

1 **N-Terminal finger stabilizes the reversible feline drug GC376 in SARS-CoV-2 M<sup>pro</sup>**

2  
3  
4 Elena Arutyunova<sup>1,4</sup>, Muhammad Bashir Khan<sup>1</sup>, Conrad Fischer<sup>2†</sup>, Jimmy Lu<sup>1,4</sup>, Tess Lamer<sup>2</sup>,  
5 Wayne Vuong<sup>2</sup>, Marco J. van Belkum<sup>2</sup>, Ryan T. McKay<sup>2</sup>, D. Lorne Tyrrell<sup>3,4</sup>, John C. Vederas<sup>2</sup>,  
6 Howard S. Young<sup>1\*</sup>, M. Joanne Lemieux<sup>1,4\*</sup>

7  
8 <sup>1</sup> Department of Biochemistry, Faculty of Medicine and Dentistry, Membrane Protein Disease  
9 Research Group, University of Alberta, Edmonton T6G 2R3, Alberta, Canada

10 <sup>2</sup>Department of Chemistry, Faculty of Science, University of Alberta, Edmonton T6G 2G2,  
11 Alberta, Canada

12 <sup>3</sup>Department of Medical Microbiology and Immunology, Faculty of Medicine and Dentistry,  
13 University of Alberta, Edmonton T6G 2E1, Alberta, Canada

14 <sup>4</sup>Li Ka Shing Institute of Virology, Faculty of Medicine and Dentistry, University of Alberta,  
15 Edmonton T6G 2E1, Alberta, Canada

16 †Present address: Department of Physical Sciences, Barry University, 11300 NE 2nd Ave., Miami  
17 Shores FL 33161, USA

18  
19 \*Corresponding authors:

20 Howard S. Young: [hyoung@ualberta.ca](mailto:hyoung@ualberta.ca)

21 M. Joanne Lemieux: [mlemieux@ualberta.ca](mailto:mlemieux@ualberta.ca)

22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

## Abstract

47  
48  
49 The main protease ( $M^{\text{pro}}$ , also known as 3CL protease) of SARS-CoV-2 is a high priority drug  
50 target in the development of antivirals to combat COVID-19 infections. A feline coronavirus  
51 antiviral drug, GC376, has been shown to be effective in inhibiting the SARS-CoV-2 main  
52 protease and live virus growth. As this drug moves into clinical trials, further characterization of  
53 GC376 with the main protease of coronaviruses is required to gain insight into the drug's  
54 properties, such as reversibility and broad specificity. Reversibility is an important factor for  
55 therapeutic proteolytic inhibitors to prevent toxicity due to off-target effects. Here we demonstrate  
56 that GC376 has nanomolar  $K_i$  values with the  $M^{\text{pro}}$  from both SARS-CoV-2 and SARS-CoV  
57 strains. Restoring enzymatic activity after inhibition by GC376 demonstrates reversible binding  
58 with both proteases. In addition, the stability and thermodynamic parameters of both proteases  
59 were studied to shed light on physical chemical properties of these viral enzymes, revealing higher  
60 stability for SARS-CoV-2  $M^{\text{pro}}$ . The comparison of a new X-ray crystal structure of  $M^{\text{pro}}$  from  
61 SARS-CoV complexed with GC376 reveals similar molecular mechanism of inhibition compared  
62 to SARS-CoV-2  $M^{\text{pro}}$ , and gives insight into the broad specificity properties of this drug. In both  
63 structures, we observe domain swapping of the N-termini in the dimer of the  $M^{\text{pro}}$ , which facilitates  
64 coordination of the drug's P1 position. These results validate that GC376 is a drug with an off-rate  
65 suitable for clinical trials.

66

67 **Keywords:** 3CL protease, coronavirus, proteolytic inhibitor, COVID-19, antivirals, enzyme  
68 mechanism, SARS, GC373

69

70

## 71 **Introduction**

72           In late 2019, a respiratory infection initially detected in China, was sparking fear of a viral  
73 outbreak [1]. This respiratory infection attributed to severe acute respiratory syndrome coronavirus  
74 2 (SARS-CoV-2), led to an ongoing coronavirus disease 2019 (COVID-19) pandemic with  
75 millions infected worldwide (<https://coronavirus.jhu.edu/map.html>). This respiratory illness was  
76 similar to a previous infection by SARS-CoV that led to a SARS outbreak in 2002/3 as well as the  
77 Middle East respiratory infection (MERS) outbreak in 2012 [2,3]. All of these outbreaks stem from  
78 related betacoronavirus infections, suggesting these strains will likely lead to future viral  
79 outbreaks. Vaccines have been developed and will be important for prevention of new infections  
80 in the future. However, even with a 95% immunity rate, there will be a significant proportion of  
81 people worldwide who will require therapeutic treatment. Antiviral development remains a priority  
82 because of importance of immediate mitigation of acute infections, vaccine hesitancy, and the  
83 inability to vaccinate some individuals. The outbreak of SARS in 2003 and MERS in 2012 along  
84 with the current pandemic reminds us that pan-inhibitors may provide a means for initial control  
85 of outbreaks, thereby preventing or quickly controlling pandemics in the future [4].

86           SARS-CoV-2 is a 30-kb positive-sense single-stranded RNA virus that is translated by the  
87 host's cellular machinery to generate two alternatively spliced long polypeptides, PP1a and PP1ab.  
88 These long polypeptides release non-structural proteins (nsps), including the RNA-dependent  
89 RNA polymerase, that are essential for viral replication after proteolytic cleavage by proteases  
90 from domain nsp3 and nsp5, respectively, a papain-like (PL<sup>pro</sup>) protease and a chymotrypsin-like  
91 main protease (M<sup>pro</sup> or 3CL<sup>pro</sup>)[5]. Similar to SARS-CoV, the SARS-CoV-2 M<sup>pro</sup> enzyme  
92 recognises the sequence of Leu-Gln↓Ser-Ala-Gly, where ↓ marks the cleavage site and this  
93 sequence is widely employed for generation of substrates for kinetic analysis and for development

94 of peptidomimetic specific probes and inhibitors [6] [7,8]. The essential role of the M<sup>pro</sup> in viral  
95 replication has resulted in a great deal of crystallographic and *in silico* studies working towards  
96 the development of antiviral therapies to treat COVID-19 [9-13].

97 Proteolytic inhibitors have been used successfully as antiviral therapeutics [14]; for  
98 example peptidomimetic inhibitors for the human immunodeficiency virus (HIV) protease and  
99 small molecule inhibitors for hepatitis C virus (HCV) protease. The HIV protease inhibitors, along  
100 with other drugs, are used in a combination therapy and play a big role in the treatment of  
101 symptoms and the subsequent reduction in spread of infection.

102 It has been recently shown by our group, as well as by other teams, that M<sup>pro</sup> of SARS-  
103 CoV-2 is a promising drug target for the development of SARS-CoV-2 antivirals [10,11,13,15].  
104 We demonstrated that the proteolytic inhibitor GC376 (a bisulphite prodrug) used to treat feline  
105 coronavirus infection and its related aldehyde inhibitor, GC373, are effective at decreasing viral  
106 load of SARS-CoV-2 in cell culture [13]. These drugs have previously been shown to be effective  
107 inhibiting the M<sup>pro</sup> of picornavirus, norovirus and coronavirus, and furthermore have been  
108 validated in animal models for both SARS and MERS [16-18]. Even though we have a  
109 considerable understanding of the efficacy of GC376 and GC373 with both SARS-CoV and  
110 SARS-CoV-2 M<sup>pro</sup> [16,17,19-23], detailed mechanistic and functional insight into the inhibitor  
111 binding process is still essential for directing broad-spectrum inhibitors in clinical trials. For  
112 example, one of desirable features for peptidomimetic proteolytic inhibitors is the reversible nature  
113 of binding since it reduces the risk of strong off-target effects and potential toxicity [24,25]. In  
114 addition, in light of the new variants, we need a clear understanding of the efficacy of GC373 and  
115 GC376 with other coronavirus M<sup>pro</sup>, and importantly a crystal structure of these inhibitors with the  
116 SARS-CoV M<sup>pro</sup> has not been determined.

117 In this study, we compare inhibition of the M<sup>pro</sup> of SARS-CoV and SARS-CoV-2 by  
118 GC376 using kinetic and structural approaches. We determine K<sub>i</sub> values are in the low nanomolar  
119 range for both SARS-CoV and SARS-CoV-2 M<sup>pro</sup>. After inhibition with GC376, NMR and activity  
120 assays demonstrate the reversible nature of inhibition for both proteases. In addition, the  
121 restoration of activity of M<sup>pro</sup> after inhibition reveal a high kinetic and thermodynamic stability for  
122 these viral proteases. We determine the crystal structures of SARS-CoV M<sup>pro</sup> inhibited with the  
123 dipeptidyl inhibitor, GC376, and aldehyde form, GC373, both of which reveal a covalent mode of  
124 inhibition similar to SARS-CoV-2 M<sup>pro</sup>. We highlight in both structures the role of the N-terminus  
125 in stabilizing the S1 subsite from domain swapping, and how this facilitates drug binding. This  
126 comparative analysis of M<sup>pro</sup> from SARS-CoV and SARS-CoV-2 provides additional insight into  
127 the mechanism of inhibition by this anti-coronaviral drug.

128

## 129 **Results**

### 130 **K<sub>i</sub> values of GC376 inhibition of M<sup>pro</sup> from both SARS-CoV and SARS-CoV-2 are in** 131 **nanomolar range**

132 Determining K<sub>i</sub> values that are reflective of drug binding affinity is a prerequisite for the  
133 prediction and evaluation of drug interactions. In our previous report, we determined the half-  
134 maximal inhibitor concentrations (IC<sub>50</sub>), values, which describe the functional strength of the  
135 inhibitor, the feline drug GC376 with both M<sup>pro</sup> of SARS-CoV and SARS-CoV-2 [13]. Here we  
136 determine K<sub>i</sub> values for the prodrug GC376 with both M<sup>pro</sup> of SARS-CoV and SARS-CoV-2. For  
137 K<sub>i</sub> determination, the inhibitory effects of increasing concentrations of GC376 on M<sup>pro</sup> from both  
138 SARS-CoV and SARS-CoV-2 were tested using the synthetic peptide FRET-substrate Abz-  
139 SVTLQSG-Y(NO<sub>2</sub>)-R followed by Michaelis-Menten kinetics. Data was plotted as reaction rate

140 versus substrate concentration (primary Lineweaver-Burk plot) and the slopes ( $K_m/V_{max}$ ) were  
141 determined by linear regression analysis. The slopes were plotted versus the concentration of  
142 GC376 to determine the inhibitory constant ( $K_i$  as y-intercept). The  $K_i$  for GC376 was 0.02  $\mu\text{M}$   
143 for SARS-CoV M<sup>pro</sup> and 0.04  $\mu\text{M}$  for SARS-CoV-2 M<sup>pro</sup>, (**Fig 1 and Table 1**).

144

### 145 **GC376 is a reversible inhibitor with M<sup>pro</sup> from both SARS-CoV and SARS-CoV-2**

146 An important factor to consider when developing a therapeutic protease inhibitor is the  
147 reversibility of compound binding [24]. Irreversible protease drugs can yield long-lasting effects  
148 by permanently blocking proteases in cells that are not the intended target and thus causing  
149 detrimental consequences resulting in side effects and antigenicity of covalently modified proteins  
150 [26]. We previously demonstrated that the bisulfite prodrug GC376 converts to the peptide  
151 aldehyde GC373 which interacts covalently with the catalytic cysteine of SARS-CoV-2 M<sup>pro</sup>[13],  
152 but did not assess experimentally whether the inhibition was reversible.

153 Reversibility of GC376 with SARS-CoV-2 M<sup>pro</sup> was evaluated first by NMR studies using  
154 <sup>13</sup>C-labelled GC373 (**Fig 2A**). HSQC experiments of samples containing only SARS-CoV-2 M<sup>pro</sup>  
155 (**Fig 2B**), inhibitor (**Fig 2C**), or both co-incubated (**Fig 2D**) provided spectra to which the  
156 reversibility experiment could be compared. Evidence of binding reversibility was acquired by  
157 HSQC experiments conducted on a co-incubated sample containing both enzyme and inhibitor that  
158 was subsequently washed with buffer. The subsequent HSQC experiment using this sample  
159 showed a disappearance of the NMR signal corresponding to the bound inhibitor (**Fig 2E**). The  
160 disappearance of this signal would only be observed in the case of a binding reversion event.

