RESEARCH ARTICLE

CORONAVIRUS

# Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease

Wenhao Dai[1,2]*, Bing Zhang[3]*, Xia-Ming Jiang[4]*, Haixia Su[1,5]*, Jian Li[1,6], Yao Zhao[3], Xiong Xie[1,5], Zhenming Jin[3], Jingjing Peng[1,5], Fengjiang Liu[3], Chunpu Li[1,5], You Li[7], Fang Bai[3], Haofeng Wang[3], Xi Cheng[1], Xiaobo Cen[7], Shulei Hu[1], Xiuna Yang[3], Jiang Wang[1,5], Xiang Liu[8], Gengfu Xiao[4], Hualiang Jiang[1,2,3,5], Zihe Rao[3], Lei-Ke Zhang[4]†, Yechun Xu[1,5]†, Haitao Yang[3]†, Hong Liu[1,2,5,6]†

SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) is the etiological agent responsible for the global COVID-19 (coronavirus disease 2019) outbreak. The main protease of SARS-CoV-2, $M^{pro}$, is a key enzyme that plays a pivotal role in mediating viral replication and transcription. We designed and synthesized two lead compounds (11a and 11b) targeting $M^{pro}$. Both exhibited excellent inhibitory activity and potent anti–SARS-CoV-2 infection activity. The x-ray crystal structures of SARS-CoV-2 $M^{pro}$ in complex with 11a or 11b, both determined at a resolution of 1.5 angstroms, showed that the aldehyde groups of 11a and 11b are covalently bound to cysteine 145 of $M^{pro}$. Both compounds showed good pharmacokinetic properties in vivo, and 11a also exhibited low toxicity, which suggests that these compounds are promising drug candidates.

In late December 2019, a cluster of pneumonia cases caused by a novel coronavirus was reported in Wuhan, China (*1*–*3*). Genomic sequencing showed that this pathogenic coronavirus is 96% identical to a bat coronavirus and shares 79.6% sequence identity to SARS-CoV (*4*–*6*). This novel coronavirus was named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses, and the pneumonia was designated as coronavirus disease 2019 (COVID-19) by the World Health Organization (WHO) on 11 February 2020 (*7*). The epidemic spread rapidly to more than 212 countries and was announced as a global health emergency by WHO (*8*). No clinically effective vaccines or specific antiviral drugs are currently available for the prevention and treatment of COVID-19 infections. The combination of α-interferon and the anti-HIV drugs lopinavir/ritonavir (Kaletra) has been used, but the curative effect remains very limited and there can be toxic side effects (*9*). Remdesivir, a broad-spectrum antiviral drug developed by Gilead Sciences Inc., is also being explored for the treatment of COVID-19, but more data are needed to prove its efficacy (*10*–*12*). Specific anti–SARS-CoV-2 drugs offering efficacy and safety are urgently needed.

A maximum likelihood tree based on the genomic sequence showed that the virus falls within the subgenus *Sarbecovirus* of the genus *Betacoronavirus* (*6*). Coronaviruses are enveloped, positive-sense, single-stranded RNA viruses. The genomic RNA of coronaviruses is ~30,000 nucleotides in length with a 5′-cap structure and a 3′-poly(A) tail, and contains at least six open reading frames (ORFs) (*13*, *14*). The first ORF (ORF 1a/b), about two-thirds of the genome length, directly translates two polyproteins, pp1a and pp1ab, so named because there is an a-1 frameshift between ORF1a and ORF1b. These polyproteins are processed by a main protease, $M^{pro}$ [also known as the 3C-like protease ($3CL^{pro}$)], and by one or two papain-like proteases, into 16 nonstructural proteins (NSPs). These NSPs engage in the production of subgenomic RNAs that encode four main

