



OPEN

## Non-covalent SARS-CoV-2 M<sup>Pro</sup> inhibitors developed from in silico screen hits

Giacomo G. Rossetti<sup>1,2</sup>, Marianna A. Ossorio<sup>1</sup>, Stephan Rempel<sup>2</sup>, Annika Kratzel<sup>3,4,5</sup>, Vasilis S. Dionellis<sup>1</sup>, Samia Barriot<sup>1</sup>, Laurence Tropa<sup>1</sup>, Christoph Gorgulla<sup>6,7,8</sup>, Haribabu Arthanari<sup>6,8</sup>, Volker Thiel<sup>3,4</sup>, Peter Mohr<sup>9</sup>, Remo Gamboni<sup>9</sup> & Thanos D. Halazonetis<sup>1,2</sup>✉

M<sup>Pro</sup>, the main protease of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is essential for the viral life cycle. Accordingly, several groups have performed in silico screens to identify M<sup>Pro</sup> inhibitors that might be used to treat SARS-CoV-2 infections. We selected more than five hundred compounds from the top-ranking hits of two very large in silico screens for on-demand synthesis. We then examined whether these compounds could bind to M<sup>Pro</sup> and inhibit its protease activity. Two interesting chemotypes were identified, which were further evaluated by characterizing an additional five hundred synthesis on-demand analogues. The compounds of the first chemotype denatured M<sup>Pro</sup> and were considered not useful for further development. The compounds of the second chemotype bound to and enhanced the melting temperature of M<sup>Pro</sup>. The most active compound from this chemotype inhibited M<sup>Pro</sup> in vitro with an IC<sub>50</sub> value of 1 μM and suppressed replication of the SARS-CoV-2 virus in tissue culture cells. Its mode of binding to M<sup>Pro</sup> was determined by X-ray crystallography, revealing that it is a non-covalent inhibitor. We propose that the inhibitors described here could form the basis for medicinal chemistry efforts that could lead to the development of clinically relevant inhibitors.

In December 2019, the Chinese health authorities reported the emergence of a disease, now referred to as COVID-19, that was declared a pandemic by the World Health Organization just a few months later<sup>1</sup>. The causative agent of COVID-19 is a novel coronavirus named SARS-CoV-2<sup>2</sup>.

Similar to other coronaviruses, SARS-CoV-2 is an enveloped, positive-sense, single-stranded RNA virus. Its genome contains at least nine open reading frames (ORFs), of which ORF1a and ORF1b give rise to polyproteins pp1a and pp1ab, respectively<sup>3</sup>. A main protease (M<sup>Pro</sup>, also called 3C-like protease) and a papain-like protease (PL<sup>Pro</sup>) cleave these polyproteins into characteristic nonstructural proteins<sup>4</sup>. M<sup>Pro</sup> is essential for the viral life cycle, as it is needed to generate the mature forms of most of the nonstructural proteins of the virus; it contains a cysteine-histidine dyad at its catalytic center and cleaves its substrates at sites comprising a glutamine followed by a residue with a small side chain. M<sup>Pro</sup> is highly conserved within the coronavirus family and is considered a good target for the development of drugs that could treat COVID-19 patients<sup>5-7</sup>.

Most efforts to identify M<sup>Pro</sup> inhibitors have focused on repurposing or modifying inhibitors of M<sup>Pro</sup>-related proteases<sup>8-19</sup>. Among the most potent of these inhibitors are GC376, boceprevir/telaprevir analogues and the PF-00835231 and PF-07321332 compounds; they are all irreversible inhibitors that form a covalent bond with the thiol group of the catalytically important residue Cys145. GC376, originally developed as a Norwalk virus 3CL<sup>Pro</sup> inhibitor<sup>20,21</sup>, inhibits SARS-CoV-2 M<sup>Pro</sup> in vitro with a half maximal inhibitory concentration (IC<sub>50</sub>) of 30 ± 8 nM<sup>12,13</sup>. Analogues of boceprevir and telaprevir, two inhibitors approved for the treatment of hepatitis C, inhibit M<sup>Pro</sup> with IC<sub>50</sub> values as low as 7.6 ± 0.1 nM in vitro and show activity in a mouse model of SARS-CoV-2

<sup>1</sup>Department of Molecular Biology, University of Geneva, 1205 Geneva, Switzerland. <sup>2</sup>FoRx Therapeutics AG, 4056 Basel, Switzerland. <sup>3</sup>Institute of Virology and Immunology, University of Bern, 3012 Bern, Switzerland. <sup>4</sup>Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland. <sup>5</sup>Graduate School for Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland. <sup>6</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Harvard University, Boston, MA 02115, USA. <sup>7</sup>Department of Physics, Faculty of Arts and Sciences, Harvard University, Cambridge, MA 02138, USA. <sup>8</sup>Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA. <sup>9</sup>NANDASI Pharma Advisors, 4123 Allschwil, Switzerland. ✉email: thanos.halazonetis@unige.ch

infection<sup>22</sup>. Finally, PF-00835231 and PF-07321332 are related compounds derived from an inhibitor of the M<sup>Pro</sup> of SARS-CoV-1<sup>14,15</sup>. PF-00835231 requires continuous intravenous infusion as a prodrug to achieve effective doses in the plasma of human patients<sup>23</sup>, while PF-07321332 can be administered orally and promises to have a significant impact on the course of the COVID-19 pandemic<sup>15</sup>.

De novo M<sup>Pro</sup> inhibitors have been identified by either in silico or physical screens. Three in silico screening studies are particularly relevant here: two studies that ranked more than one billion compounds each, but did not validate the identified hits<sup>24,25</sup>, and a third study that ranked 6.5 million compounds and validated seven top compounds, of which the most potent exhibited an IC<sub>50</sub> of 4.2 μM in vitro<sup>26</sup>. Among the studies that employed physical screens to identify M<sup>Pro</sup> inhibitors, one study screened a small library of chemical fragments for binding to M<sup>Pro</sup> by X-ray crystallography and identified several hits, which, however, were not developed further<sup>27</sup>. A second study screened a DNA-encoded library and identified compounds that inhibited M<sup>Pro</sup> with inhibition constants in the 35–45 nM range<sup>28</sup>. These inhibitors bound to M<sup>Pro</sup> covalently, as seen with the repurposed inhibitors, presumably because reactive, electrophilic compounds were included in the DNA-encoded library.