161 We then conducted a detailed study to provide the rate and percentage of reversibility, as  
162 well as the comparison of drug behaviour with SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup>.

163 Reversibility was tested by measuring catalytic activity post dialysis. Incubation of SARS-CoV  
164 M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup> with the GC376 followed by dialysis resulted in increase in enzymatic  
165 activity over time, indicative of a reversible dissociation of inhibitor (**Fig. 3**). We observed a  
166 recovery of 10% of activity after 22 hours of dialysis, which reached 30 - 40% of initial activity  
167 for SARS-CoV and 40-60% for SARS-CoV-2 after 4 days of dialysis, suggesting over time the  
168 substrate competed for the enzyme binding site. To ensure the proteins remained stable over this  
169 time period, we also monitored the stability of uninhibited enzymes, which was compared with the  
170 activity of recovered enzymes. After 4 days the residual protease activity for the uninhibited M<sup>pro</sup>  
171 of SARS-CoV and SARS-CoV-2 was 30-40%, which allowed us to conclude that the drug was  
172 fully reversible.

173

#### 174 **SARS-CoV-2 M<sup>pro</sup> has enhanced stability compared to SARS-CoV M<sup>pro</sup>**

175 After observing the high kinetic stability of both viral proteases at room temperature, we  
176 characterized their thermal stability and assessed their thermodynamic parameters including  
177 activation energies of inactivation. Thermal stability is a characteristic used to describe the kinetic  
178 stability of enzymes, and many individual proteins or protein complexes are known to have high  
179 kinetic stability [27-31]. For viral proteins, particularly the structural ones, this feature is crucial  
180 because virus particles must be able to resist harsh environmental conditions until they find a new  
181 host to infect and also remain stable during infection [10,13,32]. For example, determination of  
182 thermodynamic parameters of the HIV protease in the presence of various inhibitors was used to  
183 reveal the differences in protein stability upon forming inhibitor-protein complexes, which  
184 informed on inhibitor design [33].

185 Thermal inactivation of SARS-CoV M<sup>pro</sup> (**Fig 4A and 4B**) and SARS-CoV-2 M<sup>pro</sup> (**Fig 4D**  
186 **and 4F**) was studied at the temperature range of 24–70 °C in a time-dependant manner. The  
187 semilogarithmic plots of residual activity versus incubation time were linear at all temperatures  
188 for both proteins, which was indicative of a simple first-order monophasic kinetic process. From  
189 the slopes of semilogarithmic plots inactivation rate constants were calculated and are given in  
190 **Table 2**. For both proteases, the rate constant progressively increased with increasing temperatures  
191 whereas half-life ( $t_{1/2}$ ) and the decimal reduction time (Dt), two important parameters used in  
192 characterization of enzyme stability, decreased.

193 The dependence of inactivation rate constants on temperature was plotted using the  
194 Arrhenius equation (**Fig 4C and 4F**), from which apparent activation energies of inactivation ( $E_a$ )  
195 were calculated. Interestingly, Arrhenius plots for both proteases were not linear and showed  
196 upward curvature suggesting two denaturation processes each with its own temperature  
197 dependence and activation energy. At temperatures above 37 °C inactivation is a result of protein  
198 unfolding with high activation energy, with the rate of this process strongly dependant on  
199 temperature. At temperatures of 37 °C and below this rate becomes insignificant and other  
200 processes with low activation energy prevail. The activation energies for the high temperature  
201 range were found to be high and similar for SARS-CoV M<sup>pro</sup> ( $E_a=243.6$  kJ/mol) and SARS-CoV-  
202 2 M<sup>pro</sup> ( $E_a=234.2$  kJ/mol). However, for the low temperature range the activation energies were  
203 10—20% of those determined at high temperature, confirming that M<sup>pro</sup> inactivation involves both  
204 high- and low-activation energy processes. Interestingly, the parameters of the inactivation process  
205 at low temperature range (24–37 °C) are different for M<sup>pro</sup> from SARS-CoV and SARS-CoV-2,  
206 showing  $E_a$  of 16.4 kJ/mol and 41.4 kJ/mol and  $t_{1/2}$  (at 24 °C) of 38.5 h and 57.7 h respectively,  
207 suggesting higher stability for SARS-CoV M<sup>pro</sup>.

208 Determination of all thermodynamic parameters of inactivation can provide further  
209 information on enzyme stability.  $\Delta G$  value, the Gibbs free energy, which is the energy barrier for  
210 enzyme inactivation, is directly related to protein stability. We see a significant decrease in  $\Delta G$  for  
211 the temperatures above 55 °C indicating that the destabilization process occurs rapidly in this  
212 temperature range (**Table 2**).

213 To gain a deeper insight into the driving forces of SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup>  
214 stability, the Gibbs free energy was decomposed into its enthalpic and entropic contributions.  
215 Enthalpy,  $\Delta H$ , measures the number of non-covalent bonds broken during transition state  
216 formation for enzyme inactivation, allowing us to compare the energy landscapes of both SARS-  
217 CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup>. For temperature ranging from 37 °C to 70 °C we observed  
218 consistent high  $\Delta H$  values, which is in agreement with a temperature-dependent inactivation  
219 process. Interestingly, at the 24 °C and 37 °C temperature interval a significant jump in  $\Delta H$  occurred  
220 for both proteases, however, with different initial enthalpy values for SARS-CoV M<sup>pro</sup> and SARS-  
221 CoV-2 M<sup>pro</sup> at 24 °C (13.9 and 38.9 kJ/mol respectively), again highlighting higher stability of  
222 latter at physiological temperatures (Table 2). The compactness in the protein molecular structure  
223 as well as enzyme and solvent disorder can be inferred through the quantitative analysis of entropy  
224  $\Delta S$  values [34,35]. Small negative entropy values at 24 °C for both SARS-CoV M<sup>pro</sup> and SARS-  
225 CoV-2 M<sup>pro</sup> confirmed no disorder in protein structure upon inactivation; however, at higher  
226 temperatures all values of  $\Delta S$  were positive and similar, suggesting that unfolding is a rate-limiting  
227 step at this range (**Table 2**).

228

229 **Structural comparison of M<sup>pro</sup> from SARS-CoV and SARS-CoV-2**

230 We previously reported increased catalytic activity of SARS-CoV-2 M<sup>PRO</sup> in comparison to  
231 SARS-CoV M<sup>PRO</sup> with the catalytic turnover rate being almost 5 times higher for the former using  
232 a FRET-peptide as substrate [13]. We were interested in structural comparison of the M<sup>PRO</sup> from  
233 SARS-CoV and SARS-CoV-2, for both apo and drug-bound forms to reveal differences that  
234 account for the enhancement in activity. Crystal structures of apo-M<sup>PRO</sup> from SARS-CoV and  
235 SARS-CoV-2, and bisulphite prodrug (GC376) and the aldehyde drug (GC373) bound forms were  
236 determined. The two proteins share 96% sequence identity with only 12 out of 306 residues being  
237 different (**S1 Fig**). Therefore, as expected, there is little change in the overall structures of apo-  
238 SARS-CoV and SARS-CoV-2 M<sup>PRO</sup> (**Fig 5**), with an RMSD of 0.6 Å. We observed a new helical  
239 feature at  $\eta_2$  (residues 47-50) in SARS-CoV-2, which is unfolded in SARS-CoV, (**S1 and S2 Fig**).  
240 It is located at the entrance to the active site, near a non-conserved residue between SARS-CoV,  
241 and SARS-CoV-2 (**S2 Fig**). In the GC373-bound form of proteins, however we observed the  
242 opposite; this helix is found in the M<sup>PRO</sup> of SARS-CoV but not in SARS-CoV-2 (**S3 Fig**), suggesting  
243 a dynamic nature of this structural element.

244 Both SARS-CoV and SARS-CoV-2 M<sup>PRO</sup> form dimers, and while monomers have very low  
245 activity dimerization is necessary for full enzymatic activity and virulence [36,37]. Comparative  
246 analysis of the biological dimer of the two proteases revealed that the main differences are located  
247 at the dimer interface. In the M<sup>PRO</sup> of SARS-CoV-2, we observed a slight shift of the chymotrypsin-  
248 like domains away from each other, compared to the M<sup>PRO</sup> of SARS-CoV (**Fig 5B**). However, the  
249 biggest change is the difference in association between the dimerization domains (**Fig 5C and**  
250 **5D**). The dimer interface of SARS-CoV and SARS-CoV-2 M<sup>PRO</sup> is facilitated by several  
251 interactions between the two protomers, one of which is between the helical domain III of each  
252 protomer comprising of residues 284-286, specifically Ser-Thr-Ile (STI) in SARS-CoV M<sup>PRO</sup> and

253 Ser-Ala-Leu (SAL) in SARS-CoV-2 M<sup>pro</sup>. This unstructured loop self-associates between  
254 protomers in the dimer. Importantly, this region harbors a non-conservative residue in sequence at  
255 the dimer interface, where the Thr285 in SARS-CoV M<sup>pro</sup> is altered to Ala285 in SARS-CoV-2  
256 M<sup>pro</sup> (**Fig 5E and 5F**). The SAL-motif forms a tight van der Waals interaction and the residues  
257 from each protomer interdigitate to form a complementary interface that readily explains the  
258 observed enhanced stability.

259

## 260 **GC376 inhibited forms of SARS-CoV-2 and SARS-CoV M<sup>pro</sup> reveal a common mechanism** 261 **of inhibition**

262 We recently presented the structure of GC373 with the SARS-CoV-2 M<sup>pro</sup> [13]. The  
263 structure of SARS-CoV-2 M<sup>pro</sup> with drug GC373, as well as prodrug GC376 that converts to  
264 GC373, reflects the specificity of the enzyme for a glutamine surrogate in the P1 position and a  
265 leucine, which is preferred in the P2 position. A benzyl group is in the P3 position. Here we  
266 determined the crystal structure of the SARS-CoV M<sup>pro</sup> with the prodrug GC376 and drug GC373  
267 to examine features that determine its efficacy and compare this with the previously determined  
268 SARS-CoV-2 structure (**Fig 6**).

269 SARS-CoV M<sup>pro</sup> was incubated with GC373 and GC376, prior to crystallization. The best  
270 crystals diffracted to 2.0 Å, and the data was refined with good statistics (**Table 3**). Overall  
271 comparison of SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup> structures with GC373 showed similar  
272 agreements with the apo-M<sup>pro</sup> structures, with an RMSD of 0.6 Å (**Fig 6**). The drug binding is  
273 supported by H-bonding with the main chain of oxyanion hole residues Asn142, Gly143 and  
274 Ser144, which are identical for both proteases (**Fig 6B, S4 Fig and S5 Fig**). A good fit was  
275 observed for both the P1 and P2 positions, supported structurally by hydrogen bonding and van

276 der Waals interactions respectively with H-bonds for the P1 position being identical for M<sup>pro</sup> from  
277 SARS-CoV and SARS-CoV-2 (**Fig 6C, S4 Fig and S5 Fig**).