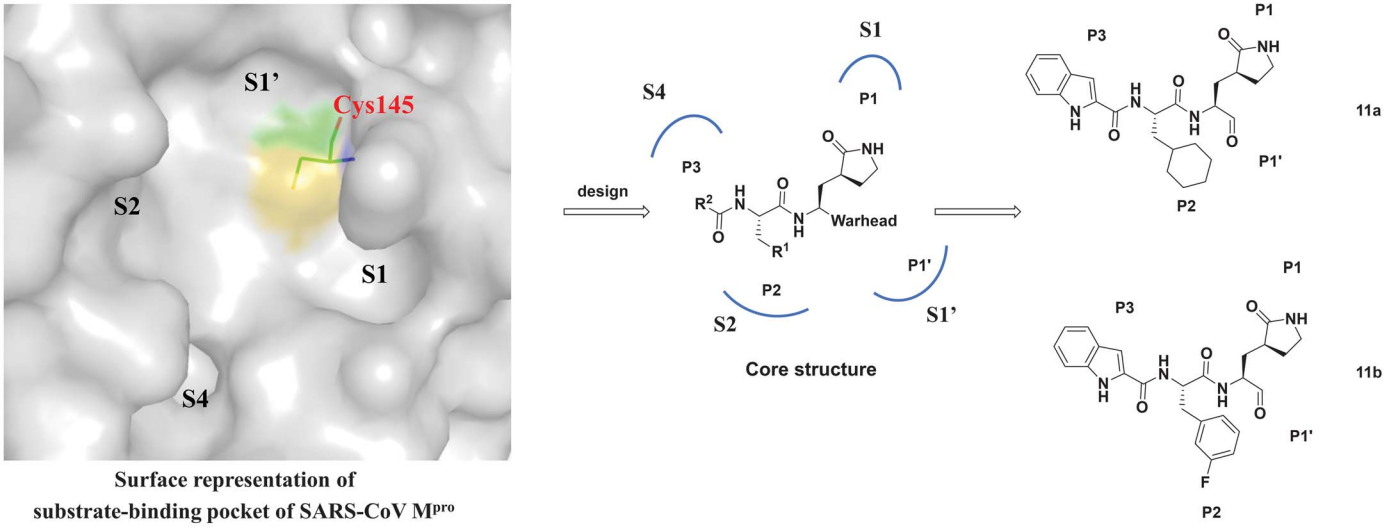


**Fig. 1. Design strategy of SARS-CoV-2 main protease ($M^{pro}$) inhibitors and the chemical structures of 11a and 11b.**

[1]State Key Laboratory of Drug Research, CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China. [2]School of Pharmacy, China Pharmaceutical University, Nanjing 210009, Jiangsu, China. [3]Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China. [4]State Key Laboratory of Virology, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan 430071, China. [5]University of Chinese Academy of Sciences, Beijing 100049, China. [6]College of Pharmacy, Nanjing University of Chinese Medicine, Qixia District, Nanjing 210023, China. [7]National Chengdu Center for Safety Evaluation of Drugs, Westchina Hospital of Sichuan University, High-Tech Development Zone, Chengdu, Sichuan 610041, China. [8]State Key Laboratory of Medicinal Chemical Biology, Frontiers Science Center for Cell Response, College of Life Sciences, College of Pharmacy, Nankai University, Tianjin 300353, China.
*These authors contributed equally to this work.
†Corresponding author. Email: hliu@simm.ac.cn (H.L.); yanght@shanghaitech.edu.cn (H.Y.); zhangleike@wh.iov.cn (L.-K.Z.); ycxu@simm.ac.cn (Y.X.)

structural proteins [envelope (E), membrane (M), spike (S), and nucleocapsid (N) proteins] and other accessory proteins (*15*, *16*). Therefore, these proteases, especially $M^{pro}$, play a vital role in the life cycle of coronaviruses.

$M^{pro}$ is a three-domain (domains I to III) cysteine protease involved in most maturation cleavage events within the precursor polyprotein (*17*–*19*). Active $M^{pro}$ is a homodimer containing two protomers. The coronavirus $M^{pro}$ features a noncanonical Cys-His dyad located in the cleft between domains I and II (*17*–*19*). $M^{pro}$ is conserved among coronaviruses, and several common features are shared among the substrates of $M^{pro}$ in different coronaviruses. The amino acids in substrates from the N terminus to the C terminus are numbered as follows: -P4-P3-P2-P1↓P1′-P2′-P3′-, with the cleavage site between P1 and P1′. In particular, a Gln is almost always required in the P1 position of the substrates. Because $M^{pro}$ has no human homolog, it is an ideal antiviral target (*20*–*22*).

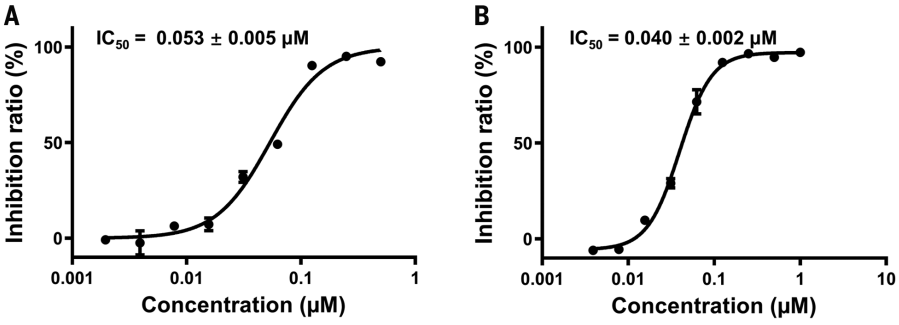


**Fig. 2. Inhibitory activity profiles of compounds 11a (A) and 11b (B) against SARS-CoV-2 $M^{pro}$.**

### Design and synthesis of 11a and 11b

The active sites of $M^{pro}$ are highly conserved among all coronavirus $M^{pro}$s and are usually composed of four sites: S1′, S1, S2, and S4 (*22*). We were able to design and synthesize inhibitors targeting SARS-CoV-2 $M^{pro}$ by analyzing the substrate-binding pocket of SARS-CoV $M^{pro}$ (PDB ID 2H2Z) (Fig. 1). The thiol of a cysteine residue in the S1′ site anchors inhibitors by a covalent linkage that is important for the inhibitors to maintain antiviral activity. In our design of new inhibitors, an aldehyde was selected as a new warhead in P1 in order to form

