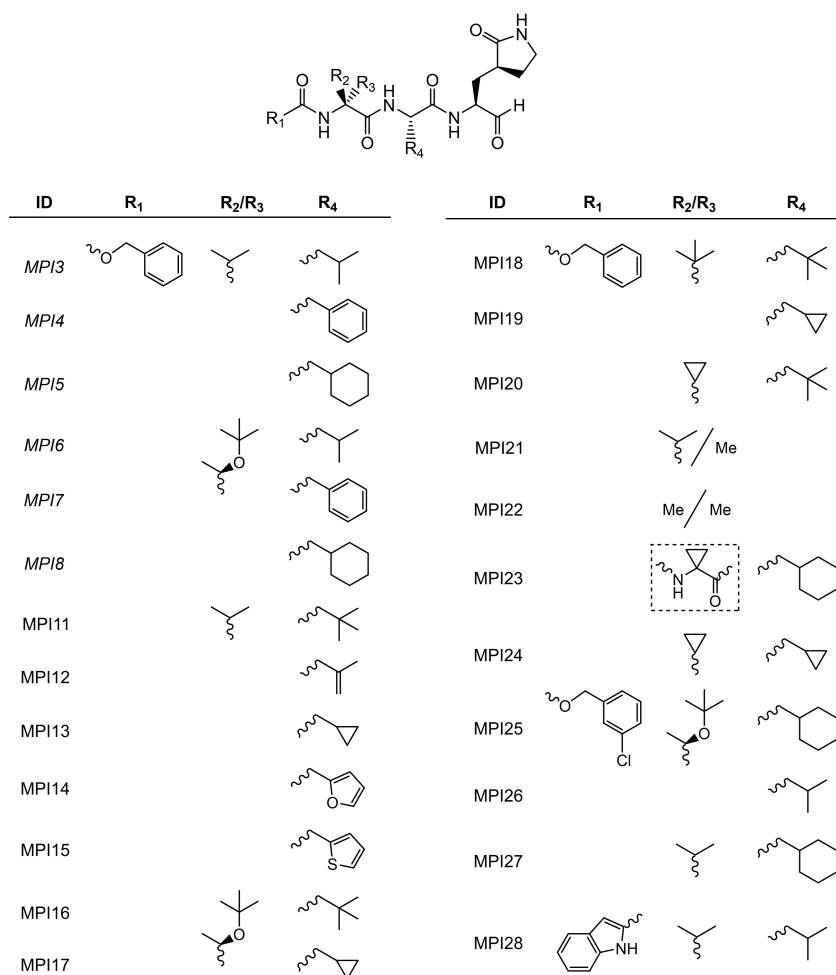

Figure 1. Structures of tripeptidyl M^{Pro} inhibitors **MPI11-28**. Inhibitors that have just a single component at R_2/R_3 shown contain a hydrogen at the R_3 position. **MPI3-8** were previously developed and are shown for comparison.

pocket for the P3 residue in a substrate. However, native M^{Pro} substrates have valine or a similar size residue at this position. Based on known opal inhibitors and substrates of M^{Pro}, we varied this site with amino acids including valine, *O-tert*-butyl-threonine, *L*-cyclopropylglycine, *L-tert*-butyl-glycine, *L*-α-methyl-valine, dimethylglycine, and 1-aminocyclopropane-1-carboxylate. The *N*-terminal protection group was chosen between CBZ, *m*-chloro CBZ, and indole-2-carboxylate due to their demonstrated
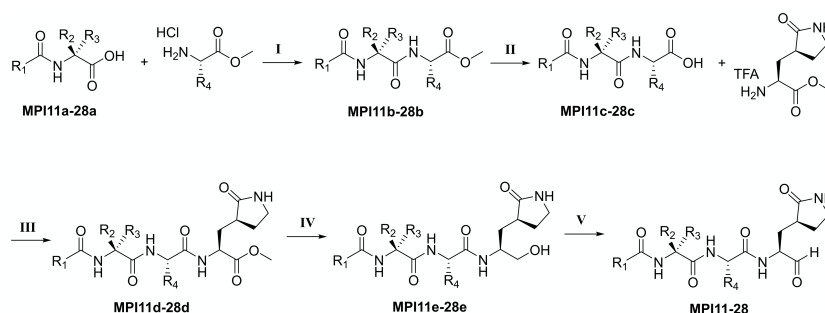
Scheme 1. Reagents and conditions: (I) HATU, DIPEA, DMF; (II) LiOH•H_2O, THF/H_2O; (III) HATU, DIPEA, DMP; (IV) $LiBH_4$, THF; (V) DMP, DCM.

contributions to antiviral potency.[6, 10] A total of 18 new M^{Pro} inhibitors, designated as MPI11-28 shown in Figure 1, were designed. We synthesized all inhibitors according to the synthetic route shown in Scheme 1. In this synthesis, a P2 amino acid ester was conjugated with a *N*-protected P3 amino acid and the afforded product was then hydrolyzed to free its *C*-terminal carboxylate for reaction with β-(S-2-oxopyrrolidin-3-yl)-alanine ester. An obtained tripeptidyl ester was reduced to afford a *C*-terminal alcohol that was then oxidized via Dess-Martin oxidation in a mild condition to make a final product.
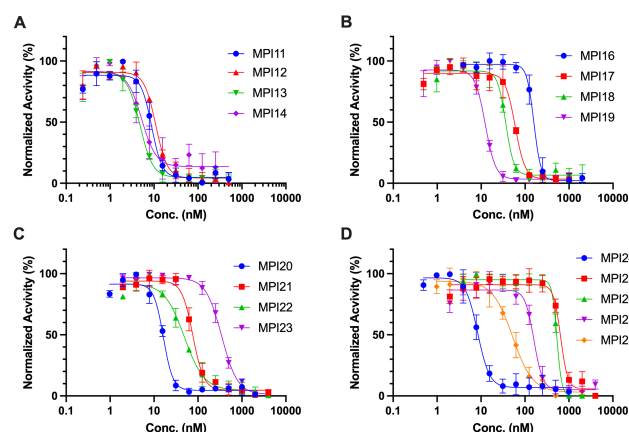

Figure 2. Inhibition curves of MPI11-28 on M^{Pro}. Triplicate experiments were performed for each compound. For all experiments, 20 or 10 nM M^{Pro} was incubated with an inhibitor for 30 min before 10 μM Sub3 was added. The M^{Pro}-catalyzed Sub3 hydrolysis rate was determined by measuring linear increase of product fluorescence (Ex: 336 nm /Em: 455 nm) for 5 min.