278

279 **The N-terminal finger of the M<sup>pro</sup> stabilizes dimer formation and coordination of the drug**  
280 **GC373**

281 A distinctive feature of M<sup>pro</sup> dimer is the interaction of N-terminal residues (“N-finger”) of  
282 protomer A with residues of domain II of protomer B. In the dimer for both protomers of SARS-  
283 CoV-2 M<sup>pro</sup> and SARS-CoV M<sup>pro</sup>, we observe the N-termini interact with residues near S1  
284 substrate-binding subsite in a hairpin adjacent to the oxyanion hole of the active site (**Fig 7**). The  
285 NH-group of Ser1 from protomer A forms strong H-bonds with the carboxylate group of Glu166  
286 (3.1 Å) and the carbonyl of Phe140 (3.3 Å) of protomer B and *vice versa*. This interaction stabilizes  
287 the enzyme, assists in the correct orientation of the oxyanion loop and S1 pocket of the substrate  
288 binding site, and thus results in enhanced catalytic efficiency, as observed in previous studies  
289 demonstrating the native N-terminal serine provides the most efficient enzyme with SARS-CoV  
290 M<sup>pro</sup> [38]. Interestingly, the H-bond distance between the Ser1 (protomer A) and Phe140 (protomer  
291 B) is closer in SARS-CoV-2 M<sup>pro</sup> (3.3 Å) compared to SARS-CoV M<sup>pro</sup> (5.5 Å) (**Fig. 8**), likely  
292 adding to its increased catalytic activity. The proper conformation of S1 pocket is also important  
293 for the drug binding and importantly, P1 position of GC373 is also stabilized by hydrogen bonding  
294 between the side chain of Glu166 (3.3 Å) and backbone carbonyl of Phe140 (3.3 Å) residues (**Fig**  
295 **8**). Thus, a hydrogen bond network between the dimer in M<sup>pro</sup> stabilizes the S1 substrate for  
296 substrate binding and hence inhibitor binding.

297 Residues adjacent to the N-terminus also play a key role in dimerization, specifically Pro9  
298 and Phe305 from protomer A, which interact with residues Pro122 and Ser123 in a strand on

299 protomer B. We also observe these interactions in all of our SARS-CoV M<sup>pro</sup> and SARS-CoV-2  
300 M<sup>pro</sup> structures bound to the inhibitor (**S7 Fig**). Mutation of Pro9 to Thr results in a monomeric  
301 species of SARS-CoV-2 M<sup>pro</sup> [39]. Together this data suggests a strong role for the N-terminus of  
302 the protease not only in function and stability, but also with inhibitor coordination.

303

## 304 **Discussion**

305 Here we show that the feline antiviral prodrug GC376 is reversible and inhibits M<sup>pro</sup> of  
306 both SARS-CoV and SARS-CoV-2 with low nanomolar K<sub>i</sub> values. While IC<sub>50</sub> values, the  
307 concentration of inhibitor at half-maximal inhibition, are very useful during drug development  
308 [40], K<sub>i</sub> values describe precise binding affinity between the inhibitor and enzyme, independent of  
309 experimental conditions, and allow for comparisons during structure-activity relationship (SAR)  
310 studies. Here we show K<sub>i</sub> values for GC376 with the SARS-CoV and SARS-CoV-2 M<sup>pro</sup> to be 20  
311 nM and 40 nM, respectively. These are lower, as expected, when compared to the IC<sub>50</sub> values of  
312 prodrug GC376 (190 nM) and drug GC373 (400 nM) with SARS-CoV-2 M<sup>pro</sup> [13]. The high  
313 degree of sequence identity between the SARS-CoV and SARS-CoV-2 M<sup>pro</sup> suggests strong  
314 conservation in proteolytic inhibition supported by K<sub>i</sub> values.

315 K<sub>i</sub> values for GC376 are in line with K<sub>i</sub> values of reported proteolytic inhibitors targeting  
316 the HCV serine protease and currently being used to treat hepatitis C such as first-generation HCV  
317 ns3/4A inhibitors Boceprevir with low nM K<sub>i</sub> values and second-generation inhibitors with  
318 subnanomolar K<sub>i</sub> values [41]. These drugs are reversible serine protease inhibitors whose  
319 development was facilitated by SAR studies [41,42]. Our K<sub>i</sub> data further supports GC376 being a  
320 broad-spectrum inhibitor [16,17,20,22], and demonstrates it is in the inhibitory range to be  
321 considered as a viable antiviral for clinical trials.

322 M<sup>pro</sup> from SARS-CoV and SARS-CoV-2 have 96% sequence identity and variant residues,  
323 with the exception of Ala285 discussed above, are conservative (**S1 Fig**). Therefore, it was not  
324 surprising that both proteins revealed similar physical chemical properties such as high thermal  
325 stability at temperatures above 37 °C with high activation energies and enthalpy independent of  
326 temperature (**Table 2**). However, at physiological temperatures (24-37 °C) we observed a  
327 difference in stability between SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup>, with the latter being more  
328 stable, exhibiting higher values of  $t_{1/2}$  (38.5 h for SARS-CoV M<sup>pro</sup> versus 57.7 h for SARS-CoV-  
329 2 M<sup>pro</sup>) and enthalpy (13.9 kJ/mol for SARS-CoV M<sup>pro</sup> versus 38.9 kJ/mol for SARS-CoV-2 M<sup>pro</sup>).  
330 A high  $\Delta H$  value is usually indicative of a larger number of noncovalent intramolecular bonds,  
331 which contribute to protein stability. Therefore, in order to understand what variant residues could  
332 be responsible for enhanced stability of SARS-CoV-2 M<sup>pro</sup> compared to SARS-CoV, we  
333 examined the regions with amino acid substitutions more closely.

334 Both the SARS-CoV and SARS-CoV-2 M<sup>pro</sup> are dimeric in nature. Early crystal structures  
335 of SARS-CoV M<sup>pro</sup> elucidated how the dimers assemble [7,43] and mutagenesis has revealed that  
336 residues at the dimer interface are important for both activity and stability [36,37,44]. From our  
337 crystal structures we observe that overall the dimerization motifs of both SARS-CoV and SARS-  
338 CoV-2 M<sup>pro</sup> are very similar, however, one key change at the domain III interface, namely  
339 Thr285Ala in SARS-CoV-2 M<sup>pro</sup>, results in a significant alteration in the distance between the  
340 domains of the protomers in the SARS-CoV-2 M<sup>pro</sup> dimer compared to SARS-CoV M<sup>pro</sup> (**Fig 5**).  
341 This mutation leads to residues in the domain III interface forming a hydrophobic zipper clearly  
342 aligning the two domains, and thus likely enhancing the  $t_{1/2}$  at low temperatures as we have  
343 observed above. The high degree of stability of the enzymes for both SARS-CoV and SARS-CoV-  
344 2 is an interesting feature that likely contributes to viral potency.

345 Another structural feature that might explain the increased activity and stability is a closer  
346 association between the N-finger Ser1 and Phe140 in the oxyanion loop in the M<sup>pro</sup> of SARS-CoV-  
347 2 compared to SARS-CoV (**Fig 8**). This interaction plays a critical role for activity since it sustains  
348 the correct conformation of the oxyanion loop, therefore precise coordination of the N-finger in  
349 both M<sup>pro</sup> of SARS-CoV and SARS-CoV-2 is a prerequisite for function. Previous work  
350 demonstrated that enzymatic activity of SARS-CoV M<sup>pro</sup> was diminished with non-native affinity  
351 tags proving the need for native N- and C-termini [6,38]. The effect was most pronounced with  
352 additional residues at the N-terminus, with the activity of the wild-type being 20-fold greater than  
353 a variant with an additional glycine at the N-terminus [38].

354 While GC376 has been crystallized with the main protease of the similar betacoronavirus  
355 MERS [18], as well as other viral proteases, including norovirus and porcine diarrhea virus  
356 (PEDV) [45], no N-finger association was observed in those crystal structures. This structural  
357 motif, however, was observed in a SARS-CoV M<sup>pro</sup> crystal structure with a Michael acceptor  
358 inhibitor, however the N-finger interaction was diminished with the addition of residues at the  
359 native N-terminus [38].

360 We demonstrated that the NH group of Ser 1 donates H-bonds to Phe140 and Glu166, the  
361 residues that coordinate the N-termini of each protomer in the dimer. Importantly, these residues  
362 also interact with the P1 position of GC373 in both SARS-CoV and SARS-CoV-2, demonstrating  
363 a strong hydrogen bond network near the active site, and stabilization of the S1 subsite pocket.  
364 This likely contributes to the high K<sub>i</sub> values for these inhibitors. The precise structural and  
365 mechanistic elucidation of the inhibitor-protease interaction and implications for M<sup>pro</sup> dimerization  
366 is paramount for the fine-tuned design of universally active inhibitor drugs. In this regard, the

367 current study provides a rationale for the precise nature of a gamma-lactam group in the P1 position  
368 of the GC373/GC376 inhibitor.

369 With coronavirus outbreaks occurring in 2002, 2015 and 2019, it is clear that broad-  
370 spectrum antivirals will be needed for the current pandemic and in the future. The development of  
371 antivirals to treat coronavirus infections remains a high priority. By comparing kinetic,  
372 thermodynamic, and structural features of M<sup>pro</sup> from SARS-CoV and SARS-CoV-2 and their  
373 binding to GC373/GC376 we revealed distinct supramolecular differences in overall protease  
374 properties, yet demonstrate comparable efficacies of GC376 with both proteases. Furthermore,  
375 reversible inhibition with the drug further supports the clinical potential of the GC376 compound.  
376 The results presented here support the use of GC376 as an antiviral with broad specificity against  
377 coronaviruses.

378

## 379 **Methods**

### 380 *Purification of SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup>*

381 Purifications of proteases were performed as described earlier [13]. Briefly, pET SUMO (small  
382 ubiquitin-like modifier) expression vector (Invitrogen) bearing M<sup>pro</sup> from SARS-CoV-2 gene with  
383 N-terminal His-SUMO tag was transformed into *E. coli* BL21 (DE3), induced with 0.5 mM  
384 isopropyl  $\beta$ -d-1-thiogalactopyranoside and the protein was expressed for 4–5 h at 37 °C. After  
385 harvesting by centrifugation (4400  $\times$  g for 10 min at 4 °C) cells were suspended in lysis buffer  
386 (20 mM Tris-HCl, pH 7.8, 150 mM NaCl, 5 mM imidazole) and lysed by sonication. The lysate  
387 was clarified by centrifugation at 17,000  $\times$  g for 30 min, and the supernatant was loaded onto Ni-  
388 NTA resin column (Qiagen). The resin was washed with 10 column volumes of lysis buffer  
389 containing 20 mM imidazole and the fusion protein was eluted with 40–500 mM imidazole in the

390 same buffer. Eluted fractions containing the protein of interest were pooled together and dialyzed  
391 against lysis buffer containing 1 mM DTT at 4 °C. The fusion protein was subsequently digested  
392 with His-tagged SUMO protease (McLab, South San Francisco, CA) at 4 °C for 1–2 h to remove  
393 the SUMO tag and the resulting cleavage mixture was then passed through Ni-NTA resin column.  
394 The flow through containing SARS-CoV-2 M<sup>pro</sup> was collected and further purified using size  
395 exclusion chromatography column (G-100, GE Healthcare,) equilibrated with 20 mM Tris, 20 mM  
396 NaCl, 1 mM DTT, pH 7.8. Fractions containing the SARS-CoV-2 M<sup>pro</sup> protein were pooled and  
397 concentrated using Amicon Ultra-15 filter with a MWCO of 10 kDa. The plasmid encoding the  
398 SARS-CoV M<sup>pro</sup> with an N-terminal His-tag upstream of a Factor Xa cleavage site was a kind gift  
399 of Dr. Michael James. The protein was expressed and purified the same way as SARS-CoV M<sup>pro</sup>-  
400 2 but Factor Xa protease (Sigma, Canada) was used (4 °C, overnight) to remove the tag.

401

#### 402 *Inhibitor and FRET Substrate Synthesis*

403 Inhibitors GC373 and GC376, and the FRET substrate Abz-SVTLQSG-Y(NO<sub>2</sub>)-R were  
404 synthesized according to methods previously described[13].

405

#### 406 *Kinetic experiments*

407 The activity determination of both proteases was performed as previously described[13] using  
408 FRET-based cleavage assay with a synthesized fluorescent substrate containing the cleavage site  
409 (indicated by the arrow, ↓) of SARS-CoV-2 M<sup>pro</sup> (Abz-SVTLQ↓SG-Tyr(NO<sub>2</sub>)-R) in 20 mM Bis-  
410 Tris, pH 7.8, 1 mM DTT activity buffer at 37 °C for 10 min. The concentration of proteases was  
411 fixed at 80 nM and the range of 0.1–500 μM was used for the substrate. Reactions were started  
412 with the enzyme and the fluorescence signal of the Abz-SVTLQ peptide cleavage product was

413 monitored at an emission wavelength of 420 nm with excitation at 320 nm, using an Flx800  
414 fluorescence spectrophotometer (BioTek, USA). The GC376 compound was dissolved in DMSO  
415 and used in a concentration range of 0.01–0.4  $\mu\text{M}$  to inhibit both proteases and measure their  
416 kinetic parameters. Kinetic data corresponding the interaction of SARS-CoV M<sup>pro</sup> and SARS CoV-  
417 2 M<sup>pro</sup> with GC376 compound were analyzed using computer-fit calculation (Prism 4.0, GraphPad  
418 Software). The slopes of the Lineweaver-Burk plots were plotted versus the inhibitor concentration  
419 and the  $K_i$  values were determined from the x-axis intercept as  $-K_i$ .