Covalent inhibitors often require extensive development to limit off-target effects and associated toxicity. Therefore, the development of non-covalent M<sup>Pro</sup> inhibitors, as an alternate strategy, is also worth pursuing. Here we characterized selected hits from the two very large in silico screens cited above<sup>24,25</sup>. The vast majority of the examined hits did not bind to or inhibit M<sup>Pro</sup> in vitro. However, a handful of hits were validated and these served as starting points for the development of a non-covalent M<sup>Pro</sup> inhibitor that was capable of inhibiting SARS-CoV-2 replication in tissue culture cells.

## Results

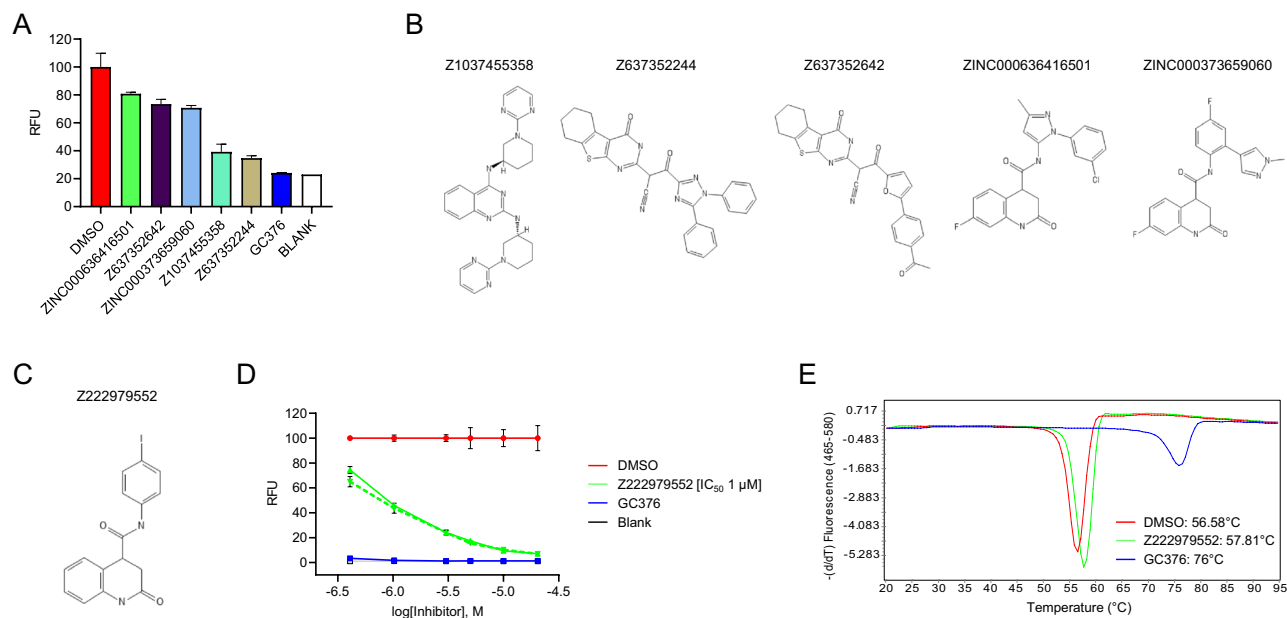
**Validation of putative SARS-CoV-2 M<sup>Pro</sup> inhibitors identified by in silico screens.** To validate candidate inhibitors of SARS-CoV-2 M<sup>Pro</sup> identified by in silico screens, we utilized an in vitro protease activity assay. An expression plasmid encoding the M<sup>Pro</sup> protease with a C-terminal His-tag<sup>8</sup> was introduced in the *E. coli* strain BL21-Gold (DE3) and recombinant M<sup>Pro</sup> was expressed and purified by affinity and size exclusion chromatography (Supplementary Fig. S1A). Protease activity was determined using a fluorescence resonance energy transfer (FRET) assay. The substrate peptide contained at its N-terminus a fluorescent dye (HiLyte-Fluor488) and at its C-terminus a quencher (QXL520); cleavage of the peptide by M<sup>Pro</sup> led to an increase in fluorescence intensity. The protease and fluorogenic substrate were assayed at 22 °C at final concentrations of 100 nM and 500 nM, respectively. Fluorescence intensity was measured every 10 min; in the absence of an M<sup>Pro</sup> inhibitor, fluorescence intensity increased linearly during the first 60 min of the reaction, whereas in the presence of the potent GC376 inhibitor<sup>12,13</sup> no increase in fluorescence intensity was observed (Supplementary Fig. S1B).

We examined hits from two in silico screening studies<sup>24,25</sup> that ranked the docking of molecules present in the REAL Space or ZINC chemical compound libraries<sup>29,30</sup>. In the first study, two in silico screens were performed, using two different three-dimensional structures to encompass the conformational flexibility of the active site of M<sup>Pro</sup><sup>24</sup>. The first screen, referred to as screen 1A, used the structure of M<sup>Pro</sup> described by Jin et al.<sup>10</sup> (pdb id: 6lu7) as target, while the second screen, referred to as screen 1B, used the structure of Dai et al.<sup>9</sup> (pdb id: 6m0k) with minor changes in the side-chain rotamers of residues Ser46, Met49 and Cys145 to capture a more open conformation of the active site. The two screens docked the same library of compounds; nevertheless, comparison of the top 1,000 hits of each screen, revealed an overlap of only 12 compounds. The second study performed one in silico screen, hereafter referred to as screen 2, of 1.3 billion compounds<sup>25</sup>, using the structure of M<sup>Pro</sup> described by Jin et al.<sup>10</sup> (pdb id: 6lu7) as target.

From the 3,808 top-ranking compounds of screen 1A, 195 compounds were manually selected for on-demand synthesis aiming for chemical diversity and drug-like features (Supplementary Table 1); whereas, from the 3,851 top-ranking compounds of screen 1B, 226 compounds were selected for on-demand synthesis (Supplementary Table 2). In addition, guided by the results of a crystallographic fragment screen<sup>27</sup> that showed a fragment containing a nitrile group deep in the active site of M<sup>Pro</sup> (pdb id: 5r82), we identified all the nitrile-containing compounds among the top 20,000 hits of screens 1A and 1B. This list included 253 compounds, 45 of which were selected for on-demand synthesis (Supplementary Table 3). Twelve of the 15 top-ranking compounds from screen 1A were also selected for on-demand synthesis (Supplementary Table 4). Finally, from screen 2, eight of the 15 top-ranking hits were selected for on-demand synthesis (Supplementary Table 4).