Kinetic characterizations of MPI11-28 on their enzymatic inhibition of M^{Pro}. We followed a previously established protocol that uses Sub3, a fluorogenic peptide substrate of M^{Pro} to determine IC_{50} values for MPI11-28.[28] In this assay, we preincubated M^{Pro} with an inhibitor for 30 min before Sub3 was added and the fluorescent product formation was recorded in a fluorescence plate reader. Except MPI15 that is insoluble in DMSO and therefore was not characterized, all other inhibitors displayed well traceable inhibition curves as shown in Figure 2. We fit all data to a four-parameter variable slope inhibition equation in GraphPad 9.0 to obtain IC_{50} values for all inhibitors. As shown in Table 1, MPI11-28 display a large variation of IC_{50} values that ranges from 4.8 to 650 nM. MPI11-14 all contain valine at the P3 site. They have the lowest IC_{50} values among all newly developed M^{Pro} inhibitors. MPI13-14 have determined IC_{50} values around 5 nM. Since 10 nM M^{Pro} is the lowest enzyme concentration we can use to do the inhibition analysis, 5 nM is technically the lowest IC_{50} value we can detect. Therefore, M^{Pro} inhibition potency for MPI13-14 is likely higher than what the numbers shown in Table 1 indicate. In comparison to MPI13-14, MPI11-12 have slightly higher IC_{50} values that are around 10 nM. In previous work, we developed another three M^{Pro} inhibitors MPI3-5 that also contain a valine at the P3 site. All three have low IC_{50} values.

All 7 inhibitors that have valine at the P3 site and a *N*-terminal CBZ protection group are among the most potent M^{Pro} inhibitors *in vitro*. A P3 valine apparently favors the enzyme inhibition kinetics. A comparison of all seven inhibitors also reveals that a residue at the P2 site with a size close to leucine leads to better M^{Pro} inhibition. β–Cyclopropyl alanine and β-(furan-2yl) alanine that have a more rigid side chain than leucine favor the enzyme inhibition kinetics more. MPI6-8 that have a P3 *O-tert*-butyl-threonine were developed previously and showed high cellular and antiviral potency. We synthesized two more M^{Pro} inhibitors with a P3 *O-tert*-butyl-threonine, MPI16-17. Both have an IC_{50} value similar to MPI6-8. Varying the P2 position does not seem to significantly change *in vitro* M^{Pro} inhibition potency among this group of inhibitors. To explore further on whether variations at the P3 site lead to different inhibition potency, we developed MPI18-24. Variants include two dialkyl glycines and 1-aminocyclopropane-1-carboxylate that are not standard *L*-amino acids. Except MPI23 that contains a P3 1-aminocyclopropane-1-carboxylate, all other inhibitors have an IC_{50} value below 100 nM. The more rigid P3 residue in MPI23 likely contributes to its less favorable M^{Pro} inhibition kinetics. MPI24 is structurally similar to MPI13. Its enzyme inhibition potency is similar to MPI13 as well. Three inhibitors MPI25-27 that have a *N*-terminal *m*-chloro CBZ group were developed and characterized as well. In comparison to their regular CBZ-containing counterpart inhibitors MPI8, MPI6, and MPI5, they display more than 5-fold higher IC_{50} values. Apparently adding a *m*-chloro substituent to the terminal CBZ group doesn't improve *in vitro* enzyme inhibition potency. MPI28 has a *N*-terminal indole-2-carboxylate. In comparison to its CBZ counterpart inhibitor MPI3, MPI28 has a 6-folder higher IC_{50} value. In comparison to indole-2-carboxylate, it is evident that CBZ serves as a better *N*-terminal group for improved *in vitro* enzyme inhibition potency.