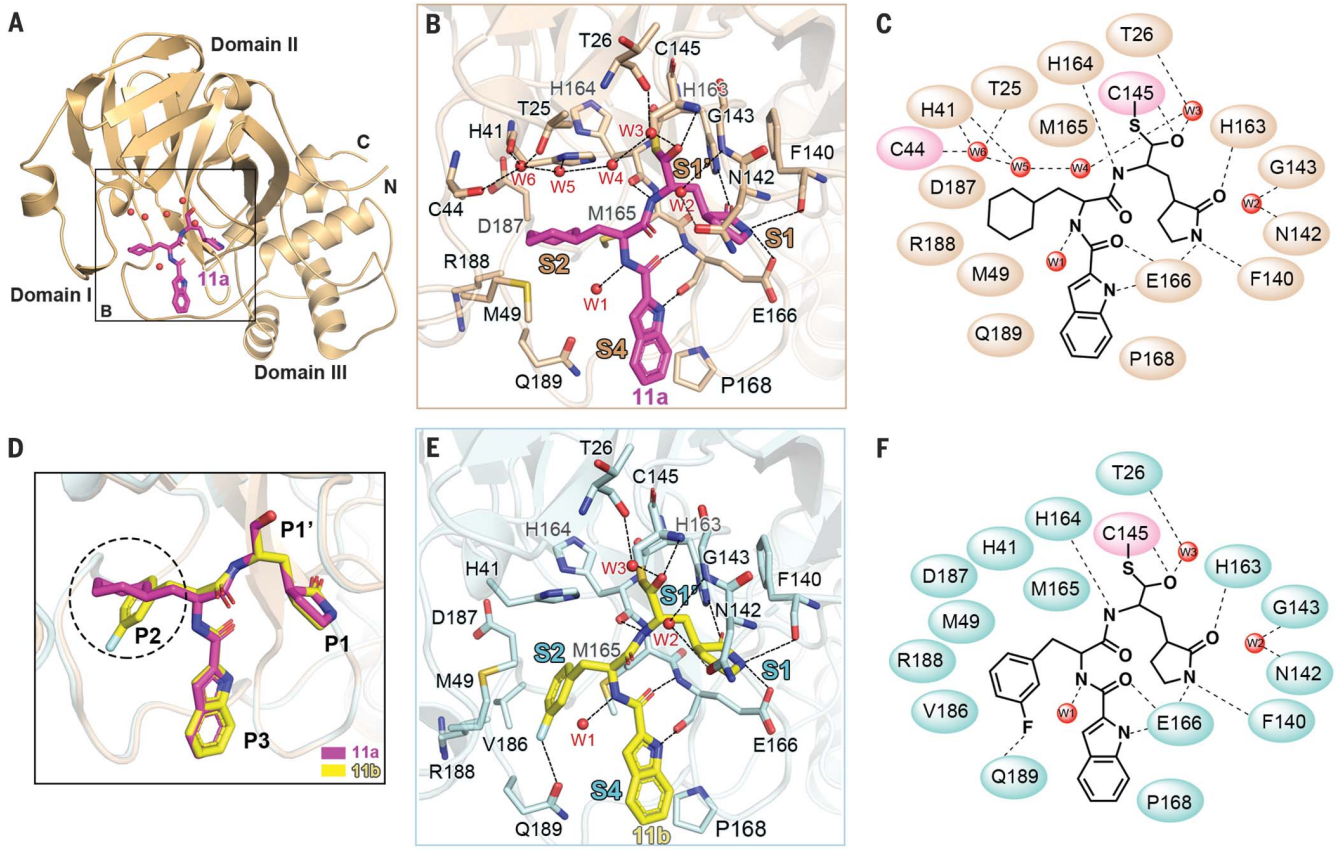


**Fig. 3. $M^{pro}$-inhibitor binding modes for 11a and 11b.** (**A**) Cartoon representation of the crystal structure of SARS-CoV-2 $M^{pro}$ in complex with **11a**. Compound **11a** is shown as magenta sticks; water molecules are shown as red spheres. (**B**) Close-up view of the **11a** binding pocket. Four subsites, S1′, S1, S2, and S4, are labeled. The residues involved in inhibitor binding are shown as wheat sticks. **11a** and water molecules are shown as magenta sticks and red spheres, respectively. Hydrogen bonds are indicated as dashed lines. (**C**) Schematic diagram of SARS-CoV-2 $M^{pro}$–**11a** interactions shown in (B). (**D**) Comparison of the binding modes of **11a** and **11b** in SARS-CoV-2 $M^{pro}$. The major difference between **11a** and **11b** is marked with a dashed circle; **11a** and **11b** are shown as magenta and yellow sticks, respectively. (**E**) Close-up view of the **11b** binding pocket. Hydrogen bonds are indicated as dashed lines. (**F**) Schematic diagram of SARS-CoV-2 $M^{pro}$–**11b** interactions shown in (E). Amino acid abbreviations: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

a covalent bond with cysteine. The reported SARS-CoV $M^{pro}$ inhibitors often have an (*S*)-γ-lactam ring that occupies the S1 site of $M^{pro}$, and this ring was expected to be a good choice in P1 (*23*). Furthermore, the S2 site of coronavirus $M^{pro}$ is usually large enough to accommodate the larger P2 fragment. To test the importance of different ring systems, we introduced a cyclohexyl or 3-fluorophenyl into P2, with the fluorine expected to enhance activity. An indole group was introduced into P3 to form new hydrogen bonds with S4 and improve drug-like properties.

