1 **Structural basis for the inhibition of SARS-CoV-2 main protease by**

2 **antineoplastic drug Carmofur**

3 Zhenming Jin[1,2*], Yao Zhao[1*], Yuan Sun[3*], Bing Zhang[1], Haofeng Wang[1,4], Yan Wu[3],

4 Yan Zhu[1], Chen Zhu[1], Tianyu Hu[1], Xiaoyu Du[1,2], Yinkai Duan[1], Jing Yu[1], Xiaobao

5 Yang[1], Xiuna Yang[1], Kailin Yang[5], Xiang Liu[6], Luke W. Guddat[7], Gengfu Xiao[3], Leike

6 Zhang[3†], Haitao Yang[1†] & Zihe Rao[1,2,6]

7 [1]Shanghai Institute for Advanced Immunochemical Studies and School of Life Science

8 and Technology, ShanghaiTech University, Shanghai, China.

9 [2]Laboratory of Structural Biology, School of Life Sciences and School of Medicine,

10 Tsinghua University, Beijing, China.

11 [3]State Key Laboratory of Virology, Wuhan Institute of Virology, Center for Biosafety

12 Mega-Science, Chinese Academy of Sciences, Wuhan, China.

13 [4]School of Life Sciences, Tianjin University, Tianjin, China.

14 [5]Taussig Cancer Center, Cleveland Clinic, Cleveland, OH, USA.

15 [6]State Key Laboratory of Medicinal Chemical Biology, Frontiers Science Center for

16 Cell Response, College of Life Sciences, College of Pharmacy, Nankai University,

17 Tianjin, China.

18 [7]School of Chemistry and Molecular Biosciences, the University of Queensland,

19 Brisbane, Australia.

20 [*]These authors contributed equally: Zhenming Jin, Yao Zhao and Yuan Sun.

21 [†]e-mail: zhangleike@wh.iov.cn; yanght@shanghaitech.edu.cn

22

23 **Abstract**

24 The antineoplastic drug Carmofur was shown to inhibit SARS-CoV-2 main protease

25 (M^{pro}). Here the X-ray crystal structure of M^{pro} in complex with Carmofur reveals that

26 the carbonyl reactive group of Carmofur is covalently bound to catalytic Cys145,

27 whereas its fatty acid tail occupies the hydrophobic S2 subsite. Carmofur inhibits viral

28 replication in cells ($EC_{50} = 24.30$ μM) and it is a promising lead compound to develop

29 new antiviral treatment for COVID-19.

30

31   COVID-19, a highly infectious viral disease, has spread since its appearance in
32   December 2019, causing an unprecedented pandemic. The number of confirmed cases
33   worldwide continues to grow at a rapid rate, but there are no specific drugs or vaccines
34   available to control symptoms or the spread of this disease at this time.

35   The etiological agent of the disease is the coronavirus SARS-CoV-2. This virus has
36   a ~30,000 nt RNA genome. The N-terminus of the viral genome encodes two
37   translational products, polyproteins 1a and 1ab (pp1a and pp1ab)[1,2], which are
38   processed into mature non-structural proteins, by the main protease (M^{pro}) and a papain-
39   like protease[3]. M^{pro} has been proposed as a therapeutic target for anti-coronavirus (CoV)
40   drug development[4-6]. We previously screened over 10,000 compounds and identified
41   Carmofur as compound that can inhibit M^{pro} *in vitro*, with an IC_{50} of 1.82 μM[7].

42   Carmofur (1-hexylcarbamoyl-5-fluorouracil) is a derivative of 5-fluorouracil (5-
43   FU) (Fig. 1a) and an approved antineoplastic agent. Carmofur has been used to treat
44   colorectal cancer since 1980s[8], and has shown clinical benefits on breast, gastric, and
45   bladder cancers[9-11]. The target for Carmofur is believed to be thymidylate synthase[12,13],
46   but it has also been shown to inhibit human acid ceramidase (AC)[14], through covalent
47   modification of its catalytic cysteine[15].