In total, we had 486 compounds synthesized and all these compounds were assayed at a final concentration of 40 μM for their ability to inhibit the protease activity of M<sup>Pro</sup>. Remarkably, only five compounds inhibited M<sup>Pro</sup> more than the pre-defined threshold level of 25% inhibition (Fig. 1A). The active compounds were: the diamino-quinazoline Z1037455358 (Fig. 1B), which is one of the 226 compounds selected from the 3,851 top-ranking compounds of screen 1B (Supplementary Table 2); the structurally-related nitriles Z637352244 and Z637352642 (Fig. 1B), which are two of the 45 nitrile-containing compounds selected from screens 1A and 1B (Supplementary Table 3); and the structurally-related dihydro-quinolinones ZINC000636416501 and ZINC000373659060 (Fig. 1B), which are two of the eight selected top-ranking compounds from screen 2 (Supplementary Table 4). All these five compounds were characterized further, as described below.

**Characterization of the diamino-quinazoline and nitrile compounds.** We first focused our efforts on compound Z1037455358, which contains a diamino-quinazoline core (Fig. 1B). The IC<sub>50</sub> of this compound in the protease assay was 26 μM (Supplementary Fig. S2A). Analogues of Z1037455358 were identified using similarity and substructure searches of the REAL Space library of chemical compounds<sup>29</sup>; 108 of these analogues were selected for on-demand synthesis (Supplementary Table 5); however, none of the analogues were more potent than the parent compound in the biochemical assay. Therefore, we retained only the parent compound for further analysis.



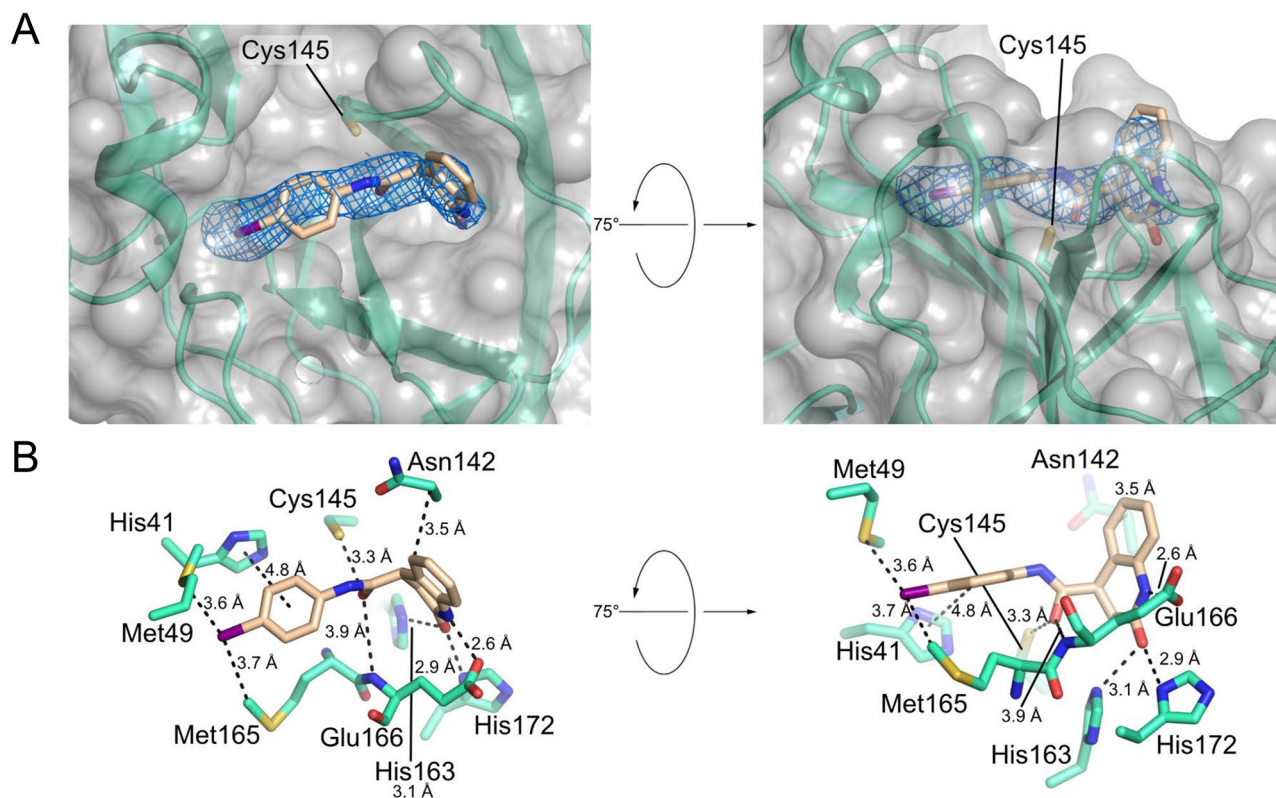
**Figure 1.** Development of  $M^{Pro}$  inhibitors from in silico screening hits. **(A)** Graph showing the inhibitory activity of the five validated in silico screening hits tested at a final concentration of 40  $\mu$ M in an in vitro protease-activity assay. After addition of the FRET substrate, fluorescence was acquired at 10 min intervals over 60 min. For each compound, the increase in fluorescence intensity was normalized to the DMSO control. GC376, a previously described  $M^{Pro}$  inhibitor, served as a positive control. Blank, reaction omitting  $M^{Pro}$ ; RFU, relative fluorescence units. **(B)** Chemical structure of the five validated compounds. **(C)** Chemical structure of Z222979552, the most active dihydro-quinolinone compound obtained after two rounds of chemical structure similarity searches of the REAL space library of molecules. **(D)** Dose-response curves for compound Z222979552 examined at 0.4, 1, 3, 5, 10 and 20  $\mu$ M final compound concentrations, in absence (solid lines) or presence (dashed lines) of 0.1  $\mu$ g protein lysate. DMSO, GC376 and blank controls are as described above. **(E)** Thermal shift assay performed in the presence of DMSO or 20  $\mu$ M of compounds Z222979552 or GC376. The graphs show the derivatives of the melting curves used to calculate the melting temperature of  $M^{Pro}$ .