420

#### 421 *NMR Experiments on Reversibility of Inhibitor Binding*

422 The  $^{13}\text{C}$ -labelled GC376 inhibitor was synthesized according to previously documented  
423 procedures, and initial HSQC NMR experiments involving only enzyme, only inhibitor, and both  
424 co-incubated were prepared as previously described [13]. The sample used for the reversibility  
425 experiment was prepared by subjecting a previously co-incubated sample containing both enzyme  
426 and inhibitor to washing steps with buffer ( $\text{D}_2\text{O}$ , 50 mM phosphate, pD 7.5 with 20 mM DTT).  
427 This involved depositing the sample in an Amicon micro-spinfilter with a 10 kDa cutoff and  
428 spinning down the sample at 6600 g for 15 min. The sample was then diluted to 300  $\mu\text{L}$  and the  
429 spin down and dilution steps were repeated once more, to a final volume of 300  $\mu\text{L}$ . This sample  
430 was then analyzed by NMR in an HSQC experiment, following protocols identical to those  
431 previously described [13].

432

#### 433 *Reversibility and Stability of 3CL Proteases from SARS-CoV and SARS-CoV-2*

434 Reversibility of 3CL protease inhibition with GC376 was determined by dialysis method. The  
435 proteases were incubated with a single concentration (20  $\mu\text{M}$ ) of the GC376 compound for 15 min

436 at RT to allow for full inhibition. Then the enzyme-inhibitor mixture was placed in a 6–8 kDa  
437 MWCO dialysis membrane (Fisher Scientific, Canada) and dialyzed against 2 L of 50 mM Tris-  
438 HCl, pH 7.8, 150 mM NaCl, 5% glycerol, 1mM DTT at RT. The dialysis buffer was changed every  
439 24 hours. Control experiments, which included dialyzing apo-proteases at the same concentration  
440 in the same dialysis buffer but different beakers, were performed simultaneously. The aliquots of  
441 dialyzing samples were taken out at certain time points and used for activity measurements. The  
442 data was represented as a percent of initial protease activity at a zero time point.

443 The thermal stability was determined by heating 2  $\mu$ M solution of M<sup>pro</sup> SARS CoV or M<sup>pro</sup> SARS-  
444 CoV-2 in 50 mM Tris-HCl, pH 7.8, 150 mM NaCl, 5% glycerol, 1mM DTT buffer in a  
445 thermostatted water-bath at various temperatures. 30  $\mu$ l protein samples were taken out at specific  
446 time points and immediately incubated on ice until activity measurements were performed as  
447 described above. Residual activities were expressed as relative to the maximal activity, which was  
448 the activity of proteases at zero time point.

449 The enzyme inactivation over time is described by a first-order equation:

$$450 \quad \ln(A_t/A_0) = -kt \quad (1)$$

451 where  $A$  represents enzyme activity at time  $t$ ,  $A_0$  is the initial activity at time zero,  $k$  is the rate  
452 constant ( $\text{min}^{-1}$ ), and  $t$  is time (min). Inactivation rate constants ( $k_d$ ) were obtained from slopes of  
453 semi-logarithmical plots of residual activity versus incubation time at each temperature. Calculated  
454 rate constants were replotted in Arrhenius plots as natural logarithms of  $k$  versus the reciprocal of  
455 absolute temperature. Arrhenius law describes the temperature dependence of rate constant as

$$456 \quad \ln(k) = -Ea/RT + c \quad (2)$$

457 where  $Ea$  is the activation energy,  $R$  is the universal gas constant ( $8.31 \text{ J mol}^{-1}\text{K}^{-1}$ ), and  $T$  is the  
458 absolute temperature.  $Ea$  was calculated from the slope of Arrhenius plot.

459 The half-life of proteases ( $t_{1/2}$ ), defined as time after which activity is reduced to 50% of initial  
460 value [46], was determined as

$$461 \quad t_{1/2} = \ln(2)/k \quad (3)$$

462 Another common way to present inactivation rate is as  $D$  value – decimal reduction time, which is  
463 the time required to reduce activity to 10% of the original value and calculated as:

$$464 \quad D = \ln(10)/k \quad (4)$$

465 The activation free energy ( $\Delta G$ ,  $\text{kJ mol}^{-1}$ ), enthalpy ( $\Delta H^\circ$ ,  $\text{kJ mol}^{-1}$ ) and entropy ( $\Delta S^\circ$ ,  $\text{kJ mol}^{-1} \text{K}^{-1}$ )  
466 were determined as

$$467 \quad \Delta G = -RT \ln(kh/k_B T) \quad (5)$$

$$468 \quad \Delta H = E_a - RT \quad (6)$$

$$469 \quad \Delta S = (\Delta H - \Delta G)/T \quad (7)$$

470 where  $h$  is the Planck constant ( $6.626 \times 10^{-34}$  Js) and  $k_B$  is the Boltzmann constant ( $1.38 \times 10^{-23}$  J  
471  $\text{K}^{-1}$ ). Experiments were performed in duplicate.

472

### 473 *Crystallization*

474 For crystallization, purified SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup> were dialysed against buffer  
475 containing 10 mM NaCl and 5mM Tris HCl pH 8.0 overnight at 4 °C. Both proteins were  
476 concentrated with a Millipore centrifugal filter (10 kDa MW cut-off) to a concentration of 9  
477 mg/mL. Protein was incubated with 5 molar excess of inhibitor at 4 °C for 2 h prior to  
478 crystallization. For SARS-CoV M<sup>pro</sup>, crystals were screened around previously known established  
479 conditions [13] with the best crystals forming with vapour diffusion hanging drop trays at room  
480 temperature at a ratio of 1:1 with mother liquor containing 10 mM CaCl<sub>2</sub>, 7% PEG 8000, 1 mM  
481 MES pH 6.0, 1mM DTT, 3% ethylene glycol and 3% DMSO (Data not shown). For SARS-CoV-

482 2 M<sup>pro</sup>, the protein was subjected to the PACT crystallization screen (Molecular Dimensions,  
483 USA), with hits identified in several conditions for both inhibitors. Best crystals were observed  
484 with hanging drop trays at room temperature at a ratio of 1:1 with mother liquor 0.2 M Sodium  
485 sulfate, 0.1 M Bis-Tris propane pH 6.5, 20% w/v PEG 3350. While the SARS-CoV-2 M<sup>pro</sup> with  
486 ligands crystallize with mother liquid containing 0.2 M Sodium chloride 0.1 M HEPES pH 7.0 20  
487 % w/v PEG 6000. Prior to freezing, crystals were incubated with 15% glycerol as a cryoprotectant  
488 for SARS-CoV-2 M<sup>pro</sup> and 20% ethylene glycol for SARS-CoV M<sup>pro</sup>. Crystals were initially  
489 screened at in-house 007 MicroMax (Rigaku Inc) with final data collection at Stanford  
490 Synchrotron Radiation Lightsource SSRL, USA, beamline 12-2 with Blu-Ice using the Web-Ice  
491 interface [47].

492

493 *Diffraction Data Collection, Phase Determination, Model Building, and Refinement.*

494 All diffraction data sets were collected using synchrotron radiation of wavelength 0.97946 Å at  
495 beamline 12-2 of Stanford Synchrotron Radiation Lightsource (SSRL) California, USA, using a  
496 Dectris PILATUS 6M detector. Several data sets were collected from the crystals of SARS-CoV-  
497 2 M<sup>pro</sup> free enzyme as well as with GC376 and GC373 treated. Numerous data sets were also  
498 collected for SARS-CoV in the presence of GC376 and GC373. XDS2 [48] and Scala were used  
499 for processing the data sets. The diffraction data set of the free SARS-CoV-2 M<sup>pro</sup> was processed  
500 at a resolution of 1.75 Å, in space group P21 (Supplementary Table 1). For the complex of SARS-  
501 CoV-2 M<sup>pro</sup> with GC376 and GC373, the data set collected, was processed at a resolution of 1.9 Å  
502 and 2.0 Å and in space group C2 (Supplementary Table 1). All three structures were determined  
503 by molecular replacement with the crystal structure of the free enzyme of the SARS-CoV-2 M<sup>pro</sup>  
504 (PDB entry 6Y7M as search model, using the Phaser program from Phenix[49], version v1.18.1-

505 3855). SARS-CoV M<sup>pro</sup> data were also processed with XDS231 and Scala at a resolution of 2.15  
506 Å and 1.90 Å for GC376 and GC373, respectively, in a space group C2. Ligand Fit from Phenix  
507 [50] was employed for the fitting of both inhibitors in the density of pre-calculated map from  
508 Phenix refinement, using the ligand code K36. Refinement of all the structures was performed  
509 with phenix.refine in Phenix software. Statistics of diffraction, data processing and model  
510 refinement are given in (Supplementary Table 1). The model was inspected with Ramachandran  
511 plots and all show good stereochemistry. Final models displayed using PyMOL molecular graphics  
512 software (Version 2.0 Schrödinger, LLC).

513

514 **Competing interests:** The authors declare no competing interests.

515

516 **Author contributions:** J.C.V., W.V. and T.L. contributed to inhibitor synthesis. T.L. contributed  
517 to FRET-substrate synthesis. E.A., M.J.v.B., J.L. and C.F. contributed to purified protein. C.F. and  
518 E.A. contributed to enzyme kinetics and reversibility studies. M.J.L., H.S.Y. E.A., J.L. and M.B.K.  
519 contributed to crystallization and structure determination. W. V. and R. T. M. contributed to  
520 labelled NMR studies. M.J.L wrote the initial draft. All authors read and approved the manuscript.

521

522 **Acknowledgements.** We would like to thank the staff at the Stanford Synchrotron Light Source,  
523 in particular Dr. Silvia Russi and Lisa Dunn. M.J.L, J.C.V and D.L.T. acknowledge funding from  
524 CIHR and NSERC (COVID-19 SOF-549297-2019). D.L.T acknowledges support from Li Ka  
525 Shing Institute of Virology and the GSK Chair in Virology. W.V. was supported by an Alberta  
526 Innovates Graduate Scholarship and an Alberta Graduate Excellence Scholarship. T.L. was  
527 supported by a CIHR Vanier Scholarship.

528  
529 Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory,  
530 is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences  
531 under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is  
532 supported by the DOE Office of Biological and Environmental Research, and by the National  
533 Institutes of Health, National Institute of General Medical Sciences (P41GM103393). The  
534 contents of this publication are solely the responsibility of the authors and do not necessarily  
535 represent the official views of NIGMS or NIH.

536

## 537 **References**

538

- 539 1. Wu F, Zhao S, Yu B, Chen YM, Wang W, et al. (2020) A new coronavirus associated with  
540 human respiratory disease in China. *Nature* 579: 265-269.
- 541 2. Lu R, Zhao X, Li J, Niu P, Yang B, et al. (2020) Genomic characterisation and epidemiology of  
542 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 395:  
543 565-574.
- 544 3. Wu A, Peng Y, Huang B, Ding X, Wang X, et al. (2020) Genome Composition and Divergence  
545 of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe* 27: 325-  
546 328.
- 547 4. Gorbalenya AE, Baker SC (2020) The species Severe acute respiratory syndrome-related  
548 coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* 5: 536-  
549 544.
- 550 5. Hilgenfeld R (2014) From SARS to MERS: crystallographic studies on coronaviral proteases  
551 enable antiviral drug design. *FEBS J* 281: 4085-4096.
- 552 6. Grum-Tokars V, Ratia K, Begaye A, Baker SC, Mesecar AD (2008) Evaluating the 3C-like  
553 protease activity of SARS-Coronavirus: recommendations for standardized assays for drug  
554 discovery. *Virus Res* 133: 63-73.
- 555 7. Muramatsu T, Takemoto C, Kim YT, Wang H, Nishii W, et al. (2016) SARS-CoV 3CL protease  
556 cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proc Natl Acad*  
557 *Sci U S A* 113: 12997-13002.
- 558 8. Rut W, Groborz K, Zhang L, Sun X, Zmudzinski M, et al. (2021) SARS-CoV-2 M(pro)  
559 inhibitors and activity-based probes for patient-sample imaging. *Nat Chem Biol* 17: 222-  
560 228.