48   The molecular details for how Carmofur inhibits M^{pro} activity were unresolved.
49   Here, we present the 1.6 Å X-ray crystal structure of SARS-CoV-2 M^{pro} in complex
50   with Carmofur (Fig. 1b, c and Supplementary Table 1). In agreement with previous
51   studies[4,5,16-18], M^{pro} forms a homodimer (protomer A and B) related by crystallographic
52   symmetry (Extended Data Fig. 1a, b). All of the residues (1–306) in the polypeptide
53   could be traced in the electron density map. Each protomer is composed of three
54   domains (Fig. 1c): domain I (residues 10–99), domain II (residues 100–184), and
55   domain III (residues 201–303); and a long loop region (residues 185–200) connects
56   domains II and III. The substrate-binding pocket lies in the cleft between domain I and
57   domain II feature the catalytic dyad residues Cys145 and His41 (Fig. 1c, d). The
58   substrate-binding pocket is divided into a series of subsites (including S1, S2, S4, and
59   S1′), each accommodating a single but consecutive amino acid residue in the substrate.
60   The first residue serine of one protomer interacts with residue Phe140 and Glu166 of
61   the other protomer to stabilize the S1 subsite (Extended Data Fig. 1c), and this structural
62   feature is essential for catalysis[7].

63   The electron density map unambiguously shows that the fatty acid moiety
64   ($C_7H_{14}NO$) of Carmofur is linked to the Sγ atom of Cys145 through a 1.8 Å covalent

bond, whereas the fatty acid tail is inserted into the S2 subsite (Fig. 1d, e). This observation suggests that the sulfhydryl group of Cys145 attacks the electrophilic carbonyl group of Carmofur, resulting in covalent modification of the cysteine residue and release of the 5-fluorouracil moiety (Fig. 1b and Extended Data Fig. 2a). In addition to the C-S covalent bond (highlighted in Fig. 1e), the inhibitor is stabilized by numerous hydrogen bonds and hydrophobic interactions (Fig. 1e and Extended Data Fig. 2b). The carbonyl oxygen of Carmofur occupies the oxyanion hole and forms hydrogen bonds (3.0 Å) with the backbone amides of Gly143 and Cys145, mimicking the tetrahedral oxyanion intermediate formed during protease cleavage (Fig. 1e). The fatty acid tail, which appears in an extended conformation, inserts into the bulky hydrophobic S2 subsite (composed of the side chains of His41, Met49, Tyr54, Met165, and the alkyl portion of the side chain of Asp187) (Fig. 1d, e). The hydrophobic interactions are mainly contributed by the side chains of His41, Met49 and Met165, all of which run parallel with the alkyl part of the fatty acid tail of the inhibitor (Fig. 1e and Extended Data Fig. 2b).

The mechanism of covalent modification by Carmofur is different from that of the inhibitor N3[7,19], which covalently modifies Cys145 through Michael addition of the vinyl group. The structures of M^{pro}-Carmofur and M^{pro}-N3 are overall similar (r.m.s.d. of 0.286 Å for all Cα atoms). The largest conformational differences occur in the substrate binding pocket, with the backbone surrounding Carmofur in a slightly more outward position compared with the M^{pro}-N3 complex structure (Extended Data Fig. 3a). Another difference is that Carmofur only occupies the S2 subsite (Fig. 1d), whereas N3 occupies four subsites (S1, S2, S4 and S1′, see Extended Data Fig. 3b, c). The lactam ring of N3 is located in the S1 subsite, which is filled by a DMSO molecule in the M^{pro}-Carmofur structure (Extended Data Fig. 3b, c). These observations demonstrate the potential for structural elaboration of Carmofur and will be useful to design more potent derivatives against the M^{pro} of SARS-CoV-2.

We previously showed that treatment with 10 µM Ebselen (EC$_{50}$ = 4.67 µM) inhibited infection of Vero cells with SARS-CoV-2 whereas Carmofur did not showed detectable antiviral activity at this concentration[7]. Here we determined the inhibitory effect of Carmofur against SARS-CoV-2 infection on Vero E6 cells, as previously described[20] (Fig. 2). By measuring viral RNA in supernatant, we determined the EC$_{50}$ for Carmofur as 24.30 µM (Fig. 2a). To verify this result, we fixed infected cells and stained them using anti-sera against viral nucleocapsid protein (NP) and observed a

99 decrease in NP levels after Carmofur treatment (Fig. 2b). We also performed

100 cytotoxicity assays for Carmofur in Vero E6 cells and determined the CC_{50} value of

101 133.4 µM (Fig. 2c). Thus, Carmofur has a favorable selectivity index (SI) of 5.36, but

102 further optimization will be required to develop into an effective drug.

