



## LETTER

# Crystal structure of SARS-CoV-2 main protease in complex with protease inhibitor PF-07321332

Dear Editor,

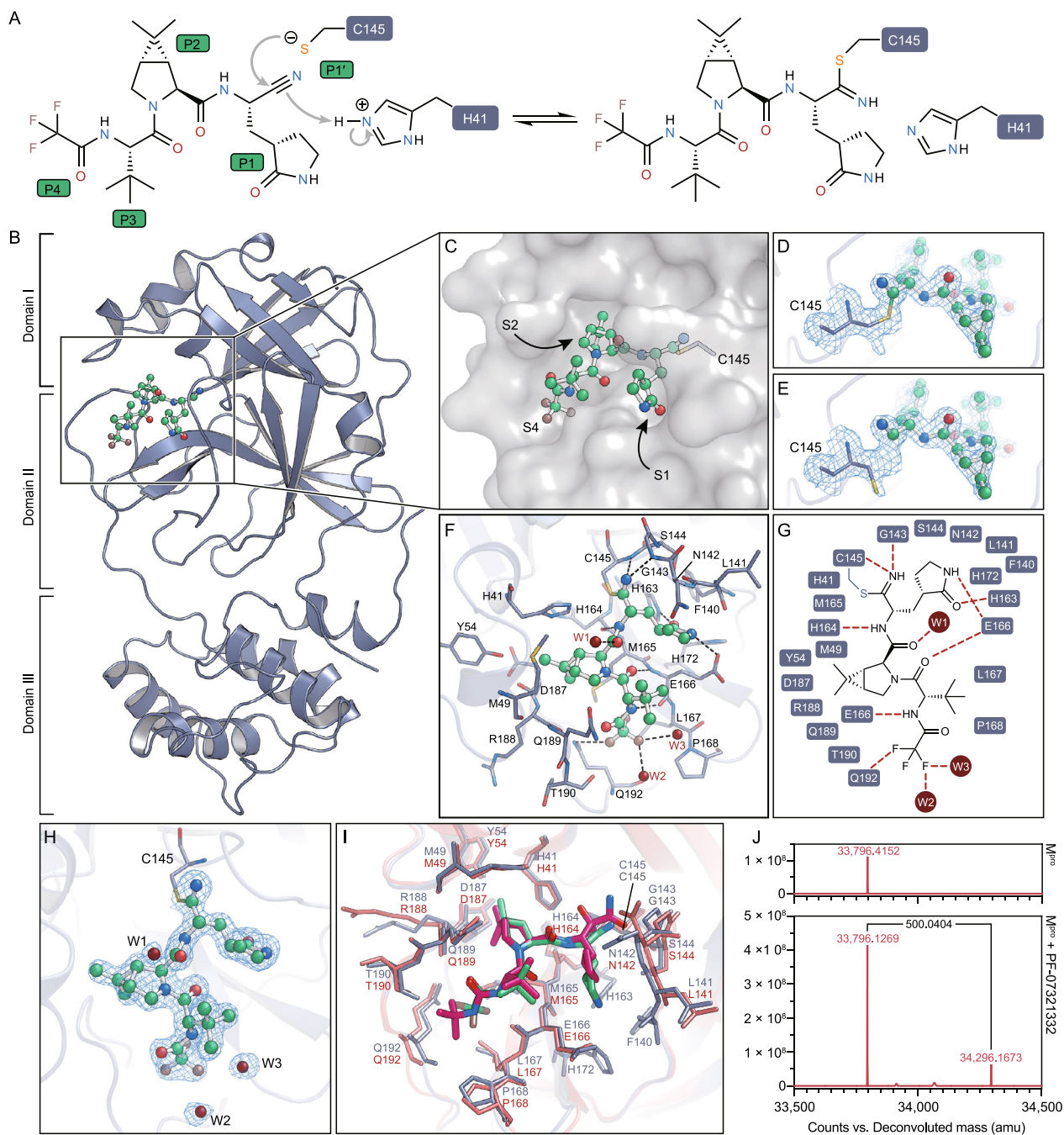
Since December 2019, the pandemic of coronavirus disease 2019 (COVID-19) has taken a heavy toll on global health, creating an urgent need to develop effective strategies for prevention and treatment. The etiological agent, known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has infected nearly 229.2 million people worldwide with more than 4.7 million deaths as of September 15, 2021. Older age and preexisting health conditions are associated with worse clinical prognosis including higher mortality rates (Zhou et al., 2020). The global race to combat this pandemic has led to rapid deployment of numerous effective vaccines against SARS-CoV-2 (Tregoning et al., 2021). However, the emergence of viral variants, including the Delta variant (B.1.617.2), compromised vaccine effectiveness with resurgence of SARS-CoV-2 infection among highly vaccinated population (Keehner et al., 2021). Therefore, development of therapeutics against the more conserved viral targets would be essential to contain the spread of COVID-19 and reduce mortality.