The synthetic route and chemical structures of the compounds (**11a** and **11b**) are shown in scheme S1. The starting material (*N*-Boc-L-glutamic acid dimethyl ester **1**) was obtained from commercial suppliers and used without further purification to synthesize the key intermediate ***3*** according to the literature (*24*). The intermediates **6a** and **6b** were synthesized from **4** and acids **5a** and **5b**. Removal of the *t*-butoxycarbonyl group from **6a** and **6b** yielded **7a** and **7b**. Coupling **7a** and **7b** with the acid **8** yielded the esters **9a** and **9b**. The peptidomimetic aldehydes **11a** and **11b** were approached through a two-step route in which the ester derivatives **9** were first reduced with $NaBH_4$ to generate the primary alcohols **10a** and **10b**, which were subsequently oxidized into aldehydes **11a** and **11b** with Dess-Martin periodinane (DMP).

### 11a and 11b are potent inhibitors of $M^{pro}$

Recombinant SARS-CoV-2 $M^{pro}$ was expressed and purified from *Escherichia coli* (*18*, *25*). A fluorescently labeled substrate—MCA-AVLQ↓SGFR-Lys(Dnp)-Lys-$NH_2$, derived from the N-terminal autocleavage sequence from the viral protease—was designed and synthesized for the enzymatic assay.

Both **11a** and **11b** exhibited high SARS-CoV-2 $M^{pro}$ inhibition activity, which reached 100% for **11a** and 96% for **11b** at 1 μM, respectively. We used a fluorescence resonance energy transfer (FRET)–based cleavage assay to determine the median inhibitory concentration ($IC_{50}$) values. The results revealed excellent inhibitory potency, with $IC_{50}$ values of 0.053 ± 0.005 μM and 0.040 ± 0.002 μM for **11a** and **11b**, respectively (Fig. 2).

### Structures of SARS-CoV-2 $M^{pro}$ in complex with 11a and 11b

To elucidate the mechanism of inhibition of SARS-CoV-2 $M^{pro}$ by **11a**, we determined the crystal structure of this complex at 1.5 Å resolution (table S1). The crystal of $M^{pro}$-**11a** belongs to the space group *C*2, and an asymmetric unit contains only one molecule (table S1). Two molecules (designated protomer A and protomer B) associate into a homodimer around a crystallographic two-fold symmetry axis (fig. S2). The structure of each protomer contains three domains with the substrate-binding site located in the cleft between domains I and II. At the active site of SARS-CoV-2 $M^{pro}$, $Cys^{145}$ and $His^{41}$ (Cys-His) form a catalytic dyad (fig. S2).

The electron density map clearly shows compound **11a** in the substrate-binding pocket of SARS-CoV-2 $M^{pro}$ in an extended conformation (Fig. 3A and fig. S3, A and B). Details