103 In conclusion, the crystal structure of M^{pro} in complex with Carmofur shows that

104 the compound directly modifies the catalytic Cys145 of SARS-CoV-2 M^{pro}. Our study

105 also provides a basis for rational design of Carmofur analogs with enhanced inhibitory

106 efficacy to treat COVID-19. Since M^{pro} is highly conserved among all CoV M^{pro}s,

107 Carmofur and its analogs may be effective against a broader spectrum of coronaviruses.

References

109 1. Wu, F. et al. A new coronavirus associated with human respiratory disease in
110 China. *Nature* **579**, 265-269 (2020).
111 2. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of
112 probable bat origin. *Nature* **579**, 270-273 (2020).
113 3. Hegyi, A. & Ziebuhr, J. Conservation of substrate specificities among
114 coronavirus main proteases. *J Gen Virol* **83**, 595-599 (2002).
115 4. Anand, K. et al. Structure of coronavirus main proteinase reveals combination
116 of a chymotrypsin fold with an extra alpha-helical domain. *Embo j* **21**, 3213-24
117 (2002).
118 5. Yang, H. et al. The crystal structures of severe acute respiratory syndrome virus
119 main protease and its complex with an inhibitor. *Proc Natl Acad Sci U S A* **100**,
120 13190-5 (2003).
121 6. Pillaiyar, T., Manickam, M., Namasivayam, V., Hayashi, Y. & Jung, S.H. An
122 Overview of Severe Acute Respiratory Syndrome-Coronavirus (SARS-CoV)
123 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy.
124 *J Med Chem* **59**, 6595-628 (2016).
125 7. Jin, Z. et al. Structure of Mpro from COVID-19 virus and discovery of its
126 inhibitors. *Nature* (2020).
127 8. Sakamoto, J. et al. An individual patient data meta-analysis of adjuvant therapy
128 with carmofur in patients with curatively resected colon cancer. *Jpn J Clin*
129 *Oncol* **35**, 536-44 (2005).
130 9. Morimoto, K. & Koh, M. Postoperative adjuvant use of carmofur for early
131 breast cancer. *Osaka City Med J* **49**, 77-83 (2003).
132 10. Gröhn, P. et al. Oral carmofur in advanced gastrointestinal cancer. *Am J Clin*
133 *Oncol* **13**, 477-9 (1990).
134 11. Nishio, S. et al. Study on effectiveness of carmofur (Mifurol) in urogenital
135 carcinoma, especially bladder cancer, as a post-operative adjuvant
136 chemotherapeutic agent. *Hinyokika Kiyo* **33**, 295-303 (1987).
137 12. Ooi, A. et al. Plasma, intestine and tumor levels of 5-fluorouracil in mice
138 bearing L1210 ascites tumor following oral administration of 5-fluorouracil,
139 UFT (mixed compound of tegafur and uracil), carmofur and 5'-deoxy-5-
140 fluorouridine. *Biol Pharm Bull* **24**, 1329-31 (2001).
141 13. Sato, S., Ueyama, T., Fukui, H., Miyazaki, K. & Kuwano, M. Anti-tumor effects
142 of carmofur on human 5-FU resistant cells. *Gan To Kagaku Ryoho* **26**, 1613-6

143   (1999).

144 14. Nguyen, H.S., Awad, A.J., Shabani, S. & Doan, N. Molecular Targeting of Acid
145   Ceramidase in Glioblastoma: A Review of Its Role, Potential Treatment, and
146   Challenges. *Pharmaceutics* **10**(2018).

147 15. Dementiev, A. et al. Molecular Mechanism of Inhibition of Acid Ceramidase by
148   Carmofur. *J Med Chem* **62**, 987-992 (2019).

149 16. Xue, X. et al. Structures of two coronavirus main proteases: implications for
150   substrate binding and antiviral drug design. *J Virol* **82**, 2515-27 (2008).

151 17. Ren, Z. et al. The newly emerged SARS-like coronavirus HCoV-EMC also has
152   an "Achilles' heel": current effective inhibitor targeting a 3C-like protease.
153   *Protein Cell* **4**, 248-50 (2013).

154 18. Wang, F., Chen, C., Tan, W., Yang, K. & Yang, H. Structure of Main Protease
155   from Human Coronavirus NL63: Insights for Wide Spectrum Anti-Coronavirus
156   Drug Design. *Sci Rep* **6**, 22677 (2016).

157 19. Yang, H. et al. Design of wide-spectrum inhibitors targeting coronavirus main
158   proteases. *PLoS Biol* **3**, e324 (2005).