Next, we examined the two structurally-related compounds Z637352244 and Z637352642, both of which contain a nitrile group (Fig. 1B). A nitrile group is also present in a fragment that was found to bind  $M^{Pro}$  by crystallographic screening<sup>27</sup>. To identify more potent compounds, we selected 301 analogues for on-demand synthesis (Supplementary Table 6) and examined them at a final concentration of 40  $\mu$ M for their ability to inhibit  $M^{Pro}$  in the protease assay. Five analogues were more active than the original compounds (Supplementary Fig. S2B). We determined the  $IC_{50}$  concentrations of the original compounds and of these five analogues. The parent compounds Z637352244 and Z637352642 exhibited  $IC_{50}$  values of 22 and 223  $\mu$ M, respectively (Supplementary Fig. S2C). Three analogues, Z56785964, Z637450230 and Z56786187, had  $IC_{50}$  values between 13–24  $\mu$ M, whereas two analogues, Z2239054061 and Z637352638, had  $IC_{50}$  values of 6.7 and 7.5  $\mu$ M, respectively (Supplementary Fig. S2C), representing a significant improvement over the parent compounds.

To further characterize the diamino-quinazoline and nitrile compounds, we examined their effect on the melting temperature of  $M^{Pro}$  using a thermal shift assay (TSA). Briefly,  $M^{Pro}$ , at a final concentration of 1  $\mu$ M, was incubated for 20 min with the inhibitors, at a final concentration of 20  $\mu$ M, and the melting temperature of  $M^{Pro}$  was determined. Compounds that bind to  $M^{Pro}$  should enhance its melting temperature<sup>31</sup>. Indeed, GC376 increased the melting temperature of  $M^{Pro}$  by 19  $^{\circ}$ C (Supplementary Fig. S3A). Surprisingly, the parent nitrile-containing compounds decreased the melting temperature of  $M^{Pro}$ , as did all their analogues, except for analogue Z637450230, which did not affect the melting temperature of  $M^{Pro}$  (Supplementary Fig. S3A). The diamino-quinazoline compound Z1037455358 also decreased the melting temperature of  $M^{Pro}$  (Supplementary Fig. S3B).

The decrease in the melting temperature of  $M^{Pro}$  by the above compounds was of concern, as it could indicate that these compounds were destabilizing  $M^{Pro}$ . Compounds that denature proteins non-specifically should lose inhibitory activity in the presence of excess carrier protein, since the latter can serve as a sink for compound sequestration. Indeed, the parent nitrile-containing compound Z637352244, its analogues Z56786187 and Z637450230, and the diamino-quinazoline compound Z1037455358, all lost activity, when the protease assay was performed in the presence of 1  $\mu$ g cell lysate (Supplementary Fig. S2D). In contrast, GC376 maintained inhibitory activity in the presence of lysate and, interestingly, so did the dihydro-quinolinones compounds ZINC000636416501 and ZINC000373659060 (Supplementary Fig. S2D).

**Characterization of the dihydro-quinolinone compounds.** Compounds ZINC000636416501 and ZINC000373659060 are related to each other and contain a dihydro-quinolinone core (Fig. 1B). Encouraged by the fact that the activity of these two compounds was not affected by the presence of cell lysate, we obtained 157 analogues (Supplementary Table 7) and examined their ability to inhibit  $M^{Pro}$ . Three analogues were significantly



**Figure 2.** Crystal structure of  $M^{pro}$  in complex with the dihydro-quinolinone compound Z222979552. (A) Surface representation of SARS-CoV-2  $M^{pro}$  with secondary structure elements colored green and compound Z222979552 shown as a stick model. The omit electron density map corresponding to the ligand compound is shown ( $\sigma = 1.0$ ). (B) Stick model of compound Z222979552 and of select residues at the active site of  $M^{pro}$  showing the key interactions between the protease and its inhibitor.

more potent than the parent compounds (Supplementary Fig. S4A). Specifically, compounds Z228770960, Z393665558 and Z225602086 had  $IC_{50}$  values of 4, 6 and 7.4  $\mu M$ , respectively, whereas the parent compounds ZINC000373659060 and ZINC000636416501 had  $IC_{50}$  values of 58 and 93  $\mu M$ , respectively (Supplementary Fig. S4B). Importantly, all three analogues retained their inhibitory activity against  $M^{pro}$  in the presence of cell lysate (Supplementary Fig. S4C).

The parent ZINC000636416501 and ZINC000373659060 compounds and their three active analogues were then examined for their ability to modulate the melting temperature of  $M^{pro}$  using the thermal shift assay. The parent compounds did not affect the melting temperature of  $M^{pro}$  (Supplementary Fig. S4D). However, the analogues increased the melting temperature of  $M^{pro}$  with the most active analogue, Z228770960, inducing an increase of 1.2  $^{\circ}C$  (Supplementary Fig. S4E). These results are consistent with the analogues forming stable complexes with  $M^{pro}$ .

To identify even more potent compounds, we performed a second round of analogue synthesis, using the first-round analogues Z228770960, Z393665558 and Z225602086 as starting points for structure similarity searches. A total of 113 second-round analogues were selected for on-demand synthesis (Supplementary Table 8), of which four were as potent or more potent than the analogues from the first round (Fig. 1C and Supplementary Fig. S5A). Specifically, compounds Z222979552, Z228166018, Z222977344 and Z222978028 had  $IC_{50}$  values of 1.0, 1.6, 2.0 and 5.8  $\mu M$ , respectively, which inhibitory activity they retained in the presence of cell lysate (Fig. 1D and Supplementary Fig. S5B). The above four analogues also increased the melting temperature of  $M^{pro}$  by up to 1.2  $^{\circ}C$  (Fig. 1E and Supplementary Fig. S5C).

The first and second rounds of analogue synthesis resulted in about 15-fold and four-fold improvements in the  $IC_{50}$  values, respectively. However, additional chemical diversity of structures that we could gain from further rounds of compound similarity searches was limited. We therefore decided to characterize the most potent dihydro-quinolinone  $M^{pro}$  inhibitor, Z222979552 ( $IC_{50} = 1.0 \mu M$ ), by determining its crystal structure in complex with  $M^{pro}$  and its antiviral activity in cell-based assays.