X-Ray Crystallography analysis of M^{Pro} bound with 16 different inhibitors. To characterize our developed inhibitors in their interactions with M^{Pro}, we crystallized M^{Pro} in its apo form, soaked apo-M^{Pro} crystals with different inhibitors and then determined structures of M^{Pro} bound with different inhibitors using X-ray crystallography. Among 17 inhibitors that we used to soak crystals, 16 were observable in the active site of M^{Pro}. All structures were determined in high resolutions (Table S1). $2F_o$-F_c electron density map around each inhibitor clearly shows a covalent interaction with C145 of M^{Pro} to generate a hemithioacetal (Figure 3A). In all formed hemithioacetals, the hemithioacetal alcohol takes an *S* configuration. In most structures, electron density around inhibitors is well shaped for modeling the P1 opal, P2, and P3 residues. For MPI12, the isopropenyl group shows weak electron density around its 1- and 3-carbon atoms indicating that the group is probably freely rotating around its 2-carbon atom in the active site of M^{Pro}. Although a few M^{Pro}-inhibitor
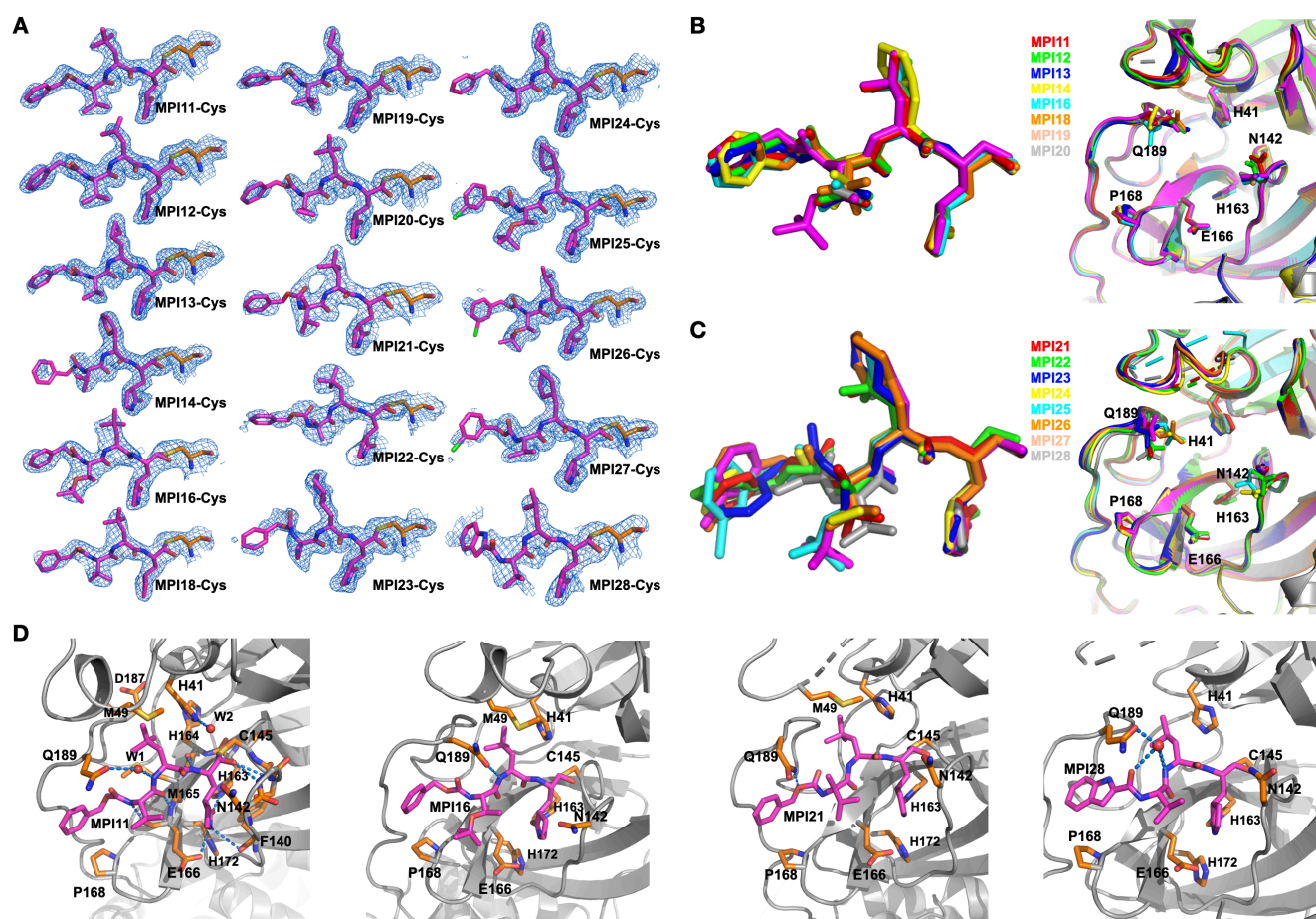
Figure 3. Crystal structures of M^{Pro} bound with 16 MPIs. (A) Contoured $2F_o$-F_c maps at the 1σ level around 16 MPIs and C145 in the active site of M^{Pro}. (B) Superposition of M^{Pro} complexes with 8 MPIs including MPI11-14, MPI16, and MPI18-20. Inhibitors on the left and proteins on the right are presented separately for clarity. (C) Superposition of M^{Pro} complexes with MPI21-28. Inhibitors and proteins are shown separately. (D) Interactions between the inhibitor and active site residues in M^{Pro} complexes with MPI11, MPI16, MPI21, and MPI28. Some of hydrogen bonds between ligands and M^{Pro} are shown as blue dotted lines.

complexes have electron density that can be used to model the *N*-terminal group, the majority have relatively weak electron density around the *N*-terminal group indicating weak binding of the *N*-terminal group at the active site. Superpositioning all M^{Pro}-inhibitor complexes shows a similar binding mode for all inhibitors to the M^{Pro} active site (Figures 3B and 3C). MPI28 displays the most difference from other inhibitors due to its unique *N*-terminal indole-2-carboxylate. All other inhibitors show very similar binding to M^{Pro} except that different orientations are adopted at the relatively flexible *N*-terminal group. In all M^{Pro}-inhibitor complexes, the protein shows a similar structure except at two residues N142 and Q189 and the M49-containing helix. Both N142 and Q189 adopt varied orientations in different M^{Pro}-inhibitor structures and the M49-containing helix is structurally rearranged and not visible in most M^{Pro}-inhibitor complexes. Due to a large number of crystal structures that are presented in this work, we will focus on only four representative M^{Pro}-inhibitor complexes to describe structure variations of M^{Pro} at the active site. In M^{Pro}-MPI11 as shown in Figure 3D, MPI11 forms extensive hydrogen bonds with M^{Pro}. The hemithioacetal hydroxyl group forms two hydrogen

bonds with two backbone amide nitrogen atoms explaining its exclusive *S* configuration. The opal side-chain amide forms three hydrogen bonds with E166, N142, and F140 and its backbone nitrogen forms a hydrogen bond with H164. The side chain of the P2 β-*tert*-butyl-alanine (4-methyl-leucine) involves van der Waals interactions with M^{Pro} residues surrounding the P2 binding pocket including M49, H41, M165, D187, and Q189. Its α-nitrogen interacts with the side chain of Q189 indirectly through a water molecule. The P3 residue of MPI11 forms two hydrogen bonds with backbone nitrogen and oxygen atoms of E166. The *N*-terminal CBZ of MPI11 interacts hydrophobically with the backbone of the Q189-containing loop and the side chain of P168. In M^{Pro}-MPI11, the side chain of N142 flips away from the opal side chain of MPI11. MPI16 interacts with M^{Pro} similar to MPI11. Notable differences are at N142 and Q189. N142 forms van der Waals interactions with the opal oxopyrrolidine ring and Q189 forms a hydrogen bond directly with the backbone nitrogen of the P2 residue. In M^{Pro}-MPI21, N142 adopts a conformation away from the opal side chain and Q189 forms a hydrogen bond with a CBZ carbamate oxygen atom. MPI28 has a different *N*-terminal group.