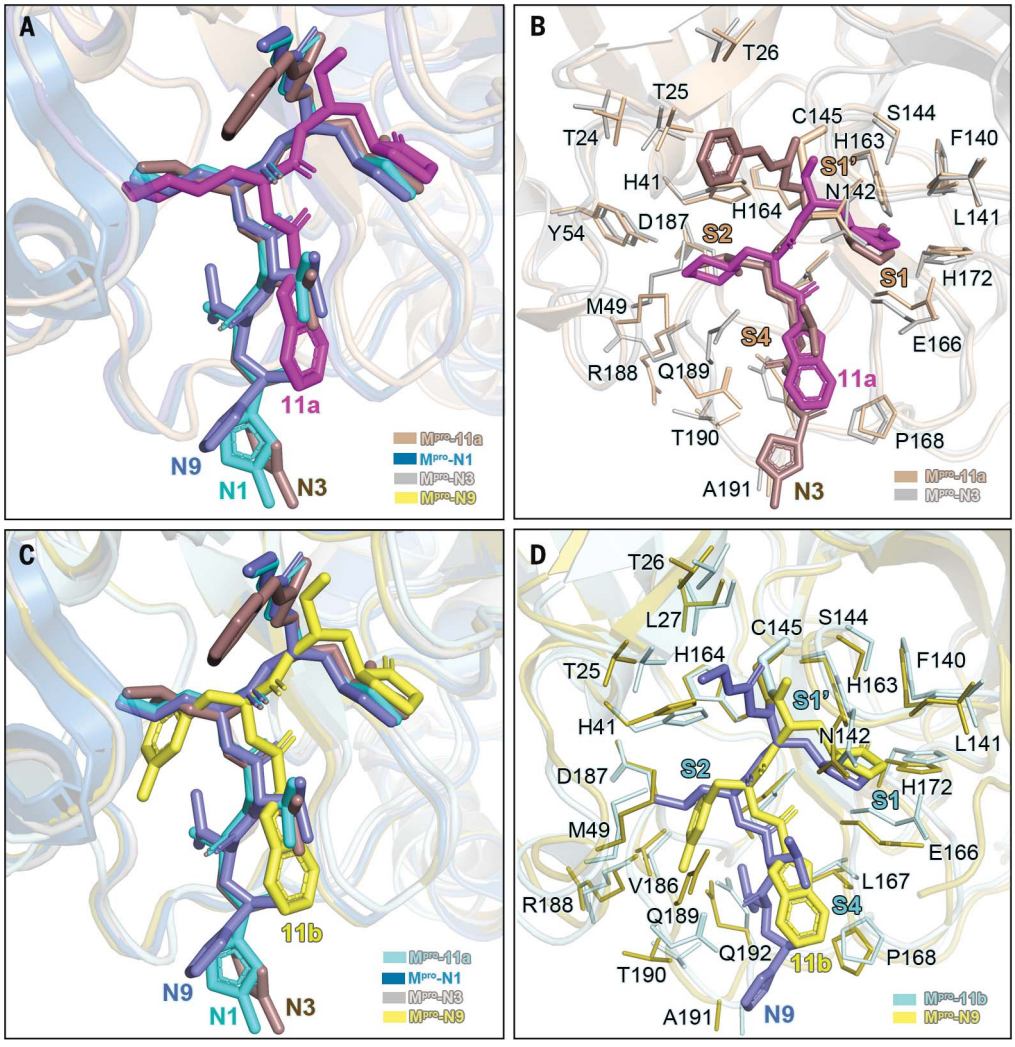


**Fig. 4. Comparison of the inhibitor binding modes in SARS-CoV $M^{pro}$ and SARS-CoV-2 $M^{pro}$.** (**A**) Comparison of binding modes of **11a** in SARS-CoV-2 $M^{pro}$ with those of **N1**, **N3**, and **N9** in SARS-CoV $M^{pro}$. SARS-CoV-2 $M^{pro}$–**11a** (wheat; PDB code 6LZE), SARS-CoV $M^{pro}$–**N1** (sky blue; PDB code 1WOF), SARS-CoV $M^{pro}$–**N3** (gray; PDB code 2AMQ), and SARS-CoV $M^{pro}$–**N9** (olive; PDB code 2AMD) are shown as cartoons. **11a**, **N1**, **N3**, and **N9** are shown in magenta, cyan, dirty violet, and salt, respectively. (**B**) Comparison of the **11a** and **N3** binding pockets. Residues in the $M^{pro}$-**11a** and $M^{pro}$-**N3** structures are colored in wheat and gray, respectively. **11a** and **N3** are shown as sticks colored in magenta and dirty violet, respectively. (**C**) Comparison of binding modes of **11b** in SARS-CoV-2 $M^{pro}$ with those of **N1**, **N3**, and **N9** in SARS-CoV $M^{pro}$. SARS-CoV-2 $M^{pro}$–**11b** (PDB code 6M0K) is shown as a pale cyan cartoon. **11b**, **N1**, **N3**, and **N9** are shown in yellow, cyan, dirty violet, and salt, respectively. (**D**) Comparison of the **11b** and **N9** binding pockets. Residues in the $M^{pro}$-**11b** and $M^{pro}$-**N9** structures are colored in pale cyan and olive, respectively. **11b** and **N9** are shown as sticks colored in yellow and salt, respectively.