- 561 9. Kneller DW, Galanie S, Phillips G, O'Neill HM, Coates L, et al. (2020) Malleability of the  
562 SARS-CoV-2 3CL M(pro) Active-Site Cavity Facilitates Binding of Clinical Antivirals.  
563 Structure 28: 1313-1320 e1313.
- 564 10. Zhang L, Lin D, Sun X, Curth U, Drosten C, et al. (2020) Crystal structure of SARS-CoV-2  
565 main protease provides a basis for design of improved alpha-ketoamide inhibitors. Science  
566 368: 409-412.
- 567 11. Jin Z, Du X, Xu Y, Deng Y, Liu M, et al. (2020) Structure of M(pro) from SARS-CoV-2 and  
568 discovery of its inhibitors. Nature 582: 289-293.
- 569 12. Kumar Y, Singh H, Patel CN (2020) In silico prediction of potential inhibitors for the main  
570 protease of SARS-CoV-2 using molecular docking and dynamics simulation based drug-  
571 repurposing. J Infect Public Health 13: 1210-1223.
- 572 13. Vuong W, Khan MB, Fischer C, Arutyunova E, Lamer T, et al. (2020) Feline coronavirus drug  
573 inhibits the main protease of SARS-CoV-2 and blocks virus replication. Nat Commun 11:  
574 4282.
- 575 14. Anderson J, Schiffer C, Lee SK, Swanstrom R (2009) Viral Protease Inhibitors. In: Kräusslich  
576 H, Bartenschlager R, editors. Antiviral Strategies: Springer. pp. 1-24.
- 577 15. Ma C, Sacco MD, Hurst B, Townsend JA, Hu Y, et al. (2020) Boceprevir, GC-376, and calpain  
578 inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease.  
579 Cell Res 30: 678-692.
- 580 16. Kim Y, Lovell S, Tiew KC, Mandadapu SR, Alliston KR, et al. (2012) Broad-spectrum  
581 antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and  
582 coronaviruses. J Virol 86: 11754-11762.
- 583 17. Kim Y, Liu H, Galasiti Kankanamalage AC, Weerasekara S, Hua DH, et al. (2016) Reversal  
584 of the Progression of Fatal Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus  
585 Protease Inhibitor. PLoS Pathog 12: e1005531.
- 586 18. Rathnayake AD, Zheng J, Kim Y, Perera KD, Mackin S, et al. (2020) 3C-like protease  
587 inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV-  
588 infected mice. Sci Transl Med 12.
- 589 19. Galasiti Kankanamalage AC, Kim Y, Weerawarna PM, Uy RA, Damalanka VC, et al. (2015)  
590 Structure-guided design and optimization of dipeptidyl inhibitors of norovirus 3CL  
591 protease. Structure-activity relationships and biochemical, X-ray crystallographic, cell-  
592 based, and in vivo studies. J Med Chem 58: 3144-3155.
- 593 20. Kim Y, Shivanna V, Narayanan S, Prior AM, Weerasekara S, et al. (2015) Broad-spectrum  
594 inhibitors against 3C-like proteases of feline coronaviruses and feline caliciviruses. J Virol  
595 89: 4942-4950.
- 596 21. Galasiti Kankanamalage AC, Kim Y, Damalanka VC, Rathnayake AD, Fehr AR, et al. (2018)  
597 Structure-guided design of potent and permeable inhibitors of MERS coronavirus 3CL  
598 protease that utilize a piperidine moiety as a novel design element. Eur J Med Chem 150:  
599 334-346.
- 600 22. Perera KD, Galasiti Kankanamalage AC, Rathnayake AD, Honeyfield A, Groutas W, et al.  
601 (2018) Protease inhibitors broadly effective against feline, ferret and mink coronaviruses.  
602 Antiviral Res 160: 79-86.
- 603 23. Chang KO, Kim Y, Lovell S, Rathnayake AD, Groutas WC (2019) Antiviral Drug Discovery:  
604 Norovirus Proteases and Development of Inhibitors. Viruses 11.
- 605 24. Drag M, Salvesen GS (2010) Emerging principles in protease-based drug discovery. Nat Rev  
606 Drug Discov 9: 690-701.

- 607 25. Hoffman RL, Kania RS, Brothers MA, Davies JF, Ferre RA, et al. (2020) Discovery of Ketone-  
608 Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic  
609 Treatment of COVID-19. *J Med Chem* 63: 12725-12747.
- 610 26. Turk B (2006) Targeting proteases: successes, failures and future prospects. *Nat Rev Drug*  
611 *Discov* 5: 785-799.
- 612 27. Sinclair JF, Ziegler MM, Baldwin TO (1994) Kinetic partitioning during protein folding yields  
613 multiple native states. *Nat Struct Biol* 1: 320-326.
- 614 28. Chen J, Wharton SA, Weissenhorn W, Calder LJ, Hughson FM, et al. (1995) A soluble domain  
615 of the membrane-anchoring chain of influenza virus hemagglutinin (HA2) folds in  
616 *Escherichia coli* into the low-pH-induced conformation. *Proc Natl Acad Sci U S A* 92:  
617 12205-12209.
- 618 29. Lai Z, McCulloch J, Lashuel HA, Kelly JW (1997) Guanidine hydrochloride-induced  
619 denaturation and refolding of transthyretin exhibits a marked hysteresis: equilibria with  
620 high kinetic barriers. *Biochemistry* 36: 10230-10239.
- 621 30. Barrientos LG, Louis JM, Botos I, Mori T, Han Z, et al. (2002) The domain-swapped dimer of  
622 cyanovirin-N is in a metastable folded state: reconciliation of X-ray and NMR structures.  
623 *Structure* 10: 673-686.
- 624 31. Fasshauer D, Antonin W, Subramaniam V, Jahn R (2002) SNARE assembly and disassembly  
625 exhibit a pronounced hysteresis. *Nat Struct Biol* 9: 144-151.
- 626 32. Malcolm BA, Lowe C, Shechosky S, McKay RT, Yang CC, et al. (1995) Peptide aldehyde  
627 inhibitors of hepatitis A virus 3C proteinase. *Biochemistry* 34: 8172-8179.
- 628 33. Maseko S, Padayachee E, Maphumulo S, Govender T, Sayed Y, et al. (2019) Kinetic and  
629 thermodynamic characterisation of HIV-protease inhibitors against E35D upward arrowG  
630 upward arrowS mutant in the South African HIV-1 subtype C protease. *J Enzyme Inhib*  
631 *Med Chem* 34: 1451-1456.
- 632 34. Gouzi H, Depagne C, Coradin T (2012) Kinetics and thermodynamics of the thermal  
633 inactivation of polyphenol oxidase in an aqueous extract from *Agaricus bisporus*. *J Agric*  
634 *Food Chem* 60: 500-506.
- 635 35. Dogan N, Tari C (2008) Characterization of three-phase partitioned exo-polygalacturonase  
636 from *Aspergillus sojae* with unique properties. *Biochemical Engineering Journal* 39: 43-  
637 50.
- 638 36. Hu T, Zhang Y, Li L, Wang K, Chen S, et al. (2009) Two adjacent mutations on the dimer  
639 interface of SARS coronavirus 3C-like protease cause different conformational changes in  
640 crystal structure. *Virology* 388: 324-334.
- 641 37. Chen S, Zhang J, Hu T, Chen K, Jiang H, et al. (2008) Residues on the dimer interface of  
642 SARS coronavirus 3C-like protease: dimer stability characterization and enzyme catalytic  
643 activity analysis. *J Biochem* 143: 525-536.
- 644 38. Xue X, Yang H, Shen W, Zhao Q, Li J, et al. (2007) Production of authentic SARS-CoV  
645 M(pro) with enhanced activity: application as a novel tag-cleavage endopeptidase for  
646 protein overproduction. *J Mol Biol* 366: 965-975.
- 647 39. Lee J, Worrall LJ, Vuckovic M, Rosell FI, Gentile F, et al. (2020) Crystallographic structure  
648 of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-  
649 terminal autoprocessing site. *Nat Commun* 11: 5877.
- 650 40. Thorarensen A, Balbo P, Banker ME, Czerwinski RM, Kuhn M, et al. (2021) The advantages  
651 of describing covalent inhibitor in vitro potencies by IC50 at a fixed time point. *IC50*

- 652           determination of covalent inhibitors provides meaningful data to medicinal chemistry for  
653           SAR optimization. *Bioorg Med Chem* 29: 115865.
- 654 41. Howe AY, Venkatraman S (2013) The Discovery and Development of Boceprevir: A Novel,  
655           First-generation Inhibitor of the Hepatitis C Virus NS3/4A Serine Protease. *J Clin Transl*  
656           *Hepatol* 1: 22-32.
- 657 42. Soumana DI, Kurt Yilmaz N, Ali A, Prachanronarong KL, Schiffer CA (2016) Molecular and  
658           Dynamic Mechanism Underlying Drug Resistance in Genotype 3 Hepatitis C NS3/4A  
659           Protease. *J Am Chem Soc* 138: 11850-11859.
- 660 43. Anand K, Ziebuhr J, Wadhvani P, Mesters JR, Hilgenfeld R (2003) Coronavirus main  
661           proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* 300: 1763-  
662           1767.
- 663 44. Shi JH, Song JX (2006) The catalysis of the SARS 3C-like protease is under extensive  
664           regulation by its extra domain. *Febs Journal* 273: 1035-1045.
- 665 45. Ye G, Wang X, Tong X, Shi Y, Fu ZF, et al. (2020) Structural Basis for Inhibiting Porcine  
666           Epidemic Diarrhea Virus Replication with the 3C-Like Protease Inhibitor GC376. *Viruses*  
667           12.
- 668 46. Sadana A (1991) *Biocatalysis: fundamentals of enzyme deactivation kinetics*. Englewood  
669           Cliffs, N.J. : Prentice Hall
- 670 47. Gonzalez A, Moorhead P, McPhillips SE, Song J, Sharp K, et al. (2008) Web-Ice: integrated  
671           data collection and analysis for macromolecular crystallography. *Journal of Applied*  
672           *Crystallography* 41: 176-184.
- 673 48. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66: 125-132.
- 674 49. Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, et al. (2019) Macromolecular  
675           structure determination using X-rays, neutrons and electrons: recent developments in  
676           Phenix. *Acta Crystallogr D Struct Biol* 75: 861-877.
- 677 50. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al. (2010) PHENIX: a  
678           comprehensive Python-based system for macromolecular structure solution. *Acta*  
679           *Crystallogr D Biol Crystallogr* 66: 213-221.
- 680

681

Protease	IC <sub>50</sub> , μM	Calculated K <sub>i</sub> , μM
SARS-CoV M <sup>pro</sup>	0.05± 0.01	0.02
SARS-CoV-2 M <sup>pro</sup>	0.19± 0.04	0.04

682

683 **Table 1.** Comparison of IC<sub>50</sub> and K<sub>i</sub> values between SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup> with  
 684 compound GC376. Data are presented as mean ± SEM, *n* = 3.