Interactions involving its *N*-terminal group are different from other inhibitors. The indole-2-carboxyl oxygen forms a hydrogen bond with a water molecule that also forms hydrogen bonds with Q189 and the backbone nitrogen of the P2 residue.
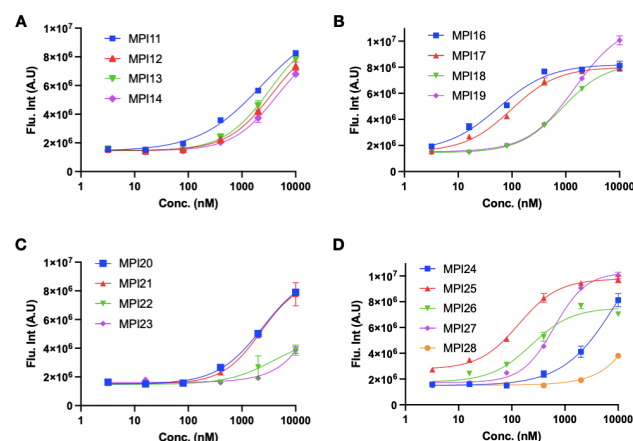

Figure 4. Cellular potency of MPI11-28 in their inhibition of M^{Pro} to drive host 293T cell survival and overall M^{Pro}-eGFP expression.

Characterizations of Cellular M^{Pro} inhibition potency of MPI11-28. M^{Pro} is acutely toxic to a human cell host. Using this unique characteristic, we previously developed a cell-based analysis to characterize cellular M^{Pro} inhibition potency for M^{Pro} inhibitors.[29] In this assay, an inhibitor with *in cellulo* potency suppresses cellular toxicity from the expression of an M^{Pro}-eGFP fusion protein, which leads to enhanced overall expression of M^{Pro}-eGFP that can be characterized by flow cytometry. We consider this assay more advantageous over a direct antiviral assay in the characterization of M^{Pro} inhibitors since a compound may inhibit host proteases such as TMPRSS2, furin, and cathepsin L that are critical for SARS-CoV-2 infection to provide false positive *in cellulo* potency results of M^{Pro} inhibition. *In cellulo* potency determined by this method for MPI5-8 agreed well with their antiviral potency that were characterized using SARS-CoV-2 (USA-WA1/2020) in Vero E6 cells. We adopted this assay to characterize MPI11-28 as well. All inhibitors were tested up to 10 μM in their inhibition of M^{Pro}-eGFP in human 293T cells. Overall cellular eGFP fluorescence were plotted against the inhibition concentration to obtain their EC_{50} values. Results are presented in Figure 4 and Table 1. As shown in Figure 4A, MPI11-14 exhibit minimal *in cellulo* potency to inhibit M^{Pro}, which is in significant contrast to their very high *in vitro* enzyme inhibition potency. Since their inhibition curves do not reach a plateau, their EC_{50} values are estimated as higher than 2 μM. MPI16-17 are the two most potent inhibitors among MPI11-28 on *in cellulo* potency with determined EC_{50} values as 56 nM and 97 nM, respectively (Figure 4B). In comparison to MPI16-17, MPI18-19 have weaker *in cellulo* potency with determined EC_{50} values as 860 nM and 1,500 nM, respectively. MPI20-23 all have weak *in*

cellulo potency (Figure 4C). MPI22-23 show very low inhibition of M^{Pro}-eGFP in 293T cells. At 10 μM, their driven M^{Pro}-eGFP expression is lower than half of the plateau level of M^{Pro}-eGFP observed for MPI16-19. For these four inhibitors, their EC_{50} values are estimated as higher than 2 μM for MPI20-21 and higher than 10 μM for MPI22-23. Among MPI24-28 (Figure 4D), MPI25-27 exhibit high *in cellulo* potency with EC_{50} values as 140 nM, 230 nM, and 630 nM respectively. On contrary to their high *in vitro* enzyme inhibition potency, MPI24 and MPI28 have very weak *in cellulo* potency with estimated EC_{50} values as higher than 2 μM and higher than 10 μM, respectively.
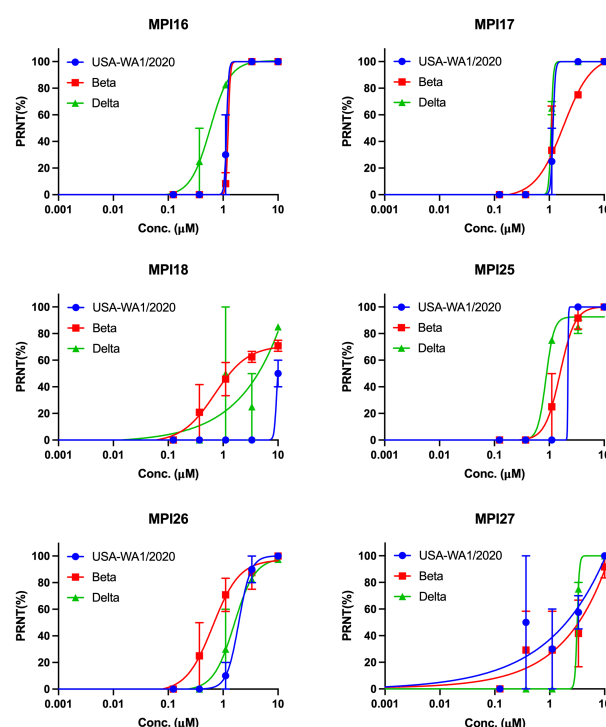

Figure 5. Plaque reduction neutralization tests (PRNTs) of MPI16-18 and MPI25-27 on their inhibition of three SARS-CoV-2 strains USA-WA1/2020, Beta and Delta in Vero E6 cells. Two repeats were conducted for each concentration.