**Structure of a dihydro-quinolinone inhibitor in complex with  $M^{pro}$ .** The crystal structure of  $M^{pro}$  in complex with compound Z222979552 was solved at a resolution of 2.5  $\text{\AA}$  and shows the compound in the active site of  $M^{pro}$  (Fig. 2A; Table 1). Unlike the previously described high-affinity  $M^{pro}$  inhibitors, compound Z222979552 did not form a covalent bond with  $M^{pro}$ . Rather, binding was mediated by hydrogen bonds, pi-stacking and hydrophobic interactions. Hydrogen bonds were observed between the dihydro-quinolinone group and the side chains of Glu166, His163 and His172, as well as between the carbonyl group of the compound and

	M <sup>Pro</sup> in complex with Z222979552 <sup>a</sup>
<b>Data collection</b>	
# Crystals/# datasets	1/1
Space group	C 1 2 1
Unit cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	114.61, 53.79, 45.44
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 100.98, 90.00
# of reflections	31,270 (2882)
# of unique reflections	9258 (912)
R <sub>merge</sub> (%)	0.17 (2.37)
CC <sub>1/2</sub>	0.98 (0.21)
I/ $\sigma$ I	8.38 (0.64)
Completeness	98.65 (97.33)
Redundancy	3.40 (3.20)
<b>Refinement</b>	
Resolution (Å)	48.53–2.50
No. of reflections	18,044 (1759)
R <sub>work</sub> /R <sub>free</sub>	0.2255/0.2783
Bond lengths (Å)	0.002
Bond angles (°)	0.45
No. of non-hydrogen atoms	2383
Protein	2354
Ligand	21
Solvent	8
<b>B-factors</b>	
Protein	87.30
Ligand	138.11
Solvent	74.72

**Table 1.** X-ray diffraction data and phasing and refinement statistics. <sup>a</sup>Values in parentheses are for the highest resolution shell.

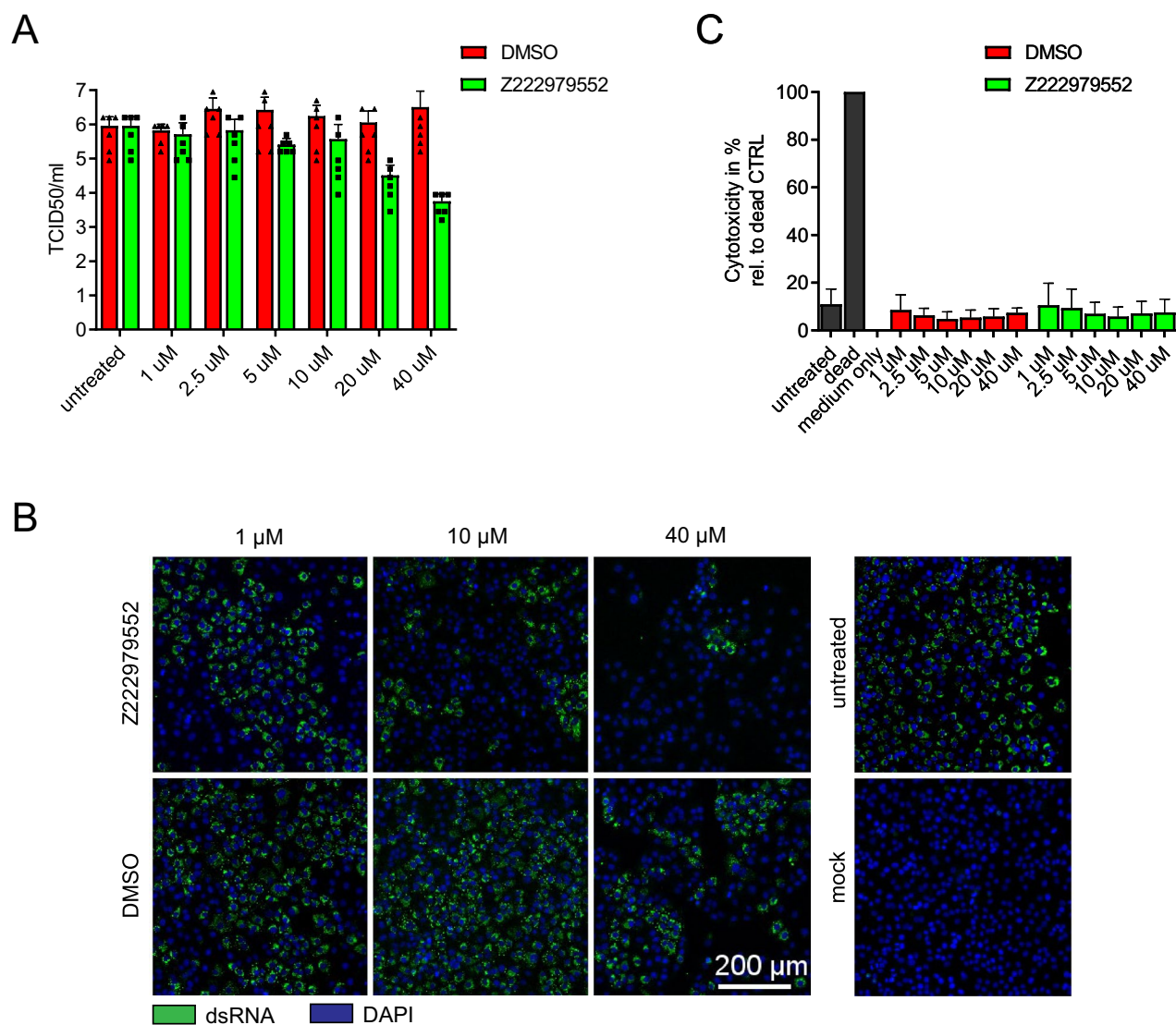
the thiol group of Cys145 and the main chain of Glu166 (Fig. 2B). A T-type pi-stacking interaction was observed between the benzene ring and His41, while the dihydro-quinolinone group, the benzene group and the iodine atom participated in hydrophobic interactions with Asn142, Met49 and Met165, respectively (Fig. 2B).