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

161

162

163
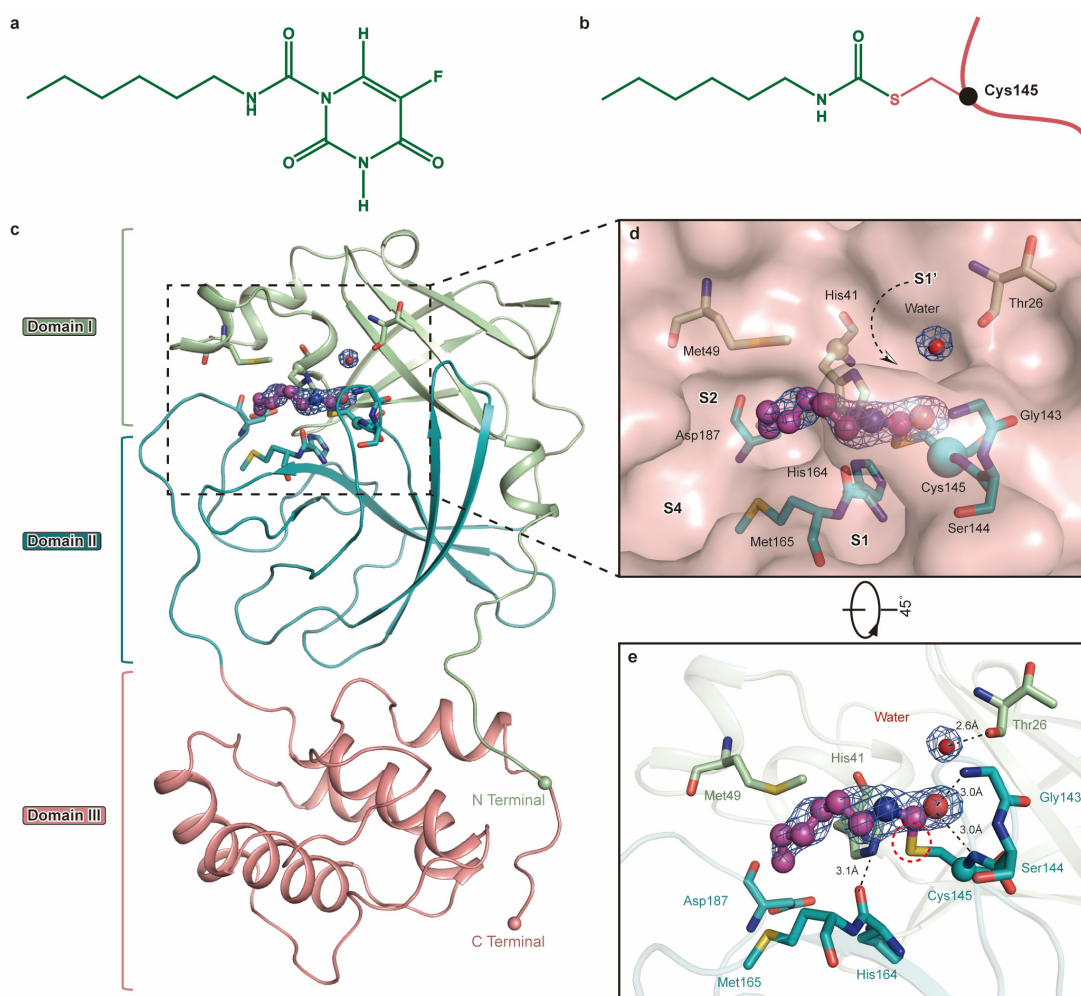
Fig. 1 | SARS-CoV-2 M^pro in complex with Carmofur. a, The chemical structure of
Carmofur. **b,** The binding mode of Carmofur to SARS-CoV-2 M^pro. The red curve
represents SARS-CoV-2 M^pro polypeptide with the sidechain of Cys145 protruding. **c,**
The structure of a single protomer. The three domains are shown in three different
colors. The catalytic center is located within the dashed square. **d,** Zoom in of the
catalytic center. The residues that participate in Carmofur binding are shown as stick
models. Carmofur is show as a ball-and-stick model with the carbons in magenta. Water
is presented as a red sphere. **e,** A rotated view of the binding site, but with the surface
removed. The red dashed circle highlights the C-S covalent bond.