Characterizations of antiviral potency of six selected inhibitors on three SARS-CoV-2 variants. Four most cellularly potent inhibitors MPI16, MPI17, MPI25, and MPI26 that contain a P3 *O-tert*-butyl-threonine and two mildly cellularly potent inhibitors MPI18 and MPI27 that don't contain a P3 *O-tert*-butyl-threonine were analyzed on their antiviral potency against three SARS-CoV-2 variants including USA-WA1/2020, Beta and Delta. We conducted plaque reduction neutralization tests in Vero E6 cells for all inhibitors. We infected Vero E6 cells by virus in the presence of an inhibitor at various concentrations for three days and then quantified viral plaque reduction. As presented in the attached figure and table, all four inhibitors that contain a P3 *O-tert*-butyl-threonine showed high antiviral potency with determined EC_{50} values

Table 1: Determined enzymatic IC_{50} and cellular EC_{50} values of M^{Pro} inhibitors

ID	Enzymatic IC_{50} (nM)	Cellular EC_{50} (µM)	Antiviral EC_{50} (µM)	PDB Entry	ID	Enzymatic IC_{50} (nM)	Cellular EC_{50} (µM)	Antiviral EC_{50} (µM)	PDB Entry
MPI3	*8.5 [b]*	*> 2 [c]*		*7JQ0*	MPI17	60	0.097	1.2[e]/1.8[f]/1.1[g]	-
MPI4	*15[b]*	*> 2 [c]*		*7JQ1*	MPI18	36	0.86	> 5[e]/0.71[f]/> 5[g]	7RVR
MPI5	*33[b]*	*0.66[c]*		*7JQ2*	MPI19	13	1.5		7RVS
MPI6	*60[b]*	*0.12[c]*		*7JQ3*	MPI20	16	> 2		7RVT
MPI7	*47[b]*	*0.19[c]*		*7JQ4*	MPI21	77	> 2		7RVU
MPI8	*105[b]*	*0.03[c]*		*7JQ5*	MPI22	49	> 10		7RVV
MPI11	8.8	> 2		7RVM	MPI23	350	> 10		7RVW
MPI12	11	> 2		7RVN	MPI24	8.4	> 2		7RVX
MPI13	4.8[d]	> 2		7RVO	MPI25	650	0.14	2.2[e]/1.6[f]/0.87[g]	7RVY
MPI14	5.2[d]	> 2		7RVP	MPI26	530	0.23	1.9[e]/0.65[f]/1.6[g]	7RVZ
MPI15[a]	n.d.	n.d.		-	MPI27	160	0.63	> 5[e]/> 5[f]/3.2[g]	7RW0
MPI16	150	0.056	1.2[e]/1.2[f]/0.58[g]	7RVQ	MPI28	52	> 10		7RW1

[a]Its IC_{50} and EC_{50} values were not determined (n.d.) due to insolubility.
[b]Data were taken from ref. [13].
[c]Data were taken from ref. [29].
[d]Reaching the detection limit.
[e]Antiviral EC_{50} value for the USA-WA1/2020 strain.
[f]Antiviral EC_{50} value for the Beta strain.
[g]Antiviral EC_{50} value for the Delta strain.

around 1 µM. Except MPI8 on the Beta strain, MPI18 and MPI27 that don't have a P3 *O-tert*-butyl-threonine showed only mild antiviral potency. These antiviral results support that P3 *O-tert*-butyl-threonine in a peptidyl aldehyde inhibitor leads to high antiviral potency.

■ **DISCUSSION**

There are 11 unique proteolytic sites in pp1a and pp1b that are hydrolyzed by M^{Pro}. Glutamine is a strictly required residue at the P1 site for all 11 sites. Due to this strict requirement and medicinal chemistry information learnt from the same function enzyme in SARS-CoV, most peptidyl inhibitors that have been developed for M^{Pro} have maintained a β-(S-2-oxopyrrolidin-3-yl)-alanine analog at the P1 site for improved potency. In our design, we keep opal at this site for all inhibitors. Our crystallography results show a well-shaped opal side chain in all M^{Pro}-inhibitor complexes that fits neatly in the P1 binding pocket of the enzyme. Since opal is optimized for this site, there is small chemical space to manipulate its side chain for improved binding. In all determined structures, the side chain of N142 adopts different conformations that do not necessarily interact with the opal oxopyrrolidine ring. It may be possible to modify the opal oxopyrrolidine ring to introduce interactions such as hydrogen bonding with the side chain of N142 for improved binding. Among all 11 M^{Pro}-targeted proteolytic sites, 9 have a P2 leucine, 1 has a P2 phenylalanine, and 1 has a P2 valine. The enzyme has an apparent preference for a P2 leucine in its substrates. A comparison of *in vitro* enzyme inhibition potency for all MPIs that we have developed also reveals that M^{Pro} prefers a P2 residue with a similar size as leucine in a peptidyl inhibitor for favorable binding. Converting the isopropyl group at the β-carbon

of the P2 leucine to cyclopropyl and 2-furanyl improves the enzyme inhibition kinetics. Although changing the P2 leucine in an inhibitor to a larger residue such as phenylalanine and β-cyclohexylalanine diminishes *in vitro* M^{Pro} inhibition potency, the level is not significant.