685

A

686

t (°C)	kd (min <sup>-1</sup> )	t <sub>1/2</sub> (min)	t <sub>1/2</sub> (h)	Dt (min)	Dt(h)	ΔH (kJ/mol)	ΔG (kJ/mol)	Δ S (kJ/mol*K)
24	0.0003	2310.5	38.5	7675.3	127.9	13.9	82.6	-0.23
37	0.0004	1732.9	28.8	5756.5	95.9	240.9	85.6	0.50
45	0.0011	630.1	10.5	2093.3	34.9	240.9	85.2	0.48
55	0.318	2.2	0.04	7.2	0.121	240.8	72.5	0.52
70	1.67	0.4	0.007	1.4	0.023	240.6	71.2	0.52

B

690

t (°C)	kd (min <sup>-1</sup> )	t <sub>1/2</sub> (min)	t <sub>1/2</sub> (h)	Dt (min)	Dt(h)	ΔH (kJ/mol)	ΔG (kJ/mol)	Δ S (kJ/mol*K)
24	0.0002	3465.7	57.7	11512.9	191.8	38.9	83.6	-0.15
37	0.0004	1732.8	28.8	5756.4	95.9	176.5	85.6	0.29
45	0.001	693.1	11.5	2302.5	38.3	176.4	85.4	0.28
55	0.23	3	0.05	10.01	0.2	176.4	73.4	0.31
70	1.2	0.6	0.009	1.9	0.03	176.2	72.2	0.32

693

694

695

696

697 **Table 2.** Thermodynamic parameters for the thermal inactivation of (A) SARS-CoV M<sup>pro</sup> and (B)

698 SARS-CoV-2 M<sup>pro</sup>. *T*, the temperature in °C, *kd*, inactivation rate constant, *t*<sub>1/2</sub>, half-life of

699 proteases (i.e., the time after which activity is reduced to one-half of the initial value), *Dt*, decimal

700 reduction time, the time required to reduce the enzymatic activity to 10% of its original value, *ΔG*,

701 activation free energy barrier, *ΔH*, activation enthalpy, *ΔS* activation entropy of thermal

702 denaturation.

703

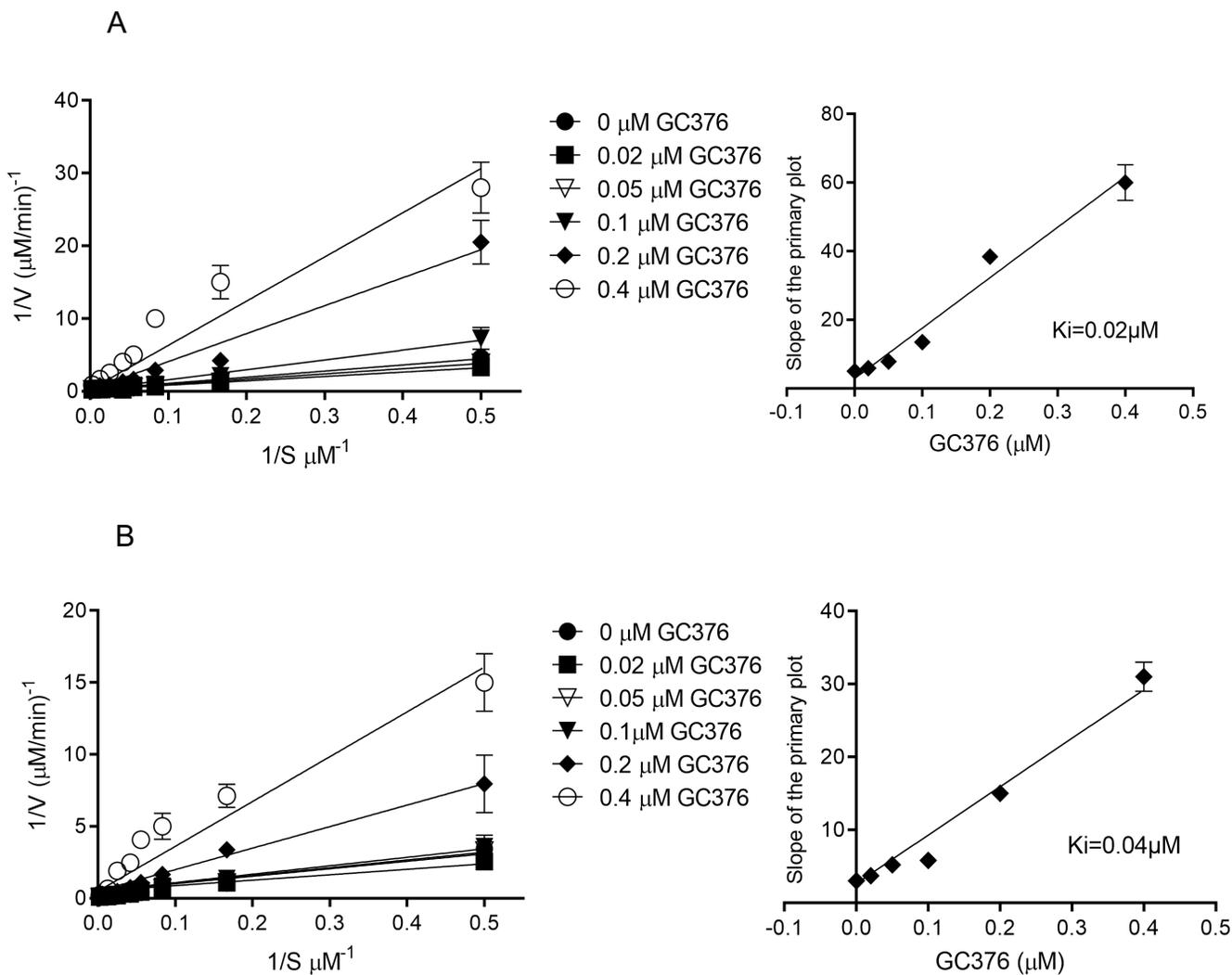
	SARS-CoV M <sup>pro</sup> GC373	SARS-CoV M <sup>pro</sup> GC376
PDB entry	7LCP	7LCQ
<b>Data collection</b>		
Space group	C2	C2
Cell dimensions		
<i>a, b, c</i> (Å)	108.28 82.19 54.06	108.27, 82.05, 53.82
$\alpha, \beta, \gamma$ (°)	90, 104.33, 90	90, 104.11, 90
Resolution (Å)	33.18 - 1.9 (1.968 - 1.9)	39.19 - 2.15 (2.22 - 2.15)
Observations	236351 (23994)	110459 (10256)
<i>R</i> <sub>merge</sub>	0.053 (1.035)	0.041 (1.76)
<i>I</i> / $\sigma$ <i>I</i>	16.12 (2.08)	14.7 (0.75)
Completeness (%)	98.04 (97.91)	96.94 (93.68)
Redundancy	6.6 (6.7)	4.5 (4.4)
CC1/2	99.90 (86.70)	99.90 (60.70)
<b>Refinement</b>		
Resolution (Å)	33.18-1.90	33.19-2.15
No. reflections	35487	24146
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	19.53/22.08	19.58/22.76
No. atoms	2438	2375
Protein	2358	2343
Ligand/ion	29	29
Water	51	3
<i>B</i> -factors	72.13	98.79
Protein	72.31	98.98
Ligand/ion	70.03	88.98
Water	65.22	72.66
R.m.s. deviations		
Bond lengths (Å)	0.010	0.016
Bond angles (°)	1.42	1.73

704 \*Values in parentheses are for highest-resolution shell. Each data were collected from single crystal

705

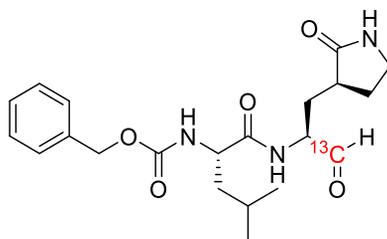
706 **Table 3.** Data collection and refinement statistics (molecular replacement) for SARS-CoV M<sup>pro</sup>

707 with drug GC373 and prodrug GC376.

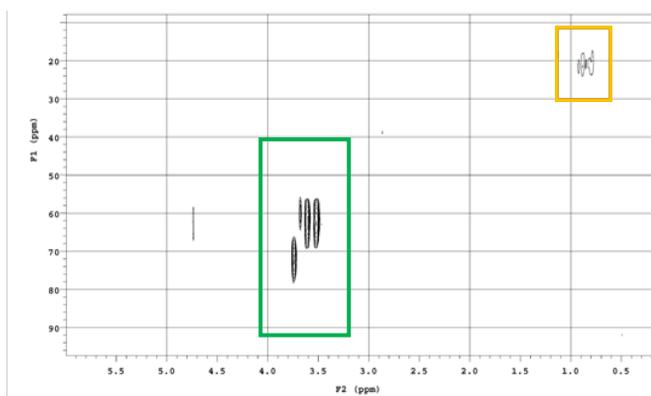


**Fig 1. Determination of  $K_i$  values of GC376 for SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup>.** Lineweaver-Burk plots (left) and the secondary plots of competitive inhibition (right) of SARS-CoV M<sup>pro</sup> (A) and SARS-CoV-2 M<sup>pro</sup> (B) by GC376. Data are presented as mean  $\pm$  SEM,  $n = 3$