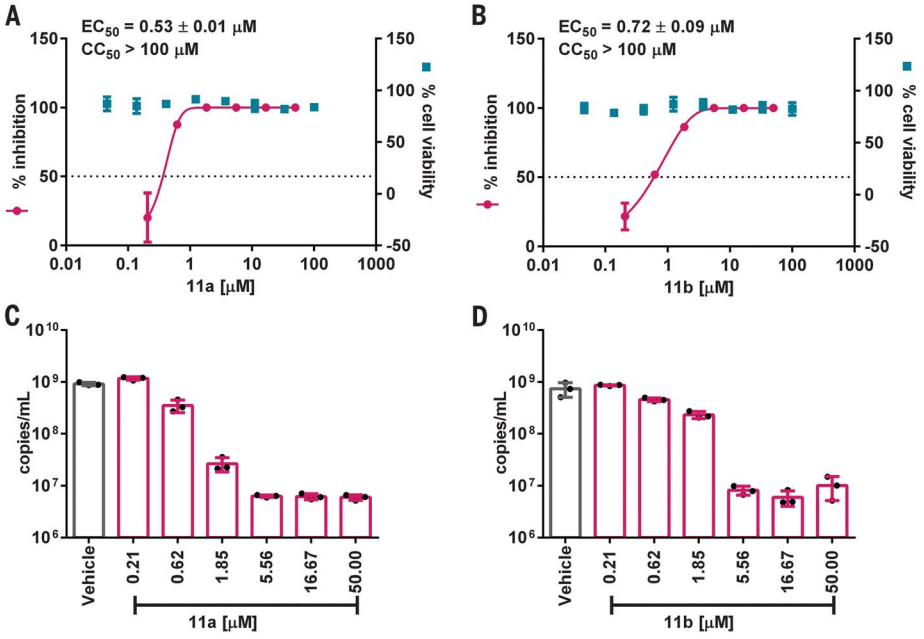


**Fig. 5. In vitro inhibition of viral main protease inhibitors against SARS-CoV-2.** (**A** and **B**) Vero E6 cells were treated with a series concentration of indicated compounds **11a** and **11b** and infected with SARS-CoV-2 at a multiplicity of infection (MOI) of 0.05. At 24 hours after infection, viral yield in the cell supernatant was quantified by plaque assay. The cytotoxicity of these compounds in Vero E6 cells was also determined by using CCK8 assays. The left and right *y* axis of each graph represents mean percent inhibition of virus yield and mean percent cell viability of the drugs, respectively. (**C** and **D**) Viral RNA copy numbers in the cell supernatants were quantified by qRT-PCR. Data are means ± SD; *n* = 3 biological replicates.

of the interaction are shown in Fig. 3, B and C. The electron density shows that the C of the aldehyde group of **11a** and the catalytic-site $Cys^{145}$ of SARS-CoV-2 $M^{pro}$ form a standard 1.8 Å C–S covalent bond. The oxygen atom of the aldehyde group also plays a crucial role in stabilizing the conformations of the inhibitor by forming a 2.9 Å hydrogen bond with the backbone of residue $Cys^{145}$ in the S1′ site. The (*S*)-γ-lactam ring of **11a** at P1 fits well into the S1 site. The oxygen of the (*S*)-γ-lactam group forms a 2.7 Å hydrogen bond with the side chain of $His^{163}$. The main chain of $Phe^{140}$ and side chain of $Glu^{166}$ also participate in stabilizing the (*S*)-γ-lactam ring by forming 3.2 Å and 3.0 Å hydrogen bonds with its NH group, respectively. In addition, the amide bonds on the chain of **11a** form hydrogen bonds with the main chains of $His^{164}$ (3.2 Å) and $Glu^{166}$ (2.8 Å), respectively. The cyclohexyl moiety of **11a** at P2 deeply inserts into the S2 site, stacking with the imidazole ring of $His^{41}$. The cyclohexyl group is also surrounded by the side chains of $Met^{49}$, $Tyr^{54}$, $Met^{165}$, $Asp^{187}$, and $Arg^{188}$, producing extensive hydrophobic interactions. The indole group of **11a** at P3 is exposed to solvent (S4 site) and is stabilized by $Glu^{166}$ through a 2.6 Å hydrogen bond. The side chains of residues $Pro^{168}$ and $Gln^{189}$ interact with the indole group of **11a** through hydrophobic interactions. Multiple water molecules (named W1 to W6) play an important role in binding **11a**. W1 interacts with the amide bonds of **11a** through a 2.9 Å hydrogen bond, whereas W2 to W6 contribute to stabilizing **11a** in the binding pocket by forming a number of hydrogen bonds with the aldehyde group of **11a** and the residues of $Asn^{142}$, $Gly^{143}$, $Thr^{26}$, $Thr^{25}$, $His^{41}$, and $Cys^{44}$.

The crystal structure of SARS-CoV-2 $M^{pro}$ in complex with **11b** is very similar to that of the **11a** complex and shows a similar inhibitor-binding mode (Fig. 3D and figs. S3, C and D, and S4A). The difference in binding mode is most probably due to the 3-fluorophenyl group of **11b** at P2. Relative to the cyclohexyl group in **11a**, the 3-fluorophenyl group undergoes a downward rotation (Fig. 3D). The side chains of residues $His^{41}$, $Met^{49}$, $Met^{165}$, $Val^{186}$, $Asp^{187}$, and $Arg^{188}$ interact with this aryl group through hydrophobic interactions, and the side chain of $Gln^{189}$ stabilizes the 3-fluorophenyl group with an additional 3.0 Å hydrogen bond (Fig. 3, E and F). In short, these two crystal structures reveal a similar inhibitory mechanism in which both compounds occupy the substrate-binding pocket and block the enzyme activity of SARS-CoV-2 $M^{pro}$.

Compounds **N1**, **N3**, and **N9** are wide-spectrum inhibitors targeting coronavirus $M^{pro}$. Compared with the binding modes of **N1**, **N3**, and **N9** in SARS-CoV $M^{pro}$ complex structures reported previously, the binding modes of **11a** and **11b** in SARS-CoV-2 $M^{pro}$ complex structures are similar and the differences among these overall structures are small (Fig. 4 and fig. S4, B to F) (*22*). The differences mainly lie in the interactions at S1′, S2, and S4 subsites, possibly because the sizes of functional groups vary at the corresponding P1′, P2, and P4 sites in the inhibitors (Fig. 4, A and C).