SARS-CoV-2 is a positive-sense, single-stranded RNA virus, and its genomic sequence shares approximately 80% identity to that of SARS-CoV, the pathogen that caused the SARS outbreak in 2003 (Lu et al., 2020). The replication of SARS-CoV-2 relies on two polyproteins pp1a and pp1ab, which are encoded by the 5' two-thirds of the viral genome (Cui et al., 2020). Both pp1a and pp1ab are co-translationally cleaved into 16 mature nonstructural proteins (nsps) by the two viral cysteine proteases, the main protease ( $M^{pro}$ , also called 3C like protease, 3CLpro) and the papain like protease (PLpro) (Jin et al., 2020; Zhao et al., 2021). SARS-CoV-2  $M^{pro}$  processes the two polyproteins through at least 11 cleavage sites. These nsps, including helicase, RNA-dependent RNA polymerase (RdRp), and methyltransferase and so on, then assemble into the replication-transcription complex to promote viral replication (Cui et al., 2020). From an evolutionary perspective, the amino acid sequence and 3-dimensional structure of  $M^{pro}$  are highly conserved throughout the subfamily *Coronavirinae*, providing a strong mechanistic basis for designing therapeutics to address the

recent emergence of SARS-CoV-2 variants with concerns for immune evasion (Wang et al., 2016; Jin et al., 2020). Examination of  $M^{pro}$  sequences across known SARS-CoV-2 variants revealed low mutation frequency (Table S1 and Fig. S1). In addition, the essential role of  $M^{pro}$  in viral life cycle and lack of its homolog in the human beings make it an attractive target for drug development against SARS-CoV-2 (Yang et al., 2005).

Using a platform consisted of structure-assisted drug design, virtual drug screening, and high-throughput screening, we previously discovered several lead compounds including ebiselen and carmofur that potently inhibit SARS-CoV-2  $M^{pro}$  (Jin et al., 2020). Given the genomic and structural similarity between SARS-CoV and SARS-CoV-2, several inhibitors rationally designed against SARS-CoV  $M^{pro}$ , such as N3 and PF-00835231, also demonstrated potent inhibitory activity against SARS-CoV-2  $M^{pro}$  (Hoffman et al., 2020; Jin et al., 2020). However, the low oral bioavailability of PF-00835231 limits its clinical potential, which has been improved using its phosphate prodrug, PF-07304814 (Boras et al., 2021). PF-07304814 is highly soluble that permits dosing with continuous infusion. Recently, Pfizer announced the development of the second-generation orally available SARS-CoV-2  $M^{pro}$  inhibitor, PF-07321332, which has been evaluated in a dose escalation phase 1 study for safety (NCT04756531) (Vandyck and Deval, 2021). The phase 2/3 clinical trials (NCT04960202 and NCT05011513) to evaluate the safety and efficacy of PF-07321332/ritonavir combination among non-hospitalized adult patients with COVID-19 are currently in progress. Co-administration of low dose ritonavir is expected to slow down the metabolic degradation of PF-07321332 in the body.