**Inhibition of viral replication by a dihydro-quinolinone M<sup>Pro</sup> inhibitor.** To further characterize compound Z222979552, we examined whether it could inhibit SARS-CoV-2 replication in Vero E6 cells. Consistent with its ability to inhibit M<sup>Pro</sup> in vitro, Z222979552 suppressed SARS-CoV-2 replication in Vero E6 cells, resulting in a more than a 100-fold decrease of SARS-CoV-2 titers at the highest concentration tested (Fig. 3A). In a second assay we monitored by immunofluorescence the presence of viral double-stranded RNA (dsRNA) in Vero E6 cells infected with SARS-CoV-2. Treatment of the cells with Z222979552 prevented the formation of dsRNA intermediates of SARS-CoV-2 RNA synthesis (Fig. 3B). We also performed a cytotoxicity assay with non-infected Vero E6 cells and observed that Z222979552 is not cytotoxic (Fig. 3C). Finally, we validated the cell-based assays using remdesivir as a positive control (Supplementary Fig. S6). Taken together, the above findings indicate that compound Z222979552 has antiviral activity in cells.

## Discussion

COVID-19 has had a significant impact on our society. The vaccines have significantly curtailed the spread of the pandemic<sup>32</sup>, but challenges still remain regarding vaccine acceptance by the public and durability of vaccine efficacy over time<sup>33,34</sup>. Thus, there continues to be an urgent need for novel medicines to treat COVID-19 and M<sup>Pro</sup> is considered a valuable target for the development of SARS-CoV-2 antivirals, because it is required for viral replication<sup>5–7</sup>. Equally importantly, M<sup>Pro</sup> is highly conserved in evolution<sup>5</sup>, which means that inhibitors developed against M<sup>Pro</sup> of one coronavirus may inhibit replication of multiple members of the coronavirus family. Indeed, SARS-CoV-2 M<sup>Pro</sup> inhibitors entering the clinic were derived from compounds that were originally developed as inhibitors of SARS-CoV-1 M<sup>Pro</sup>, but whose development was, unfortunately, halted, when the SARS-CoV-1 epidemic waned<sup>14,15,23</sup>.

All the inhibitors described here were obtained from the REAL Space or the ZINC chemical compound libraries, which together encompass more than 20 billion make-on-demand compounds<sup>29,30</sup>. Despite their large size, these two libraries represent a very small part of the universe of chemical space. Therefore, we anticipate that the compounds that we have identified can serve as starting points for medicinal chemistry efforts, particularly



**Figure 3.** Compound Z222979552 reduces SARS-CoV-2 replication in Vero E6 cells in a dose-dependent manner. **(A)** SARS-CoV-2 titers upon treatment with compound Z222979552 or DMSO control. Viral titers are displayed as fifty-percent tissue culture infective dose (TCID<sub>50</sub>)/ml 24 h post-infection. TCID<sub>50</sub> values correspond to the viral titers (log<sub>10</sub> scale) required to kill fifty-percent of infected host cells. The experiment was performed in triplicate and the results are shown as means and standard deviations. **(B)** Immunofluorescence staining of infected Vero E6 cells for double-stranded RNA (dsRNA). The cells were treated with compound Z222979552 or DMSO and were stained 24 h post-infection. Green, dsRNA; blue, DAPI. One representative image out of three biological replicates is shown. **(C)** Z222979552-mediated cytotoxicity, determined using Vero E6 cells treated with the compound or DMSO for 24 h. The experiment was performed in triplicate and the results are shown as means and standard deviations.

because these compounds have drug-like features and activity in cell-based assays. Moreover, their mechanism of action is understood from the crystal structure of compound Z222979552 in complex with M<sup>Pro</sup>.

One interesting feature of the inhibitors described here is that they are non-covalent, in contrast to the vast majority of previously described M<sup>Pro</sup> inhibitors, including the PF-07321332 inhibitor, which form covalent bonds with the catalytic cysteine. The therapeutic significance of having a non-covalent M<sup>Pro</sup> inhibitor is unclear at this moment, but the possibility of developing such inhibitors to complement the covalent inhibitors can only offer more therapeutic opportunities.

## Methods

**Protein expression and purification of SARS-CoV-2 M<sup>Pro</sup>.** The M<sup>Pro</sup> construct<sup>8</sup> provided by Rolf Hilgenfeld was transformed into *E. coli* strain BL21-Gold (DE3) (Agilent). Transformed clones were picked to prepare pre-starter cultures in 2 mL YT medium with ampicillin (100 μg/ml), at 37 °C for 8 h. The pre-starter culture was then inoculated into fresh 120 mL YT medium with ampicillin (100 μg/ml) and incubated at 37 °C

overnight. The next day, the starter culture was inoculated into 1,600 mL YT medium with ampicillin (100 µg/ml) and incubated at 37 °C until OD<sub>600</sub> reached a value between 0.6 and 0.8. 1 mM isopropyl-D-thiogalactoside (IPTG) was then added to induce the overexpression of M<sup>Pro</sup> at 30 °C for 5 h. The bacteria were harvested by centrifugation at 8260 × g, 4 °C for 15 min, resuspended in Binding Buffer (25 mM BTP [pH 6.8]; 300 mM NaCl; 2 mM DTT; 1 mM EDTA; 3% DMSO) and then lysed using an Emulsiflex-C3 homogenizer (Avestin). The lysate was clarified by ultracentrifugation at 137,088 × g, 4 °C for 1 h and loaded onto a HisTrap FF column (Cytiva) using an Äkta protein purification system (Cytiva). When all the supernatant containing M<sup>Pro</sup> had passed through the column, the column was washed with 80 mL binding buffer to remove non-specifically bound proteins and then M<sup>Pro</sup> was eluted using an imidazole gradient (0–500 mM) in Binding Buffer. The M<sup>Pro</sup> fractions were concentrated using 3 kDa Amicon Ultra Centrifugal Filters (Merck Millipore) and the M<sup>Pro</sup> protein was further purified by size exclusion chromatography using a HiLoad Superdex 200 column (Cytiva) attached to a SMART protein purification system (Pharmacia).

**Compounds.** We initially selected for on-demand synthesis 485 compounds that were identified by two in silico screening studies as putative M<sup>Pro</sup> inhibitors; these compounds are referred to as parent compounds (Supplementary Tables 1–4). After evaluating the activity of the parent compounds in vitro, we selected an additional 686 compounds that were analogues of the few active parent compounds (Supplementary Tables 5–8). All the above compounds were purchased from Enamine, their purity was ≥ 90% and they were synthesized on-demand. The compounds were dissolved in DMSO at a concentration of 2 mM and were stored at -20 °C. GC376 was purchased from BPSBioscience.