### Antiviral activity of 11a and 11b

To further substantiate the enzyme inhibition results, we evaluated the ability of these compounds to inhibit SARS-CoV-2 in vitro (Fig. 5 and fig. S5). As shown in Fig. 5, compounds **11a** and **11b** exhibited good anti–SARS-CoV-2 infection activity in cell culture, with half-maximal effective concentration ($EC_{50}$) values of 0.53 ± 0.01 μM and 0.72 ± 0.09 μM, respectively, by plaque assay. Neither compound caused cytotoxicity, with half cytotoxic concentration ($CC_{50}$) values of >100 μM, yielding selectivity indices for **11a** and **11b** of >189 and >139, respectively. We also used immunofluorescence and quantitative real-time polymerase chain reaction (qRT-PCR) to monitor the antiviral activity of **11a** and **11b**. The results showed that **11a** and **11b** exhibit a good antiviral effect on SARS-CoV-2 (Fig. 5 and fig. S5).

### Pharmacokinetic and toxicity studies

To explore the further druggability of the compounds **11a** and **11b**, we evaluated both compounds for their pharmacokinetic properties. As shown in table S2, compound **11a** given to mice intraperitoneally (5 mg/kg) and intravenously (5 mg/kg) displayed a half-life ($T_{1/2}$) of 4.27 hours and 4.41 hours, respectively, and we observed a high maximal concentration ($C_{max}$ = 2394 ng/ml) and a good bioavailability of 87.8% when the compound **11a** was given intraperitoneally. Metabolic stability of **11a** in mice was also good (clearance = 17.4 ml $min^{-1}$ $kg^{-1}$). When administered intraperitoneally (20 mg/kg), subcutaneously (5 mg/kg), and intravenously (5 mg/kg), compound **11b** also showed good pharmacokinetic properties (its bioavailability exceeded 80% when given both intraperitoneally and subcutaneously, and it displayed a longer $T_{1/2}$ of 5.21 hours when given intraperitoneally). Considering the danger of COVID-19, we selected intravenous drip administration for further study because of its comparatively high area under the curve (AUC) value and rapid effect. Relative to **11a** administrated intravenously in CD-1 mice, **11b** displayed a shorter $T_{1/2}$ (1.65 hours) and a faster clearance rate (clearance = 20.6 ml $min^{-1}$ $kg^{-1}$). Compound **11a** was selected for further investigation with intravenous drip dosing in Sprague-Dawley (SD) rats and beagle dogs. The results showed (table S3) that **11a** exhibited long $T_{1/2}$ values (SD rat, 7.6 hours; beagle dog, 5.5 hours), low clearance rates (rat, 4.01 ml $min^{-1}$ $kg^{-1}$; dog,

5.8 ml $min^{-1}$ $kg^{-1}$), and high AUC values (rat, 41,500 hours·ng/ml; dog, 14,900 hours·ng/ml). The above pharmacokinetic results indicate that compound **11a** warrants further study.

An in vivo toxicity study (table S4) of **11a** was carried out on SD rats and beagle dogs. The acute toxicity of **11a** was measured in SD rats. No SD rats died after receiving 40 mg/kg by intravenous drip administration. When the dosage was raised to 60 mg/kg, one of four SD rats died. The dose range toxicity of **11a** was studied for 7 days at dosing levels of 2, 6, and 18 mg/kg in SD rats and at 10 to 40 mg/kg in beagle dogs. All animals received once-daily (QD) dosing by intravenous drip, and all animals were clinically observed at least once a day. No obvious toxicity was observed in either group. These data indicate that **11a** is a good candidate for further clinical study.

## REFERENCES AND NOTES


1. N. Zhu *et al*., *N. Engl. J. Med.* **382**, 727–733 (2020).
2. Q. Li *et al*., *N. Engl. J. Med.* **382**, 1199–1207 (2020).
3. J. F. W. Chan *et al*., *Lancet* **395**, 514–523 (2020).
4. P. Zhou *et al*., *Nature* **579**, 270–273 (2020).
5. F. Wu *et al*., *Nature* **579**, 265–269 (2020).
6. R. Lu *et al*., *Lancet* **395**, 565–574 (2020).
7. A. E. Gorbalenya *et al*., bioRxiv 2020.02.07.937862 [preprint]. 11 February 2020.
8. World Health Organization, "WHO Director-General's opening remarks at the media briefing on COVID-19-11 March 2020" (2020); www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.
9. B. Cao *et al*., *N. Engl. J. Med.* NEJMoa2001282 (2020).
10. M. L. Holshue *et al*., *N. Engl. J. Med.* **382**, 929–936 (2020).
11. M. Wang *et al*., *Cell Res.* **30**, 269–271 (2020).
12. J. Cohen, *Science* 10.1126/science.abb0659 (27 January 2020).
13. Y. Chen, Q. Liu, D. Guo, *J. Med. Virol.* **92**, 418–423 (2020).
14. S. Hussain *et al*., *J. Virol.* **79**, 5288–5295 (2005).
15. R. Ramajayam, K.-P. Tan, P.-H. Liang, *Biochem. Soc. Trans.* **39**, 1371–1375 (2011).
16. Z. Ren *et al*., *Protein Cell* **4**, 248–250 (2013).
17. K. Anand *et al*., *EMBO J.* **21**, 3213–3224 (2002).
18. H. Yang *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13190–13195 (2003).
19. K. Anand, J. Ziebuhr, P. Wadhwani, J. R. Mesters, R. Hilgenfeld, *Science* **300**, 1763–1767 (2003).
20. F. G. Hayden *et al*., *Antimicrob. Agents Chemother.* **47**, 3907–3916 (2003).
21. Y. Kim *et al*., *PLOS Pathog.* **12**, e1005531 (2016).
22. H. Yang *et al*., *PLOS Biol.* **3**, e324 (2005).
23. L. Zhang *et al*., *J. Med. Chem.* acs.jmedchem.9b01828 (2020).
24. Y. Zhai *et al*., *J. Med. Chem.* **58**, 9414–9420 (2015).
25. X. Xue *et al*., *J. Mol. Biol.* **366**, 965–975 (2007).