173

174
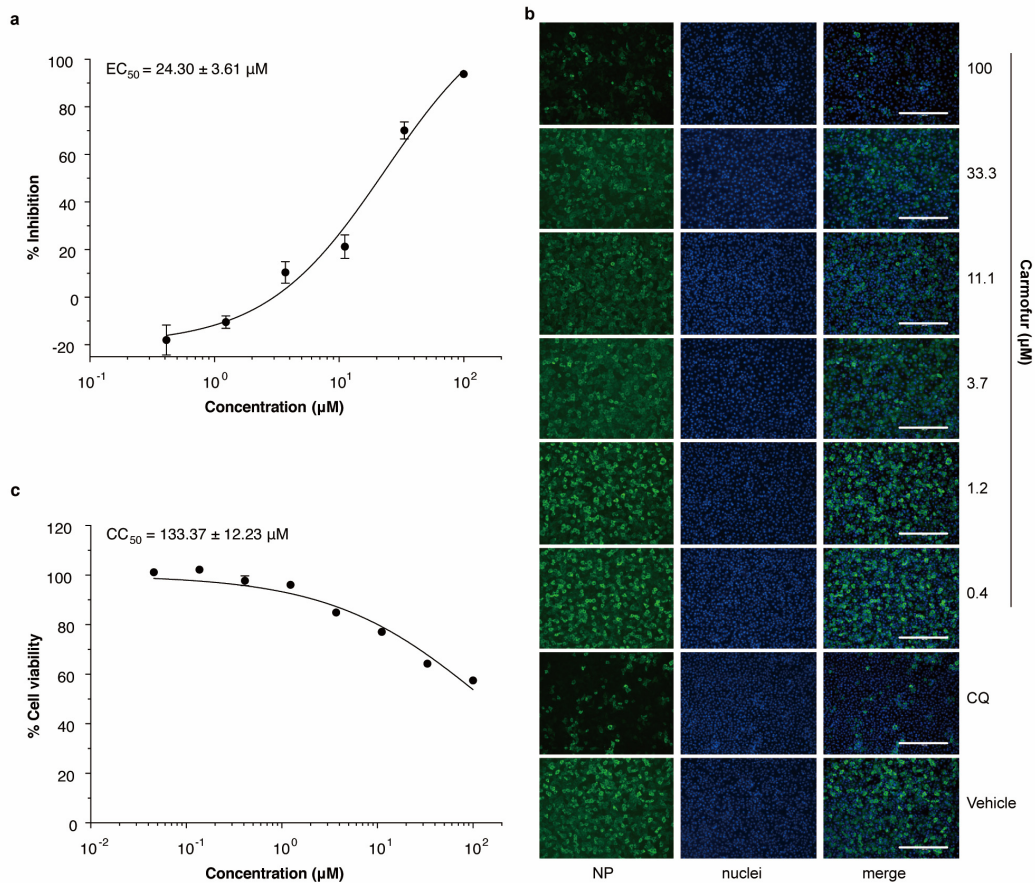
Fig. 2 | Inhibition of SARS-CoV-2 by Carmofur in Vero E6 cells. Vero E6 cells infected with SARS-CoV-2 at a MOI of 0.05 were treated with different concentrations of Carmofur. **a,** qRT-PCR assays were performed to measure viral copy number from cellular supernatant. The Y-axis of the graph indicates percentage inhibition of virus relative to DMSO (vehicle) treated sample. Data was shown as mean ± s.e.m., n = 6 biological replicates. **b,** Immunofluorescence for intracellular NP. At 24 hours post infection, cells were fixed, and intracellular NP levels were monitored by immunofluorescence. Chloroquine (CQ, 10 μM) was used as a positive control. Results are shown as representative of three biological replicates. Bars: 400 μm. **c,** Cell viability was measured using a Cell Counting Kit-8 (a commonly used kit for quantitation of viable cell number to measure proliferation and cytotoxicity) assay. The Y-axis represents percentage of cell viability relative to DMSO (vehicle) treated sample. Data was shown as mean ± s.e.m., n = 3 biological replicates. Data for graphs in **a** and **c** are available as source data.

Methods

Cloning, protein expression and purification of SARS-CoV-2 M^{pro}. The cell cultures were grown and the protein expressed according to a previous report[7]. The cell pellets were resuspended in lysis buffer (20mM Tris-HCl pH 8.0, 150 mM NaCl, 5% Glycerol), lysed by high-pressure homogenization, and then centrifuged at 25,000g for 30 min. The supernatant was loaded onto Ni-NTA affinity column (Qiagen, Germany), and washed by the lysis buffer containing 20 mM imidazole. The His-tagged M^{pro} was eluted by lysis buffer that included 300 mM imidazole. The imidazole was then removed through desalting. Human rhinovirus 3C protease was added to remove the C-terminal His tag. SARS-CoV-2 M^{pro} was further purified by ion exchange chromatography. The purified M^{pro} was transferred to 10 mM Tris-HCl pH 8.0 through desalting and stored at -80 degrees until needed.

Crystallization, data collection and structure determination. SARS-CoV-2 M^{pro} was concentrated to 5 mg/ml incubated with 0.3 mM Carmofur (Selleck, USA) for 1 hour and the complex was crystallized by hanging drop vapor diffusion method at 20 °C. The best crystals were grown using a well buffer containing 0.1 M MES pH 6.0, 5% polyethylene glycol (PEG) 6000, and 3% DMSO. The cryo-protectant solution was the reservoir but with 20% glycerol added.