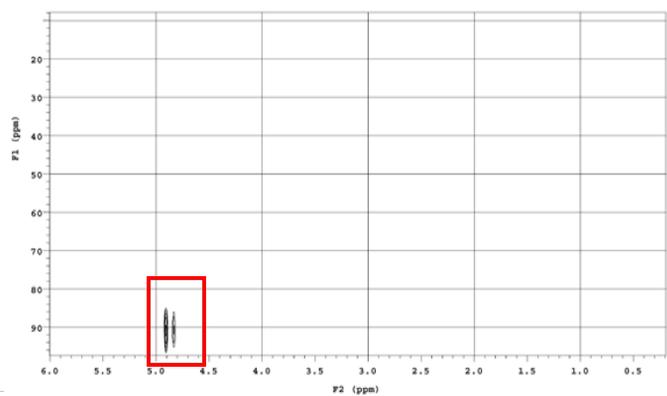
A



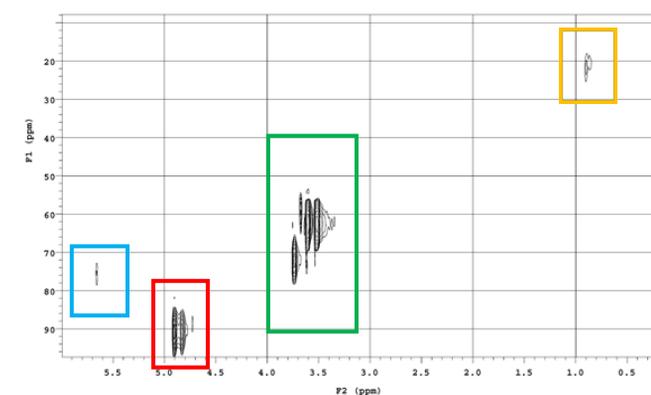
B



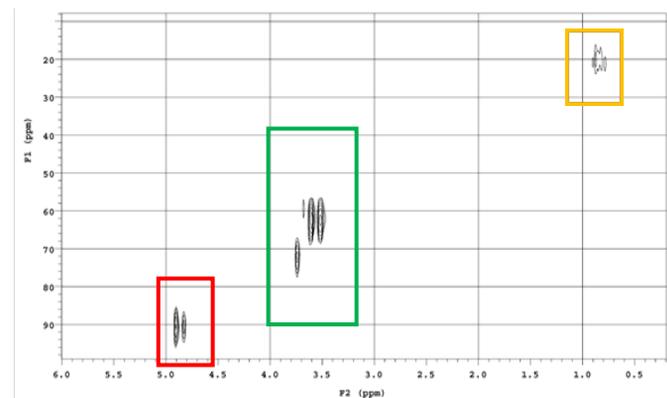
C



D

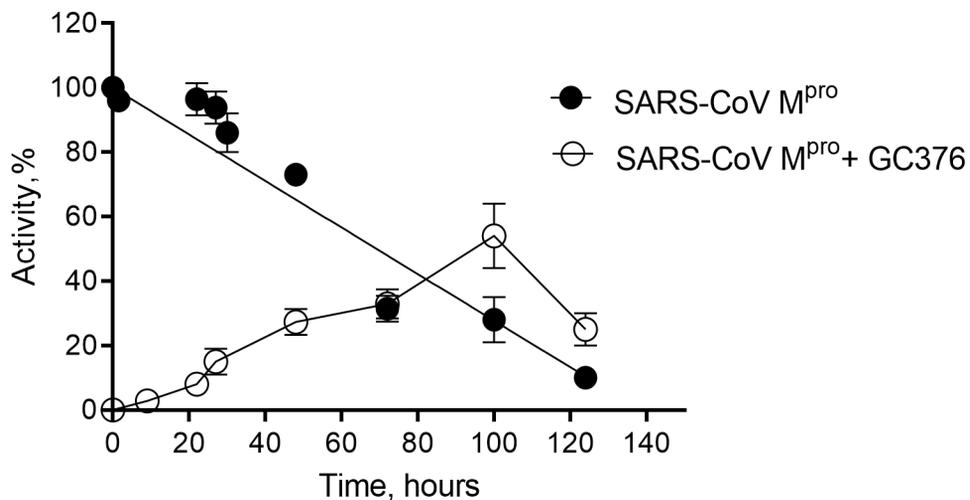


E

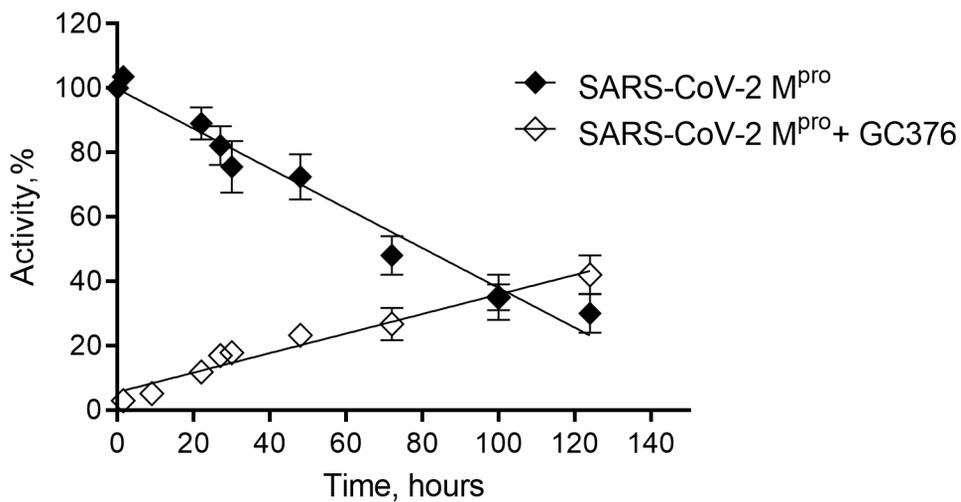


**Figure 2:** HSQC NMR experiments examining reversibility of GC376 binding. (A) Structure of  $^{13}\text{C}$ -labelled GC376. (B) HSQC spectra of SARS-CoV-2 M<sup>pro</sup> in deuterated buffer. (C) HSQC spectra of  $^{13}\text{C}$ -labelled GC376. (D) Co-incubation of SARS-CoV-2 M<sup>pro</sup> with  $^{13}\text{C}$ -labelled GC376. (E) Co-incubated sample after washing step with buffer. Boxes: Blue = bound inhibitor; Red = free inhibitor; Green = DTT (from buffer); Orange = SARS-CoV-2 M<sup>pro</sup>.