M^{Pro} shows almost no preference to the P3 residue in its targeted proteolytic sites in pp1a and pp1b. We explored a variety of chemical variants at this site including dialkyl glycines and 1-aminocyclopropane-1-carboxylate that have no α-proton. Although M^{Pro} displays a clear preference for a P3 valine, other substituents at this site are well tolerated. Given the diverse structures of tolerable chemical compositions at this site, D-amino acids might be introduced at this site for the development of novel M^{Pro} inhibitors. This will need to be explored further. All our crystal structures show weak electron density around the *N*-terminal group indicating its weak binding to the enzyme. Adding a *m*-chloride to the *N*-terminal CBZ group also leads to significant decrease in *in vitro* enzyme inhibition potency. Our study indicates that both CBZ and indole-2-carboxyl are not optimal chemical groups for interactions with M^{Pro}. Some smaller groups need to be explored at this site.

Although most inhibitors we have developed in this study exhibit very high *in vitro* enzyme inhibition potency and two inhibitors have IC_{50} values reaching the characterization limit, just a few inhibitors show *in cellulo* potency to inhibit M^{Pro} in 293T cells. Inhibitors that have a P3 *O-tert*-butyl-threonine all show high *in cellulo* potency. All other P3 substituents lead to low *in cellulo* potency. Many factors may contribute to this phenomenon. Since all MPIs are peptidyl inhibitors, a small or native P3 residue might be prone to proteolytic degradation by host proteases leading to low *in cellulo* potency. This is

partially supported by characterizable *in cellulo* potency for MPI18-19 that have a P3 β-*tert*-butyl-alanine. It is also possible that a P3 *O-tert*-butyl-threonine introduces more favorable cellular permeability than other P3 substituents into an M^Pro inhibitor. MPI21-23 that have a P3 dialkylglycine are expected to be more resistant to proteolytic digestion by human proteases than other MPIs. Their low *in cellulo* potency might be due to low cellular permeability. Based on *in cellulo* potency of MPIs, we can also derive that an *N*-terminal CBZ group works better than the other two groups that have been tested at this site and a P2 β-cyclohexylalanine favors high *in cellulo* potency.

■ CONCLUSIONS

In combination with our previous studies,[13,29] we can conclude that a P3 *O-tert*-butyl-threonine in peptidyl aldehyde inhibitors for M^Pro is optimal for high *in cellulo* and antiviral potency. This new and critical finding will be highly useful for the development of novel antiviral lead series and drug candidates with potential to treat COVID-19. However, chemical compositions at the P2 site and the *N*-terminal group in peptidyl aldehyde inhibitors for optimal *in cellulo* and antiviral potency need to be explored further.

■ EXPERIMENTAL SECTION

ASSOCIATED CONTENT

Supporting Information.
The Supporting Information is available free of charge on the ACS Publications website.

Supplementary information for the synthesis of MPI11-28, NMR spectrascopies and HPLC chromatographies of MPI11-28, X-ray data collection and processing parameters, and all original flow cytometry graphs. All compounds are >95% pure by HPLC analysis.
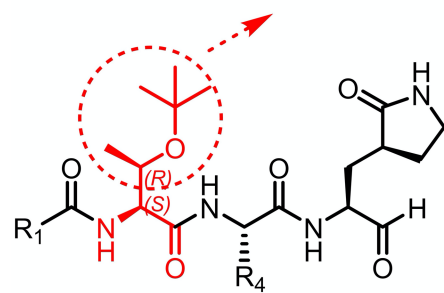
ACKNOWLEDGMENT

REFERENCE

1. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X., Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020,** *581* (7807), 215-220.

2. Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q., Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020,** *367* (6485), 1444-1448.

3. Kim, D.; Lee, J. Y.; Yang, J. S.; Kim, J. W.; Kim, V. N.; Chang, H., The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020,** *181* (4), 914-921 e10.

4. Zhou, P.; Yang, X. L.; Wang, X. G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H. R.; Zhu, Y.; Li, B.; Huang, C. L.; Chen, H. D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R. D.; Liu, M. Q.; Chen, Y.; Shen, X. R.; Wang, X.; Zheng, X. S.; Zhao, K.; Chen, Q. J.; Deng, F.; Liu, L. L.; Yan, B.; Zhan, F. X.; Wang, Y. Y.; Xiao, G. F.; Shi, Z. L., A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020,** *579* (7798), 270-273.

5. Morse, J. S.; Lalonde, T.; Xu, S.; Liu, W. R., Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV. *Chembiochem* **2020,** *21* (5), 730-738.

6. Dai, W.; Zhang, B.; Jiang, X. M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L. K.; Xu, Y.; Yang, H.; Liu, H., Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **2020,** *368* (6497), 1331-1335.

7. Fu, L.; Ye, F.; Feng, Y.; Yu, F.; Wang, Q.; Wu, Y.; Zhao, C.; Sun, H.; Huang, B.; Niu, P.; Song, H.; Shi, Y.; Li, X.; Tan, W.; Qi, J.; Gao, G. F., Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat Commun* **2020,** *11* (1), 4417.

8. Dai, W.; Jochmans, D.; Xie, H.; Yang, H.; Li, J.; Su, H.; Chang, D.; Wang, J.; Peng, J.; Zhu, L.; Nian, Y.; Hilgenfeld, R.; Jiang, H.; Chen, K.; Zhang, L.; Xu, Y.; Neyts, J.; Liu, H., Design, Synthesis, and Biological Evaluation of Peptidomimetic Aldehydes as Broad-Spectrum Inhibitors against Enterovirus and SARS-CoV-2. *J Med Chem* **2021,** 10.1021/acs.jmedchem.0c02258.

9. Qiao, J.; Li, Y. S.; Zeng, R.; Liu, F. L.; Luo, R. H.; Huang, C.; Wang, Y. F.; Zhang, J.; Quan, B.; Shen, C.; Mao, X.; Liu, X.; Sun, W.; Yang, W.; Ni, X.; Wang, K.; Xu, L.; Duan, Z. L.; Zou, Q. C.; Zhang, H. L.; Qu, W.; Long, Y. H.; Li, M. H.; Yang, R. C.; Liu, X.; You, J.; Zhou, Y.; Yao, R.; Li, W. P.; Liu, J. M.; Chen, P.; Liu, Y.; Lin, G. F.; Yang, X.; Zou, J.; Li, L.; Hu, Y.; Lu, G. W.; Li, W. M.; Wei, Y. Q.; Zheng, Y. T.; Lei, J.; Yang, S., SARS-CoV-2 M(pro) inhibitors with antiviral activity in a transgenic mouse model. *Science* **2021,** *371* (6536), 1374-1378.