## ACKNOWLEDGMENTS

We thank J. Halpert and LetPub (www.letpub.com) for linguistic assistance during the preparation of this manuscript, and the staff from beamlines BL17U1, BL18U1, and BL19U1 at Shanghai Synchrotron Radiation Facility for assistance during data collection. **Funding:** Supported by National Natural Science Foundation of China grants 21632008, 21672231, 21877118, 31970165, 91953000, and 81620108027; Strategic Priority Research Program of the Chinese Academy of Sciences grants XDA12040107 and XDA12040201; Chinese Academy of Engineering and Ma Yun Foundation grant 2020-CMKYGG-05; Science and Technology Commission of Shanghai Municipality grants 20431900100 and 20431900200; National Key R&D Program of China grants 2017YFC0840300, 2020YFC0841400, 2020YFA0707500, and 2017YFB0202604; Department of Science and Technology of Guangxi Zhuang Autonomous Region grant 2020AB40007; and Frontier Biotechnologies Inc. **Author contributions:** H.Y. and H.L. conceived the project; Y.X., L.-K.Z., H.Y., and H.L. designed the experiments; W.D. and J.L. designed and synthesized the compounds; X.-M.J. and H.S. tested the inhibitory activities; X.X., J.P., C.L., S.H., and J.W. performed the chemical experiments and collected the data; B.Z., Y.Z., Z.J., F.L., F.B., H.W., X.C., X.L., and X.Y. collected the diffraction data and solved the crystal structure; Y.L. and X.C. performed the toxicity experiments; G.X., H.J., Z.R., L.-K.Z., Y.X., H.Y., and H.L. analyzed and discussed the data; and L.-K.Z., Y.X., H.Y., and H.L. wrote the manuscript. **Competing interests:** The Shanghai Institute of Materia Medica has applied for PCT and Chinese patents that cover **11a**, **11b**, and related peptidomimetic aldehyde compounds. **Data and materials availability:** All data are available in the main text or the supplementary materials. The PDB accession number for the coordinates of SARS-CoV-2 $M^{pro}$ in complex with **11a** is 6LZE, and that for the coordinates of SARS-CoV-2 $M^{pro}$ in complex with **11b** is 6M0K. The plasmid encoding the SARS-CoV-2 $M^{pro}$ will be freely available. Compounds **11a** and **11b** are available from H.L. under a material transfer agreement with Shanghai Institute of Materia Medica. There is currently an international effort to join forces to design better inhibitors of SARS-CoV-2 $M^{pro}$ as described in the following website: https://covid.postera.ai/covid. 

## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6497/1331/suppl/DC1
Materials and Methods
Scheme S1
Figs. S1 to S5
Tables S1 to S4
References (*26*–*29*)

18 March 2020; accepted 20 April 2020
Published online 22 April 2020
10.1126/science.abb4489

Science

# Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease

Wenhao Dai, Bing Zhang, Xia-Ming Jiang, Haixia Su, Jian Li, Yao Zhao, Xiong Xie, Zhenming Jin, Jingjing Peng, Fengjiang Liu, Chunpu Li, You Li, Fang Bai, Haofeng Wang, Xi Cheng, Xiaobo Cen, Shulei Hu, Xiuna Yang, Jiang Wang, Xiang Liu, Gengfu Xiao, Hualiang Jiang, Zihe Rao, Lei-Ke Zhang, Yechun Xu, Haitao Yang and Hong Liu

*Science* **368** (6497), 1331-1335.
DOI: 10.1126/science.abb4489originally published online April 22, 2020

**Promising antiviral protease inhibitors**

With no vaccine or proven effective drug against the virus that causes coronavirus disease 2019 (COVID-19), scientists are racing to find clinical antiviral treatments. A promising drug target is the viral main protease $M^{pro}$, which plays a key role in viral replication and transcription. Dai *et al.* designed two inhibitors, 11a and 11b, based on analyzing the structure of the $M^{pro}$ active site. Both strongly inhibited the activity of $M^{pro}$ and showed good antiviral activity in cell culture. Compound 11a had better pharmacokinetic properties and low toxicity when tested in mice and dogs, suggesting that this compound is a promising drug candidate.

*Science*, this issue p. 1331



*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.