To elaborate its inhibitory mechanism, we determined the crystal structure of SARS-CoV-2  $M^{pro}$  in complex with PF-07321332 at 1.6 Å resolution (Table S2). PF-07321332 is a reversible covalent inhibitor carrying a nitrile warhead that targets SARS-CoV-2  $M^{pro}$  (Fig. 1A). The overall structure of SARS-CoV-2  $M^{pro}$  is a functional dimer, consistent with our previous findings (Yang et al., 2005; Wang et al., 2016; Jin et al., 2020). All of the residues (1–306) could be traced according to the electron density map. Each protomer



◀ **Figure 1. The structure of SARS-CoV-2 M<sup>pro</sup> in complex with PF-07321332.** (A) A putative inhibition mechanism of PF-07321332. The catalytic C145 and H41 of SARS-CoV-2 M<sup>pro</sup> are highlighted with dark blue squares. P1', P1, P2 and P4 are highlighted with green squares. (B) The overall structure of a SARS-CoV-2 M<sup>pro</sup> protomer. The three domains of M<sup>pro</sup> are labeled. The substrate-binding pocket is located within the black box. PF-07321332 is shown as ball-and-stick model with the carbon atoms in bright green, oxygen atoms in bright red, and nitrogen atoms in blue, and fluorine atom in brown. (C) The zoom-in view of the substrate-binding pocket. PF-07321332 forms a covalent bond to C145. Three substrate-binding subsites (S1, S2 and S4) are labeled. (D) The S<sub>γ</sub> atom of the catalytic C145 forms a 1.8-Å C-S covalent bond with the nitrile carbon of PF-07321332. The 2F<sub>o</sub>-F<sub>c</sub> density map contoured at 1.2σ is shown in the light blue mesh. (E) The catalytic C145 in another state where it does not form a covalent bond with PF-07321332. The 2F<sub>o</sub>-F<sub>c</sub> density map contoured at 1.2σ is shown in the light blue mesh. (F) The detailed binding model in the SARS-CoV-2 M<sup>pro</sup>-PF07321332 complex structure. The residues that participate in inhibitor binding are shown as sticks. (G) Schematic diagram of the interactions between PF-07321332 and SARS-CoV-2 M<sup>pro</sup>. (H) The polder density map of PF-07321332, which is colored as the light blue mesh and contoured at 3.0 σ. (I) The comparison of SARS-CoV-2 M<sup>pro</sup>-PF07321332 complex structure with that of SARS-CoV-2 M<sup>pro</sup>-boceprevir. (J) The molecular weights of apo SARS-CoV-2 M<sup>pro</sup> and PF-07321332 treated M<sup>pro</sup> determined from tandem mass spectrometry. The mass shift (Δm) of the protein is labeled.

contains three domains (Fig. 1B). Domain I (residues 10–99) and domain II (residues 100–184) are primarily consisted of β-barrel folds. Domain III (residues 201–303) contains five α-helices, and it is relatively independent and connected to domain II through a long loop region (residues 185–200). The substrate-binding pocket that features a non-canonical dyad of cysteine 145 and histidine 41 (C145 and H41) lies in the cleft between domains I and II (Fig. 1B and C).