**SARS-CoV-2 M<sup>Pro</sup> protease activity assay.** M<sup>Pro</sup> protease assays were performed in duplicate in Falcon 384-well optitux flat bottom, TC-treated microplates (Corning) in a final volume of 10 µL. M<sup>Pro</sup>, at a final concentration of 100 nM, was preincubated for 20 min at room temperature (RT) with the compounds in assay buffer (5 mM HEPES pH 7.5, 0.1 mg/mL BSA, 0.01% Triton, 2 mM DTT) under gentle agitation. The FRET substrate, HiLyte-Fluor488-ESATLQSGLRKAK(QXL520)-NH<sub>2</sub> (Eurogentec), was then added at a final concentration of 500 nM and incubated for 2 min at RT with gentle agitation prior to the start of fluorescence measurement. Compounds and FRET substrate were dispensed with an acoustic liquid dispenser (Gen5-Acoustic Transfer System; EDC Biosystems). The fluorescence intensity was measured kinetically for 7 cycles, every 10 min at 22 °C, using a Spark 10 M microplate reader (Tecan) and excitation and emission wavelengths of 485 and 528 nm, respectively.

**SARS-CoV-2 M<sup>Pro</sup> thermal shift binding assay.** Thermal shift assays were performed in duplicate in LightCycler 480 multiwell plates 96, white (Roche) in a final volume of 20 µL. M<sup>Pro</sup> protease, at a final concentration of 1 µM, was preincubated for 20 min at room temperature (RT) under gentle agitation with the compounds (final concentration: 20 µM) in assay buffer (10 mM HEPES pH 7.5; 150 mM NaCl). Protein unfolding was monitored with 5X SYPRO Orange (Sigma) binding dye. Compounds and SYPRO Orange were dispensed with an acoustic liquid dispenser (Gen5-Acoustic Transfer System; EDC Biosystems). Fluorescence (excitation wavelength: 465 nm; emission wavelength: 580 nm) was measured over a temperature gradient ranging from 20 to 95 °C, with incremental steps of 0.05 °C/s and 11 acquisitions per °C. The melting curves and peaks were obtained using the melting temperature (T<sub>m</sub>) calling analysis of the LightCycler 480 Software (release 1.5.1.62; Roche; [https://lifescience.roche.com/en\\_ch/brands/realtime-pcr-overview.html#software](https://lifescience.roche.com/en_ch/brands/realtime-pcr-overview.html#software)).

**Crystallization and model building.** Purified M<sup>Pro</sup> was concentrated to 5.6 mg/mL and crystallized using the hanging vapor diffusion method. Plate-shaped crystals grew in star-like clusters within one week in drops mixed in a 1:1 ratio with reservoir solution (0.1 M Tris [pH 8.0], 25% (v/v) PEG-3350 and 0.2 M LiCl) at 293 K. For soaking and cryoprotection, the crystal clusters were moved into 2 µL reservoir solution supplemented with 25% ethylene glycol. Then, 0.4 µL inhibitor solution was added resulting in a final compound concentration of 10 mM. Soaking was allowed to proceed for 2 h, after which single crystals were broken off the clusters and flash-frozen in liquid nitrogen. Diffraction data was collected at the Swiss Light Source (SLS), beamline PXI (X06SA) with λ = 1.000 Å at T = 100 K. The data were processed using the XDS package<sup>35</sup> up to a resolution of 2.5 Å. The structure was solved using Phaser<sup>35</sup> from the Phenix suite (v.19.1.) with 6wqf as a search model. In Phenix, the model was iteratively refined using Phenix Refine<sup>36</sup> with manual adjustments done in Coot6 (v.0.9.3). The three-dimensional model and molecular restraints of the ligand were generated using eLBOW7 from the Phenix suite<sup>37</sup>. According to the Ramachandran statistics for the final model 95.38% of residues are in favored regions, 3.96% are in allowed regions, and 0.77% of residues are outliers. All figures concerning structural data were prepared with OpenSource Pymol by Schrodinger (v.1.20; <https://github.com/schrodinger/pymol-open-source>).

**Viral replication assays.** Vero E6 cells (kindly provided by Doreen Muth, Marcel Müller and Christian Drosten, Charité, Berlin, Germany) were propagated in Dulbecco's modified EMEM (DMEM), supplemented with 10% heat-inactivated fetal bovine serum, 1% nonessential amino acids, 100 µg/mL streptomycin, 100 IU/mL penicillin and 15 mM HEPES at 37 °C in a humidified incubator with 5% CO<sub>2</sub>. SARS-CoV-2 (SARS-CoV-2/München-1.1/2020/929, kindly provided by Daniela Niemeyer, Marcel Müller and Christian Drosten) passage 1 was used for infection of the cells.

Two million Vero E6 cells were plated per well of a 96-well plate; 24 h later, the cells were infected with SARS-CoV-2/München-1.1/2020/929 passage 1 at an MOI of 0.1 for 1 h at 37 °C and then washed 3 times with PBS. Z222979552 and remdesivir (or respective volumes of DMSO) were added to cells in following concentrations: 0, 1, 2.5, 5, 10, 20 and 40 µM. 24 h post-infection infectious supernatant was serially diluted and the 50% tissue

culture infectious dose (TCID<sub>50</sub>) per ml was determined 96 h later using the Spearman-Kärber algorithm, as previously described<sup>38</sup>. Cytotoxic effects of Z222979552, remdesivir or the corresponding DMSO volumes were determined using the CytoTox 96 Non-Radioactive Cytotoxicity Assay (Promega).

Vero E6 cells infected with SARS-CoV-2 were fixed with 4% formalin and then permeabilized in PBS supplemented with 50 mM NH<sub>4</sub>Cl, 0.1% (w/v) Saponin and 2% (w/v) Bovine Serum Albumin. The cells were stained using a mouse monoclonal antibody against dsRNA (SCICONS, clone J2) and Alexa-Fluor 488-labeled donkey-anti mouse IgG (H + L) (JacksonImmuno) secondary antibody. Images were acquired with an EVOS FL Auto 2 Imaging System, using a 10× objective lens and processed using Fiji software packages<sup>39</sup> version 1.53j (<https://imagej.net/software/fiji/downloads>) and assembled with the FigureJ plugin<sup>40</sup> version 1.36 (<https://imagejdocu.list.lu/plugin/utilities/figurej/start>).