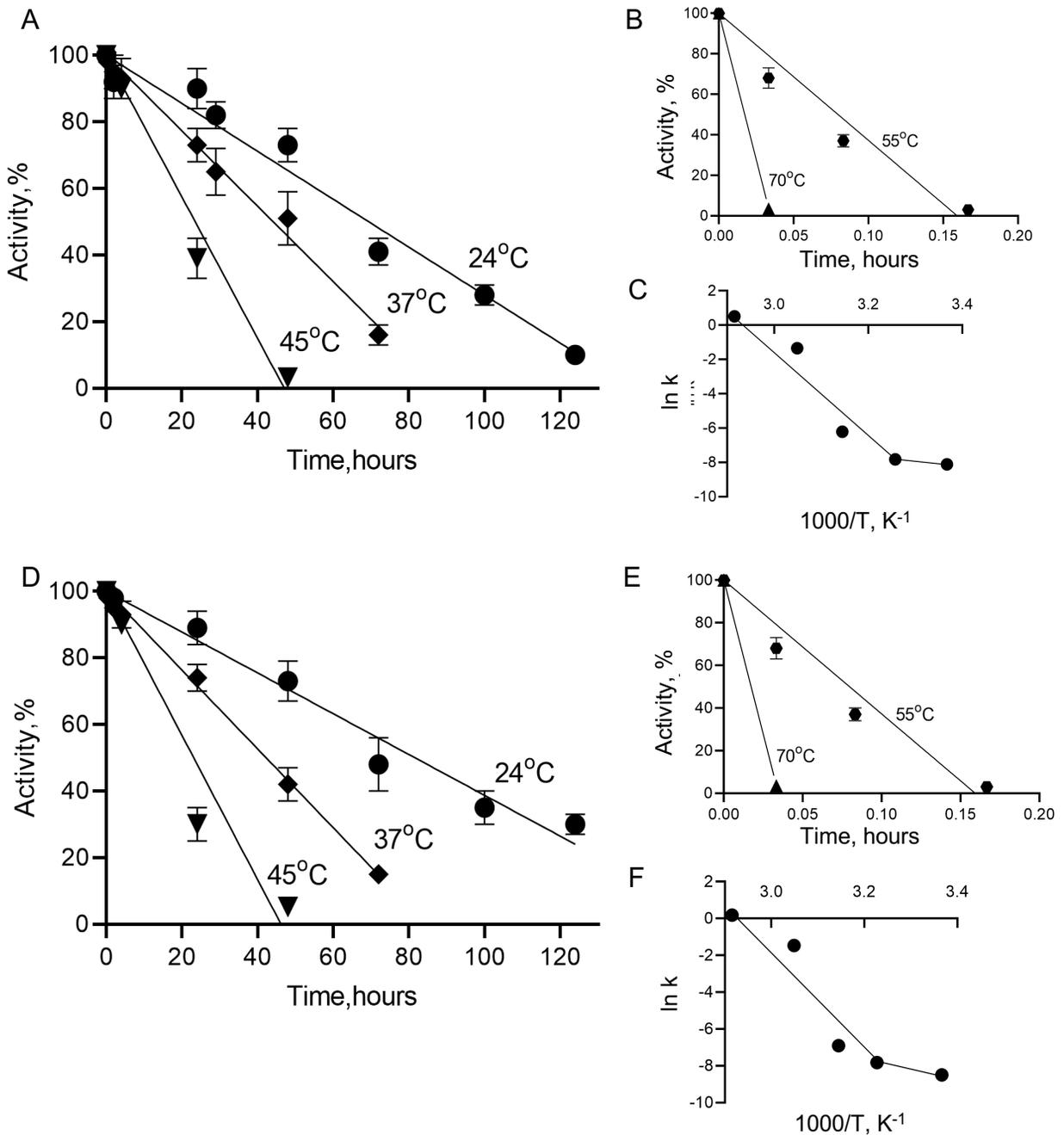
A



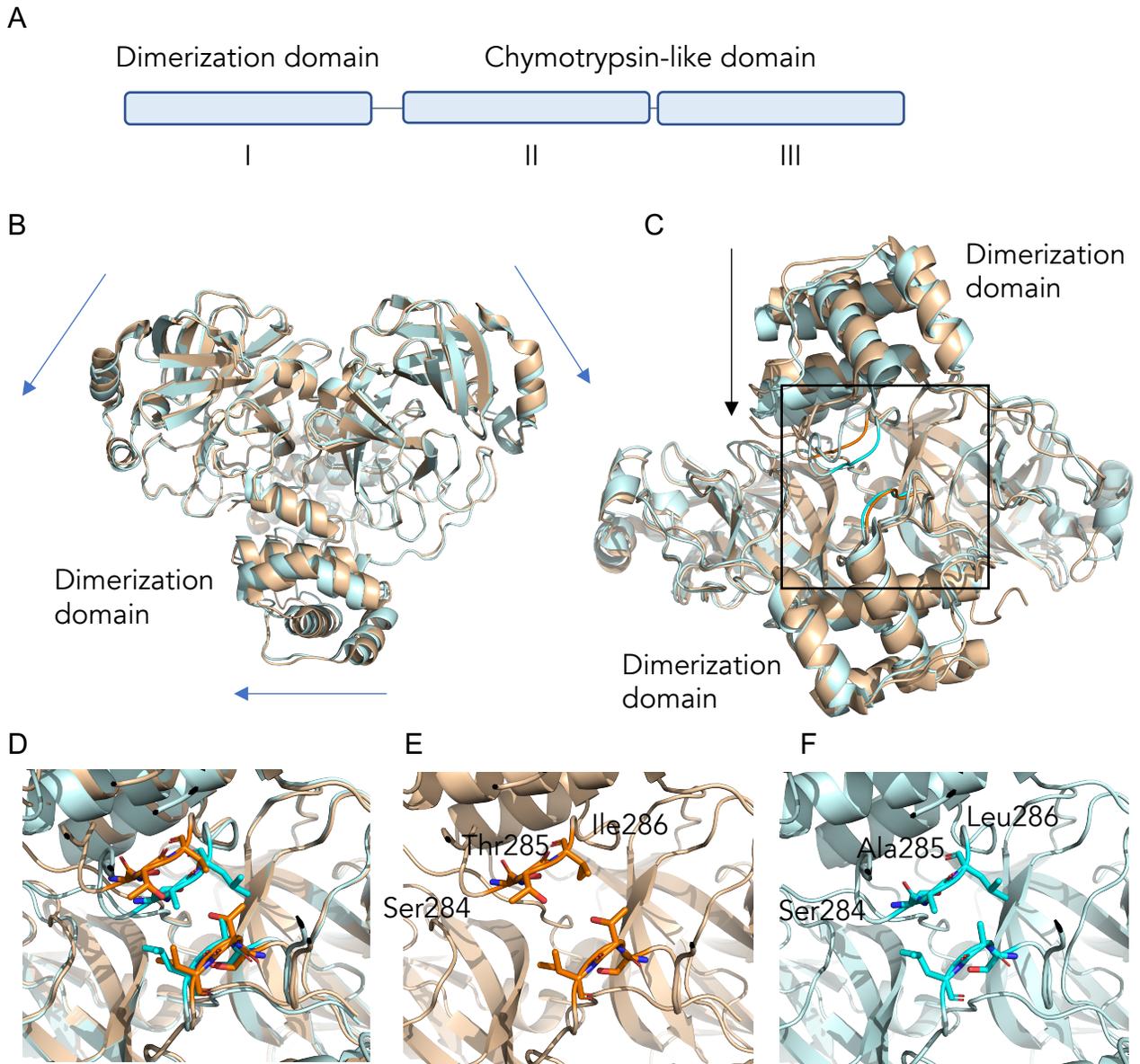
B



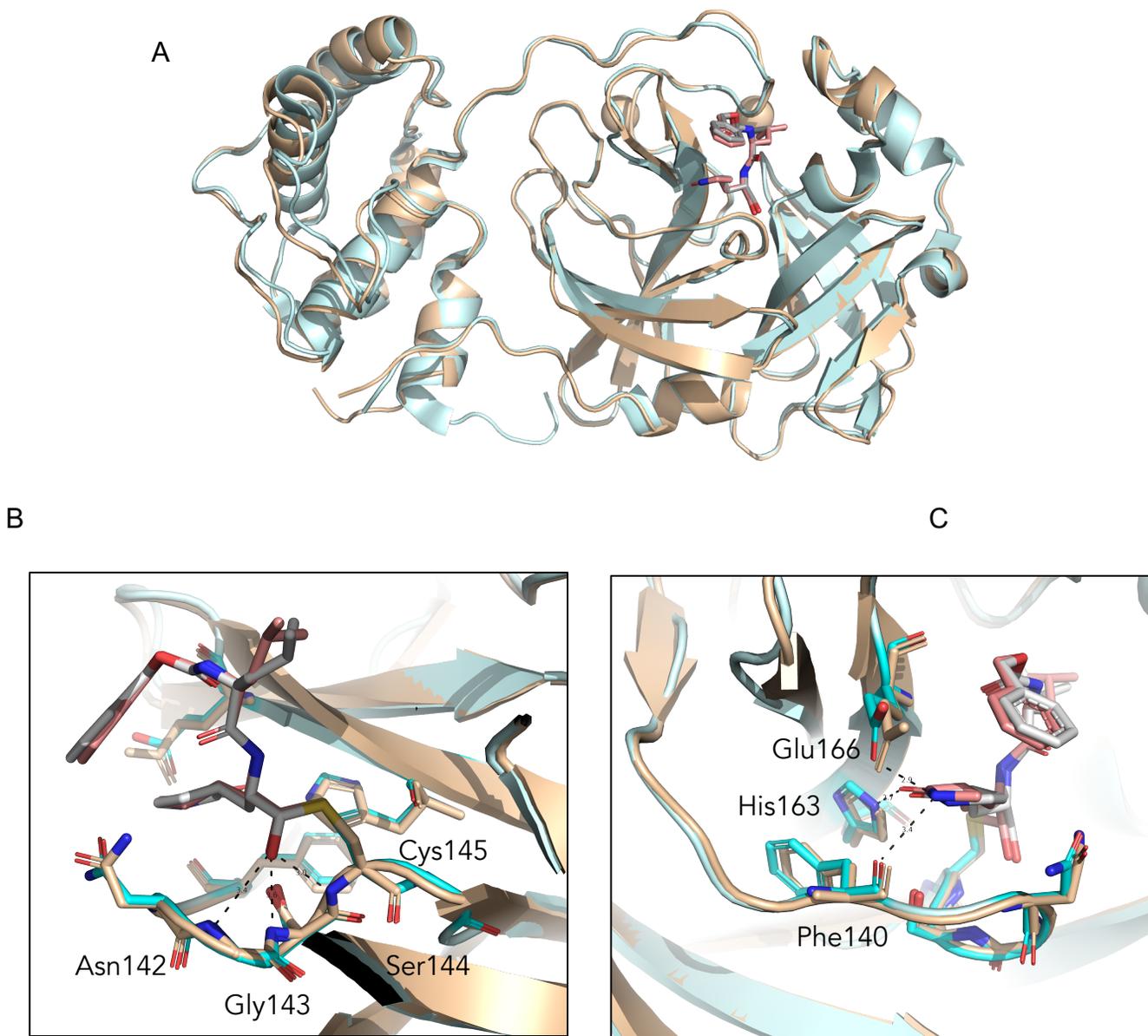
**Fig 3. Reversibility of GC376 with SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup>.** The dependence of activity of SARS-CoV M<sup>pro</sup> (A) and SARS-CoV-2 M<sup>pro</sup> (B) incubated alone and with the bound GC376 compound on time. Data are presented as mean  $\pm$  SEM,  $n = 3$



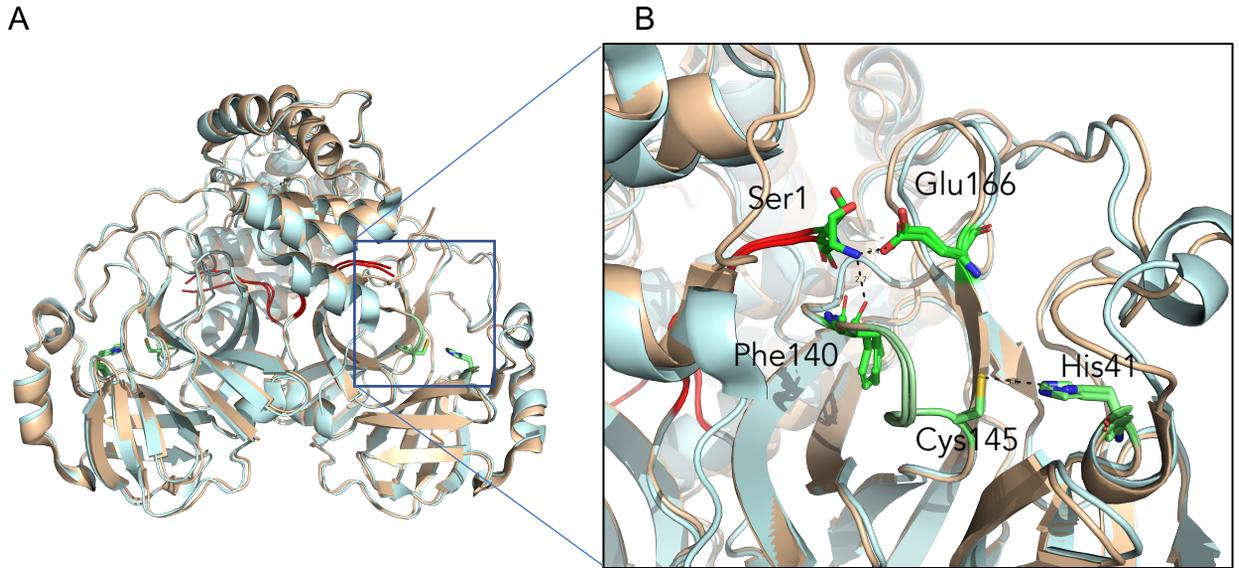
**Fig 4. Thermal stability study of SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup>.** Time course of residual activities of SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup> in temperature ranges of 24-45°C (A and D, respectively) and in 55-70°C (B and E) and Arrhenius plots for SARS-CoV M<sup>pro</sup> and SARS-CoV-2 M<sup>pro</sup> (C and F, respectively). Data are presented as mean ± SEM, *n* = 2



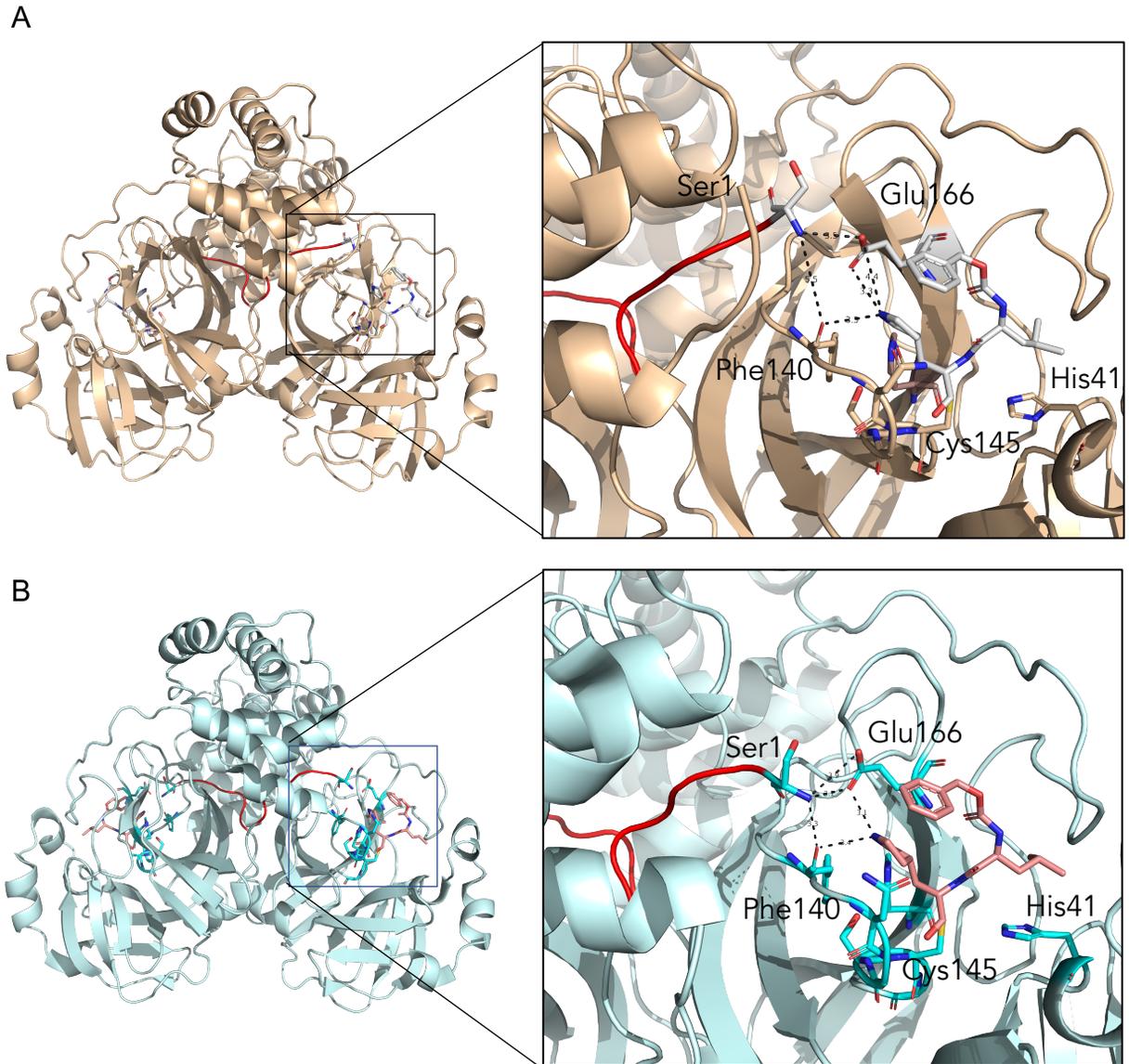
**Fig 5. Differences observed between SARS-CoV and SARS-CoV-2 M<sup>pro</sup> structures.** (A) Overall domain organization of M<sup>pro</sup>. (B) Overlay of SARS-CoV M<sup>pro</sup> (2DUC.pdb) and SARS-CoV-2 M<sup>pro</sup> (6WTM.pdb) structures reveals small global shifts. (C) The largest structural change is the closer distance between the dimer interface of M<sup>pro</sup> in SARS-CoV-2 compared to SARS-CoV. (D) A close examination of the dimerization loop in both SARS-CoV M<sup>pro</sup> (orange) and SARS-CoV-2 M<sup>pro</sup> (blue). (E) In SARS-CoV M<sup>pro</sup> a Thr285 in the STI sequence at the dimer interface participates in dimerization via hydrophobic interactions while the M<sup>pro</sup> in SARS-CoV-2 (F) has an alanine in a SAL motif resulting in a zippered interdigitation of the hydrophobic residues and closer association of the dimerization domains.



**Figure 6. Comparison of SARS-CoV and SARS-CoV-2 M<sup>pro</sup> structures with GC373 drug.** (A). Overall 3-dimensional structures show similarities between SARS-CoV M<sup>pro</sup> (wheat) and SARS CoV-2 M<sup>pro</sup> (cyan) with an RMSD of 0.6 Å. GC373 binds covalently with the catalytic Cys145 of the M<sup>pro</sup> of both SARS-CoV (7LCP.pdb) and SARS-CoV-2 (6TWK.pdb) and shows (B) similar oxanyonion hole coordination by Ser144, Gly143 and Asn142 and (C) drug adduct coordination with side chains of His163, Glu166 and backbone of Phe140.



**Figure 7. The N-terminus of one protomer interacts with the active site region of the other protomer.** (A) Comparison of SARS-CoV M<sup>pro</sup> (wheat) and SARS-CoV-2 M<sup>pro</sup> (cyan) structures reveals the N-terminus of each protomer (red) participates in domain swapping in the other protomer. (B) Hydrogen bonding with the N-terminal Ser1 occurs with the side chain of Glu166 and backbone oxygen of Phe140. This influences the region adjacent to the catalytic residues Cys145 and His41, and the oxyanion hole, colored in green.



**Figure 8 . The N-terminal finger of the M<sup>Pro</sup> stabilizes dimer formation and coordination of the drug GC373.** The N-finger of M<sup>Pro</sup> facilitates coordination of drug GC373 in both in SARS-CoV (7LCQ.PDB) (A) and SARS-CoV-2 (6WTJ.PDB) (B). Overall of 3-dimensional structures show the N-terminus (red) inserts into the second protomer in SARS-CoV M<sup>Pro</sup> (wheat) and in SARS-CoV-2 M<sup>Pro</sup> (cyan). Both residues that coordinate the N-finger, Phe140 and Glu166 also interact with the P1 position of the drug in M<sup>Pro</sup> of both SARS-CoV and SARS-CoV-2.