The unambiguous electron density map demonstrates that the nitrile carbon of PF-07321332 is linked to the S<sub>γ</sub> atom of C145 through a standard 1.8-Å C-S covalent bond (Fig. 1D and H). This suggests that the thiol group of catalytic C145 attacks the electrophilic nitrile group at P1' site of PF-07321332, resulting in the formation of a thioimidate adduct (Fig. 1A). This result is also supported by tandem mass spectrometry analysis, which revealed the presence of a mass shift of 500 Dalton (Fig. 1J). The reversibility of the covalent bond was supported by the dual conformations of the catalytic cysteine based on the electron density map, with the thiol group of cysteine not directly bonded to nitrile group of the inhibitor in the alternative conformation (Fig. 1E). The imine nitrogen of the thioimidate moiety occupies the oxyanion hole that is stabilized by the backbone NH of G143 and C145 through two hydrogen bonds

(Fig. 1F and G). The classical (S)-γ-lactam ring at P1 position fits into the S1 subsite in a similar pattern to several other rationally designed inhibitors such as N3, 11a, and 11b (Dai et al., 2020; Jin et al., 2020). The hydrogen bond between the oxygen atom of the lactam ring and Nε2 atom of H163 stabilizes the inhibitor (Fig. 1F and G). Besides, the Oε1 atom of E166 also interacts with the NH group of the lactam ring through a hydrogen bond (Fig. 1F and G). The amide nitrogen at P1 site forms a hydrogen bond with the main-chain carbonyl oxygen of H164 and the amide oxygen forms another hydrogen bond with an ordered water molecule W1 (Fig. 1F and G).

Similar to the structure of SARS-CoV-2 M<sup>pro</sup> in complex with boceprevir (Fu et al., 2020), the S2 subsite is occupied by the rigid dimethylcyclopropylproline (DMCP) group (Fig. 1C, F, and I). This hydrophobic group is surrounded by the side chains of H41, M49, Y54, M165 and D189, and the main chains of D187 and R188, resulting in extensive hydrophobic interactions (Fig. 1F and G). The hydrophobic tert-butyl group at P3 was exposed to the solvent with limited interactions with M<sup>pro</sup> (Fig. 1C). The trifluoroacetyl group at P4 site is accommodated by the S4 sub-pocket. The amide nitrogen at P4 site forms a hydrogen bond with the main chain carbonyl oxygen of E166 (Fig. 1F and 1G). Compared with SARS-CoV-2 M<sup>pro</sup>-boceprevir complex structure, the trifluoromethyl group of PF-07321332 forms additional hydrogen bonds by interacting with the Nε2 atom Q192 and two ordered water molecules W2 and W3, which provides stronger interactions to anchor the P4 site (Fig. 1F–H). Therefore, in addition to the S-C covalent bond connecting the nitrile carbon with the catalytic C145, the inhibitor PF-07321332 is further stabilized through a network of hydrogen bonds and hydrophobic interactions, which enhance its binding to the active site of SARS-CoV-2 M<sup>pro</sup>.

Protease inhibitors have been proven to be an effective pharmaceutical strategy to combat both hepatitis C virus (HCV) and human immunodeficiency virus (HIV) infections in the clinic. Peptidomimetic inhibitors, such as boceprevir and telaprevir, carry a warhead of α-ketoamide and form a reversible covalent bond with the catalytic serine residue at the active site of HCV NS3/4A protease (Manns and von Hahn, 2013). Such mechanism-based inactivation of viral protease through forming a covalent bond enhances the selectivity to the viral target. We previously demonstrated that N3, a peptidomimetic compound with a Michael acceptor that forms an irreversible covalent bond with the catalytic cysteine, exhibited broad-spectrum inhibition of the M<sup>pro</sup> across both human and animal CoVs, including the SARS-CoV-2 (Yang et al., 2005; Wang et al., 2016; Jin et al., 2020). The complex structure of SARS-CoV-2 M<sup>pro</sup> with PF-07321332 demonstrates the efficacy of the warhead to form a covalent bond with the cysteine residue at the active site. The lactam ring, DMCP group, and trifluoroacetyl group enhance the binding of the inhibitor at the S1, S2, and S4 subsites of the substrate-binding pocket. Similar to PF-00835231 and N3 (Yang et al., 2005; Hoffman et al., 2020), PF-07321332 may exhibit wide spectrum

inhibition against M<sup>pro</sup> beyond SARS-CoV-2, highlighting the potential to be used for other human CoVs.