X-ray data were collected on beamline BL17U1 at Shanghai Synchrotron Radiation Facility (SSRF) at 100 K and at a wavelength of 0.97918 Å using an Eiger X 16M image plate detector. Data integration and scaling were performed using the program XDS[21]. The structure was determined by molecular replacement (MR) with the PHASER[22] and Phenix 1.17.1[23] using the SARS-CoV-2 M^{pro} (PDB ID: 6LU7) as a search template. The model from MR was subsequently subjected to iterative cycles of manual model adjustment with Coot 0.8[24] and refinement was completed with Phenix REFINE[25]. The inhibitor, Carmofur, was built according to the omit map. The phasing and refinement statistics are summarized in Supplementary Table 1.

Antiviral and cytotoxicity assays for Carmofur. A clinical isolate of SARS-CoV-2 (nCoV-2019BetaCoV/Wuhan/WIV04/2019) was propagated in Vero E6 cells, and viral titer was determined as described previously[20]. Vero E6 cells were from ATCC with authentication. The authentication was performed by morphology check under microscopes and growth curve analysis. We confirm that all cells were tested as

221   mycoplasma negative. For the antiviral assay, pre-seeded Vero E6 cells (5×10^4

222   cells/well) were pre-treated with the different concentration of Carmofur for 1 h and the

223   virus was subsequently added (MOI of 0.05) to allow infection for 1 h. Next, the virus-

224   drug mixture was removed, and cells were further cultured with fresh drug containing

225   medium. At 24 h post infection, the cell supernatant was collected and vRNA in

226   supernatant was subjected to qRT-PCR analysis, while cells were fixed and subjected

227   to immunofluorescence to monitor intracellular NP level as described previously[20]. For

228   cytotoxicity assays, Vero E6 cells were suspended in growth medium in 96-well plates.

229   The next day, appropriate concentrations of Carmofur were added to the medium. After

230   24 h, the relative numbers of surviving cells were measured by the CCK8 (Beyotime,

231   China) assay in accordance with the manufacturer's instructions. All experiments were

232   performed in triplicate, and all the infection experiments were performed at biosafety

233   level-3 (BSL-3).

234   **Reporting Summary.** Further information on experimental design is available in the

235   Nature Research Reporting Summary linked to this article.

236 **Data availability**

237   Coordinates and structure factors for SARS-CoV-2 M[pro] in complex with Carmofur

238   have been deposited in Protein Data Bank (PDB) with accession number 7BUY. Source

239   data for Fig. 2 are available with the paper online.

240 **References**

241   21.   Kabsch, W. *XDS. Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).

242   22.   McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-
243   674 (2007).

244   23.   Liebschner, D. et al. Macromolecular structure determination using X-rays,
245   neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D*
246   *Struct Biol* **75**, 861-877 (2019).

247   24.   Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development
248   of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

249   25.   Afonine, P.V. et al. Towards automated crystallographic structure refinement
250   with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352-67 (2012).

251 **Acknowledgements**

Author contributions

H.Y. and Z.R. conceived the project; Z.J., Y.Zhao, H.Y. and Z.R. designed the experiments; Z.J., Y.Zhao, H.W., Y.Zhu, C.Z., X.D., J.Y. and Xiuna Yang cloned, expressed, purified and crystallized proteins; Y.Zhao, Z.J., B.Z. and T.H. collected the diffraction data; Y.Zhao, B.Z. and X.L. solved the crystal structure; Y.S. and Y.W. performed cell-based antiviral and cytotoxicity assays; Y.D. and L.Z. performed qRT-PCR and cytotoxicity assay analysis; Z.J., Y.Zhao, Y.D., Xiaobao Yang, K.Y., X.L., L.W.G., G.X., L.Z., H.Y. and Z.R. analyzed and discussed the data; Z.J., Y.Zhao, K.Y., L.W.G., L.Z., H.Y. and Z.R. wrote the manuscript.