10. Hoffman, R. L.; Kania, R. S.; Brothers, M. A.; Davies, J. F.; Ferre, R. A.; Gajiwala, K. S.; He, M.; Hogan, R. J.; Kozminski, K.; Li, L. Y.; Lockner, J. W.; Lou, J.; Marra, M. T.; Mitchell, L. J., Jr.; Murray, B. W.; Nieman, J. A.; Noell, S.; Planken, S. P.; Rowe, T.; Ryan, K.; Smith, G. J., 3rd; Solowiej, J. E.; Steppan, C. M.; Taggart, B., Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. *J Med Chem* **2020,** *63* (21), 12725-12747.

11. Ma, C.; Sacco, M. D.; Hurst, B.; Townsend, J. A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M. T.; Chen, Y.; Wang, J., Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res* **2020,** *30* (8), 678-692.

12. Sacco, M. D.; Ma, C.; Lagarias, P.; Gao, A.; Townsend, J. A.; Meng, X.; Dube, P.; Zhang, X.; Hu, Y.; Kitamura, N.; Hurst, B.; Tarbet, B.; Marty, M. T.; Kolocouris, A.; Xiang, Y.; Chen, Y.; Wang, J., Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against M(pro) and cathepsin L. *Sci Adv* **2020,** *6* (50), eabe0750.

13. Yang, K. S.; Ma, X. R.; Ma, Y.; Alugubelli, Y. R.; Scott, D. A.; Vatansever, E. C.; Drelich, A. K.; Sankaran, B.; Geng, Z. Z.; Blankenship, L. R.; Ward, H. E.; Sheng, Y. J.; Hsu, J. C.; Kratch, K. C.; Zhao, B.; Hayatshahi, H. S.; Liu, J.; Li, P.; Fierke, C. A.; Tseng, C. K.; Xu, S.; Liu, W. R., A Quick Route to Multiple Highly Potent SARS-CoV-2 Main Protease Inhibitors*. *ChemMedChem* **2021,** *16* (6), 942-948.

14. Ma, X. R.; Alugubelli, Y. R.; Ma, Y.; Vatansever, E. C.; Scott, D. A.; Qiao, Y.; Yu, G.; Xu, S.; Liu, W. R., MPI8 is Potent against SARS-CoV-2 by Inhibiting Dually and Selectively the SARS-CoV-2 Main Protease and the Host Cathepsin L*. *ChemMedChem* **2021**, 10.1002/cmdc.202100456.

15. Douangamath, A.; Fearon, D.; Gehrtz, P.; Krojer, T.; Lukacik, P.; Owen, C. D.; Resnick, E.; Strain-Damerell, C.; Aimon, A.; Abranyi-Balogh, P.; Brandao-Neto, J.; Carbery, A.; Davison, G.; Dias, A.; Downes, T. D.; Dunnett, L.; Fairhead, M.; Firth, J. D.; Jones, S. P.; Keeley, A.; Keseru, G. M.; Klein, H. F.; Martin, M. P.; Noble, M. E. M.; O'Brien, P.; Powell, A.; Reddi, R. N.; Skyner, R.; Snee, M.; Waring, M. J.; Wild, C.; London, N.; von Delft, F.; Walsh, M. A., Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nat Commun* **2020,** *11* (1), 5047.

16. Ghahremanpour, M. M.; Tirado-Rives, J.; Deshmukh, M.; Ippolito, J. A.; Zhang, C. H.; Cabeza de Vaca, I.; Liosi, M. E.; Anderson, K. S.; Jorgensen, W. L., Identification of 14 Known Drugs as Inhibitors of the Main Protease of SARS-CoV-2. *ACS Med Chem Lett* **2020,** *11* (12), 2526-2533.

17. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020,** *582* (7811), 289-293.

18. Jin, Z.; Zhao, Y.; Sun, Y.; Zhang, B.; Wang, H.; Wu, Y.; Zhu, Y.; Zhu, C.; Hu, T.; Du, X.; Duan, Y.; Yu, J.; Yang, X.; Yang, X.; Yang, K.; Liu, X.; Guddat, L. W.; Xiao, G.; Zhang, L.; Yang, H.; Rao, Z., Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nat Struct Mol Biol* **2020,** *27* (6), 529-532.

19. Ma, C.; Hu, Y.; Townsend, J. A.; Lagarias, P. I.; Marty, M. T.; Kolocouris, A.; Wang, J., Ebselen, Disulfiram, Carmofur, PX-12, Tideglusib, and Shikonin Are Nonspecific Promiscuous SARS-CoV-2 Main Protease Inhibitors. *ACS Pharmacol Transl Sci* **2020,** *3* (6), 1265-1277.

20. Rathnayake, A. D.; Zheng, J.; Kim, Y.; Perera, K. D.; Mackin, S.; Meyerholz, D. K.; Kashipathy, M. M.; Battaile, K. P.; Lovell, S.; Perlman, S.; Groutas, W. C.; Chang, K. O., 3C-like protease inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV-infected mice. *Sci Transl Med* **2020**, 12(557), eabc5332.

21. Zhu, W.; Xu, M.; Chen, C. Z.; Guo, H.; Shen, M.; Hu, X.; Shinn, P.; Klumpp-Thomas, C.; Michael, S. G.; Zheng, W., Identification of SARS-CoV-2 3CL Protease Inhibitors by a Quantitative High-Throughput Screening. *ACS Pharmacol Transl Sci* **2020,** *3* (5), 1008-1016.