## Data availability

Atomic coordinates and structure factors for the crystal structure of M<sup>pro</sup> in complex with Z222979552 have been deposited in the Protein Data Bank under accession code 7P2G.

Received: 18 November 2021; Accepted: 24 January 2022

Published online: 15 February 2022

## References

- Zhu, N. *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
- Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **5**, 536–544 (2020).
- Chan, J. F. *et al.* Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg. Microbes Infect.* <https://doi.org/10.1080/22221751.2020.1719902> (2020).
- Chen, Y., Liu, Q. & Guo, D. Emerging coronaviruses: genome structure, replication, and pathogenesis. *J. Med. Virol.* **92**, 418–423 (2020).
- Ullrich, S. & Nitsche, C. The SARS-CoV-2 main protease as drug target. *Bioorg. Med. Chem. Lett.* **30**, 127377 (2020).
- Wu, C. *et al.* Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm. Sin. B* **10**, 766–788 (2020).
- Vandyck, K. & Deval, J. Considerations for the discovery and development of 3-chymotrypsin-like cysteine protease inhibitors targeting SARS-CoV-2 infection. *Curr. Opin. Virol.* **49**, 36–40 (2021).
- Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved  $\alpha$ -ketoamide inhibitors. *Science* **368**, 409–412 (2020).
- Dai, W. *et al.* Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **368**, 1331–1335 (2020).
- Jin, Z. *et al.* Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).
- Jin, Z. *et al.* Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nat. Struct. Mol. Biol.* **27**, 529–532 (2020).
- Vuong, W. *et al.* Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat. Commun.* **11**, 4282 (2020).
- Ma, C. *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 (2020).
- Hoffman, R. L. *et al.* Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19. *J. Med. Chem.* **63**, 12725–12747 (2020).
- Owen, D. R. *et al.* An oral SARS-CoV-2 M<sup>pro</sup> inhibitor clinical candidate for the treatment of COVID-19. *Science* <https://doi.org/10.1126/science.aba4784> (2021).
- Sacco, M. D. *et al.* Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against Mpro and cathepsin L. *Sci. Adv.* **6**, 1005 (2020).
- Riva, L. *et al.* Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature* **586**, 113–119 (2020).
- Günther, S. *et al.* X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. *Science* **372**, 642–646 (2021).
- Zaidman, D. *et al.* An automatic pipeline for the design of irreversible derivatives identifies a potent SARS-CoV-2 M<sup>pro</sup> inhibitor. *Cell Chem. Biol.* <https://doi.org/10.1016/j.chembiol.2021.05.018> (2021).
- Kim, Y. *et al.* Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. *J. Virol.* **86**, 11754–11762 (2012).
- Kim, Y. *et al.* Broad-spectrum inhibitors against 3C-like proteases of feline coronaviruses and feline caliciviruses. *J. Virol.* **89**, 4942–4950 (2015).
- Qiao, J. *et al.* SARS-CoV-2 Mpro inhibitors with antiviral activity in a transgenic mouse model. *Science* **371**, 1374–1378 (2021).
- de Vries, M. *et al.* A comparative analysis of SARS-CoV-2 antivirals characterizes 3CLpro inhibitor PF-00835231 as a potential new treatment for COVID-19. *J. Virol.* **2**, 10084 (2021).
- Gorgulla, C. *et al.* A multi-pronged approach targeting SARS-CoV-2 proteins using ultra-large virtual screening. *iScience* **24**, 102021 (2021).
- Ton, A., Gentile, F., Hsing, M., Ban, F. & Cherkasov, A. Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3 billion compounds. *Mol. Inf.* **39**, e2000028 (2020).
- Clyde, A. *et al.* High throughput virtual screening and validation of a SARS-CoV-2 main protease non-covalent inhibitor. *bioRxiv* <https://doi.org/10.1101/2021.03.27.437323> (2021).
- Douangamath, A. *et al.* Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nat. Commun.* **11**, 5047 (2020).
- Srinivas, C. S. *et al.* DNA-encoded chemistry technology yields expedient access to SARS-CoV-2 Mpro inhibitors. *PNAS* **118**, e2111172118 (2021).
- Enamine REAL. Available online: <https://enamine.net/library-synthesis/real-compounds>.
- Sterling, T. & Irwin, J. J. ZINC 15—ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).
- Lo, M.-C. *et al.* Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.* **332**, 153–159 (2004).
- Haas, E. J. *et al.* Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. *Lancet* **397**, 1819–1829 (2021).



33. Forman, R., Shah, S., Jeurissen, P., Jit, M. & Mossialos, E. COVID-19 vaccine challenges: What have we learned so far and what remains to be done?. *Health Policy* **125**, 553–567 (2021).
34. Falsey, A. R. *et al.* SARS-CoV-2 neutralization with BNT162b2 vaccine dose 3. *N. Engl. J. Med.* **15**, 10058 (2021).
35. Kabsch, W. X. D. S. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
36. Adams, P. D. *et al.* PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
37. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. Electronic ligand builder and optimization workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 1074–1080 (2009).
38. George, V. G., Hierholzer, J. C. & Ades, E. W. Cell culture. In *Virology Methods Manual* (eds Mahy, B. W. J. & Kangro, H. O.) 3–24 (Academic Press, 1996).
39. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
40. Mutterer, J. & Zinck, E. Quick-and-clean article figures with FigureJ. *J. Microsc.* **252**, 89–91 (2013).

## Acknowledgements

The authors thank Rolf Hilgenfeld for providing the plasmid expressing M<sup>pro</sup> in bacteria and the beam line scientists at the Swiss Light Source for help with collection of X-ray diffraction data. Funding for this project was provided by the Fondation Aclon and by overhead funds from the ERC Project ONIDDAC (Principal Investigator: T.D.H.).

## Author contributions

G.G.R. and M.O. performed the protease and thermal shift assays; S.R. determined the crystal structure of M<sup>pro</sup> with bound inhibitor; A.K. performed the viral replication assays; S.B. and L.T. expressed and purified M<sup>pro</sup> protein; C.G. and H.A. analyzed and provided in silico screening hits; P. M. and R. G. selected hits and analogues for on-demand synthesis; V.S.D., V.T., G.G.R. and T.D.H. designed and supervised experiments; T.D.H. conceived the study; all authors contributed to writing the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-06306-4>.

**Correspondence** and requests for materials should be addressed to T.D.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022