Competing interests

The authors declare no competing interests.
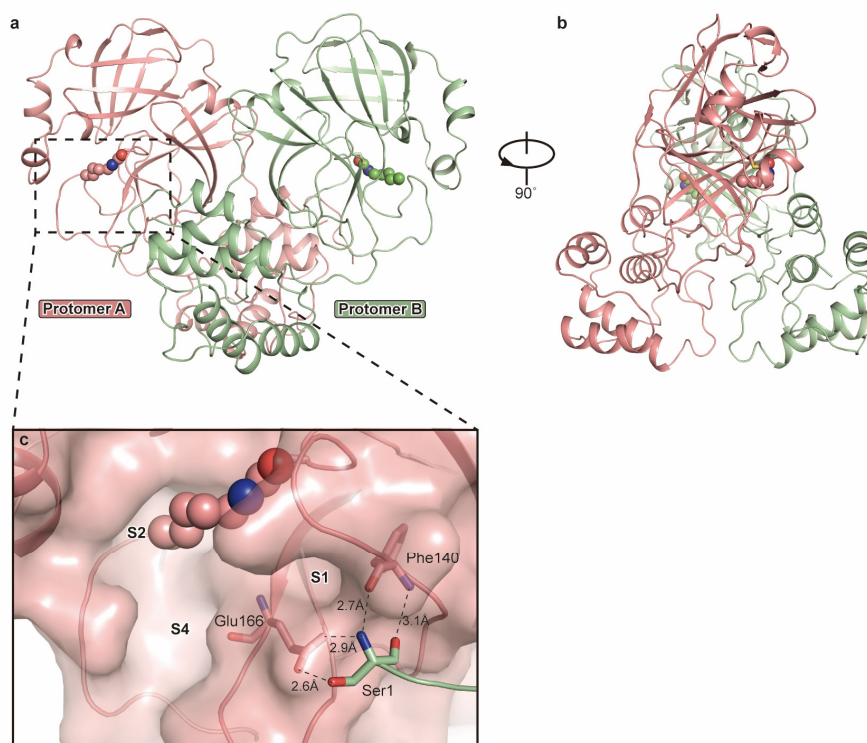
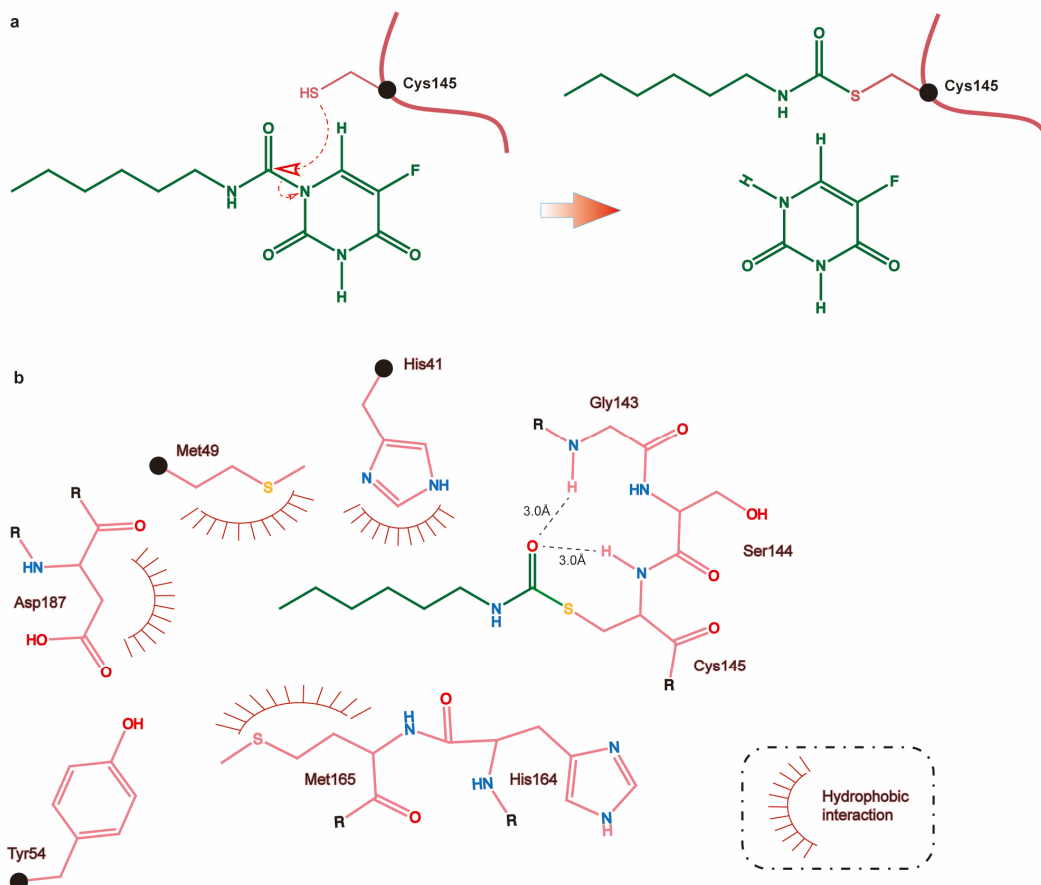**Extended Data Fig. 1 | Overall structure of SARS-CoV-2 M^pro in complex with Carmofur. a,** The overall structure of SARS-CoV-2 M^pro in complex with Carmofur. The salmon and green represent the different protomers. The Carmofur atoms are shown as solid spheres. **b,** The side view of the complex. **c,** The first serine participates in the formation of the dimer.