22. Gunther, S.; Reinke, P. Y. A.; Fernandez-Garcia, Y.; Lieske, J.; Lane, T. J.; Ginn, H. M.; Koua, F. H. M.; Ehrt, C.; Ewert, W.; Oberthuer, D.; Yefanov, O.; Meier, S.; Lorenzen, K.; Krichel, B.; Kopicki, J. D.; Gelisio, L.; Brehm, W.; Dunkel, I.; Seychell, B.; Gieseler, H.; Norton-Baker, B.; Escudero-Perez, B.; Domaracky, M.; Saouane, S.; Tolstikova, A.; White, T. A.; Hanle, A.; Groessler, M.; Fleckenstein, H.; Trost, F.; Galchenkova, M.; Gevorkov, Y.; Li, C.; Awel, S.; Peck, A.; Barthelmess, M.; Schlunzen, F.; Lourdu Xavier, P.; Werner, N.; Andaleeb, H.; Ullah, N.; Falke, S.; Srinivasan, V.; Franca, B. A.; Schwinzer, M.; Brognaro, H.; Rogers, C.; Melo, D.; Zaitseva-Doyle, J. J.; Knoska, J.; Pena-Murillo, G. E.; Mashhour, A. R.; Hennicke, V.; Fischer, P.; Hakanpaa, J.; Meyer, J.; Gribbon, P.; Ellinger, B.; Kuzikov, M.; Wolf, M.; Beccari, A. R.; Bourenkov, G.; von Stetten, D.; Pompidor, G.; Bento, I.; Panneerselvam, S.; Karpics, I.; Schneider, T. R.; Garcia-Alai, M. M.; Niebling, S.; Gunther, C.; Schmidt, C.; Schubert, R.; Han, H.; Boger, J.; Monteiro, D. C. F.; Zhang, L.; Sun, X.; Pletzer-Zelgert, J.; Wollenhaupt, J.; Feiler, C. G.; Weiss, M. S.; Schulz, E. C.; Mehrabi, P.; Karnicar, K.; Usenik, A.; Loboda, J.; Tidow, H.; Chari, A.; Hilgenfeld, R.; Uetrecht, C.; Cox, R.; Zaliani, A.; Beck, T.; Rarey, M.; Gunther, S.; Turk, D.; Hinrichs, W.; Chapman, H. N.; Pearson, A. R.; Betzel, C.; Meents, A., X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. *Science* **2021,** *372* (6542), 642-646.

23. Hattori, S. I.; Higashi-Kuwata, N.; Hayashi, H.; Allu, S. R.; Raghavaiah, J.; Bulut, H.; Das, D.; Anson, B. J.; Lendy, E. K.; Takamatsu, Y.; Takamune, N.; Kishimoto, N.; Murayama, K.; Hasegawa, K.; Li, M.; Davis, D. A.; Kodama, E. N.; Yarchoan, R.; Wlodawer, A.; Misumi, S.; Mesecar, A. D.; Ghosh, A. K.; Mitsuya, H., A small molecule compound with an indole moiety inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat Commun* **2021,** *12* (1), 668.

24. Kuo, C. J.; Chao, T. L.; Kao, H. C.; Tsai, Y. M.; Liu, Y. K.; Wang, L. H.; Hsieh, M. C.; Chang, S. Y.; Liang, P. H., Kinetic characterization and inhibitor screening for the proteases leading to identification of drugs against SARS-CoV-2. *Antimicrob Agents Chemother* **2021**, 65(4), e02577-20.

25. Lockbaum, G. J.; Reyes, A. C.; Lee, J. M.; Tilvawala, R.; Nalivaika, E. A.; Ali, A.; Kurt Yilmaz, N.; Thompson, P. R.; Schiffer, C. A., Crystal Structure of SARS-CoV-2 Main Protease in Complex with the Non-Covalent Inhibitor ML188. *Viruses* **2021,** *13* (2), 174.

26. Zaidman, D.; Gehrtz, P.; Filep, M.; Fearon, D.; Gabizon, R.; Douangamath, A.; Prilusky, J.; Duberstein, S.; Cohen, G.; Owen, C. D.; Resnick, E.; Strain-Damerell, C.; Lukacik, P.; Covid-Moonshot, C.; Barr, H.; Walsh, M. A.; von Delft, F.; London, N., An automatic pipeline for the design of irreversible derivatives identifies a potent SARS-CoV-2 M(pro) inhibitor. *Cell Chem Biol* **2021**, 10.1016/j.chembiol.2021.05018.

27. Zhang, C. H.; Stone, E. A.; Deshmukh, M.; Ippolito, J. A.; Ghahremanpour, M. M.; Tirado-Rives, J.; Spasov, K. A.; Zhang, S.; Takeo, Y.; Kudalkar, S. N.; Liang, Z.; Isaacs, F.; Lindenbach, B.; Miller, S. J.; Anderson, K. S.; Jorgensen, W. L., Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations. *ACS Cent Sci* **2021,** *7* (3), 467-475.

28. Vatansever, E. C.; Yang, K. S.; Drelich, A. K.; Kratch, K. C.; Cho, C. C.; Kempaiah, K. R.; Hsu, J. C.; Mellott, D. M.; Xu, S.; Tseng, C. K.; Liu, W. R., Bepridil is potent against SARS-CoV-2 in vitro. *Proc Natl Acad Sci U S A* **2021,** *118* (10), e2012201118.

29. Cao, W.; Cho, C.-C. D.; Geng, Z. Z.; Ma, X. R.; Allen, R.; Shaabani, N.; Vatansever, E. C.; Alugubelli, Y. R.; Ma, Y.; Ellenburg, W. H.; Yang, K. S.; Qiao, Y.; Ji, H.; Xu, S.; Liu, W. R., Cellular Activities of SARS-CoV-2 Main Protease

Inhibitors Reveal Their Unique Characteristics. *bioRxiv* **2021**, 2021.06.08.447613.

O-*tert*-butyl-threonine at P3 site is a key component
to achieve high cellular and antiviral potency

SARS-CoV-2 M^{Pro} inhibitors