280

Extended Data Fig. 2 | Inhibition of M^{pro} by Carmofur. a, Putative inhibition
mechanism. Red curve represents M^{pro} polypeptide and the black sphere represent the
Cα of C145. **b,** Schematic diagram of M^{pro}-Carmofur interactions. Black spheres
represent Cα atoms.

286
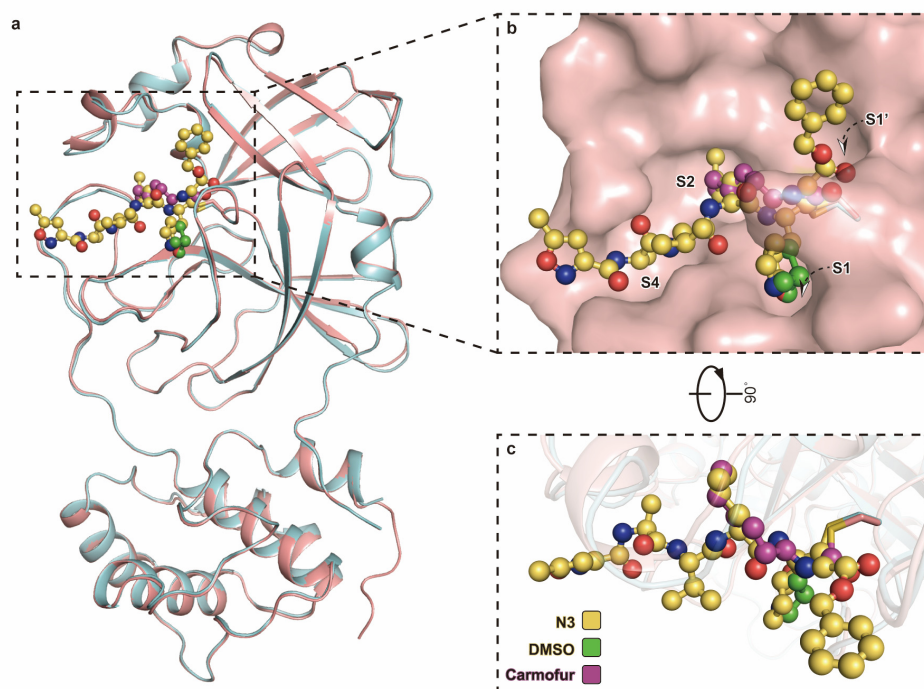
Extended Data Fig. 3 | Binding mode of Carmofur and N3 to M^{pro}. a, Overall structural comparison between the M^{pro}-Carmofur and M^{pro}-N3 complexes. The salmon cartoon represents the Carmofur bound structure and the light cyan represents the N3 bound structure. Carmofur, N3 and DMSO are represented by the purple, yellow and green balls and sticks, respectively. **b,** The binding pocket of M^{pro}. Carmofur and N3 are represent in the same way as in panel **a**. **c,** Schematic diagram of Carmofur and N3.

294 **Supplementary Table 1. Data collection and refinement statistics.**

	PDB code: 7BUY
Data Collection	
Space group	C2
Cell dimensions	
a, b, c (Å)	98.03, 81.65, 51.64
α, β, γ (°)	90, 114.88, 90
Resolution (Å)	25.64 - 1.6 (1.66 - 1.60)[a]
R_{merge}	0.03824 (0.5052)
R_{meas}	0.04164 (0.5474)
R_{pim}	0.01622 (0.2089)
$I / \sigma I$	24.05 (3.60)
$CC_{1/2}$	0.999 (0.946)
Completeness (%)	99.79 (99.96)
Redundancy	3.4 (3.4)
Refinement	
Resolution (Å)	25.64-1.6
No. reflections	48563 (4843)
R_{work} / R_{free}	0.1859 / 0.2034
No. atoms	
Protein	2367
Ligand/ion	29
Water	205
B factors	
Protein	37.91
Ligand/ion	52.55
Water	46.4
R.m.s. deviations	
Bond lengths (Å)	0.016
Bond angles (°)	1.37
Methods	
Ramachandran (%)	
Favored	97.37
Allowed	2.63
Outliers	0

Data was collected from a single crystal. [a] Values in parentheses are for highest-resolution shell.

295