In conclusion, the 1.6-Å crystal structure of SARS-CoV-2 M<sup>pro</sup> in complex with the protease inhibitor PF-07321332 reveals critical insight for pharmaceutical development. With the emergence of SARS-CoV-2 variants with immune evasion, therapeutics targeting a conserved but essential viral target among the variants may synergize the COVID-19 vaccines. The improved oral bioavailability would enhance the clinical utility of PF-07321332 as a therapeutic to prevent progression among early-stage non-hospitalized patients or as a prophylaxis in the pre- and post-exposure setting. Further optimization of such covalent peptidomimetic inhibitors would generate potent drugs to contain the current COVID-19 pandemic and increase public health preparedness for potential future pandemic.

## FOOTNOTES

We are grateful to the staff at the BL19U1 beamline of the Shanghai Synchrotron Radiation Facility (SSRF), where the diffraction data were collected. We also appreciate the advice from Dr. Xiang Liu of Nankai University and Dr. Lijie Wu of ShanghaiTech University. This work was supported by grants from National Key R&D Program of China (Grant No. 2017YFC0840300 to Z.R.), National Key R&D Program of China (Grant No. 2020YFA0707500 to H.Y.), Science and Technology Commission of Shanghai Municipality (Grant Nos. 20431900200 and 20XD1422900), and Department of Science and Technology of Guangxi Zhuang Autonomous Region (Grant No. 560 2020AB40007).

H.Y., D.M., and Z.R. conceived the project; Y.Z., D.M., Z.R., and H.Y. designed the experiments; Y.Z., Q.Z., Yan Z., and L.F. cloned, expressed, and purified the proteins; Q.Z., Y.Z., and M.S. performed the crystallization experiment; Y.Z., Q.Z., and J.Z. collected the diffraction data; Y.Z. solved the crystal structure; C.F. synthesized PF-07321332; C.P. performed intact protein and MS/MS analyses; Y. Z., Q.Z., C.F., R.Z., X.Z., Y.D., H.W., X.Y., L.Z., K.Y., D.M., Z.R., and H.Y. analyzed and discussed the data; Y.Z., K.Y., D.M., Z.R., and H. Y. wrote the manuscript.

Coordinates and structure factors of the SARS-CoV-2 M<sup>pro</sup>-PF-07321332 complex have been deposited in Protein Data Bank (PDB) with the accession number 7VH8.

The authors declare no competing interests.

Yao Zhao<sup>1</sup> , Chao Fang<sup>2</sup> , Qi Zhang<sup>1,3,4</sup> ,  
Ruxue Zhang<sup>5</sup> , Xiangbo Zhao<sup>2</sup>, Yinkai Duan<sup>1</sup> ,  
Haofeng Wang<sup>1</sup> , Yan Zhu<sup>1</sup> , Lu Feng<sup>1</sup> , Jinyi Zhao<sup>1</sup> ,  
Maolin Shao<sup>1</sup> , Xiuna Yang<sup>1</sup> , Leike Zhang<sup>5</sup> ,  
Chao Peng<sup>6</sup> , Kailin Yang<sup>7</sup> , Dawei Ma<sup>2,✉</sup> ,  
Zihe Rao<sup>1,8,9,✉</sup> , Haitao Yang<sup>1,✉</sup> 

<sup>1</sup> Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China

- <sup>2</sup> State Key Laboratory of Bioorganic & Natural Products Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200032, China
- <sup>3</sup> CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai 200031, China
- <sup>4</sup> University of Chinese Academy of Sciences, Beijing 100049, China
- <sup>5</sup> State Key Laboratory of Virology, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan 430071, China
- <sup>6</sup> National Facility for Protein Science in Shanghai, Zhangjiang Lab, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China
- <sup>7</sup> Taussig Cancer Center, Cleveland Clinic, Cleveland, OH 44195, USA
- <sup>8</sup> Laboratory of Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100091, China
- <sup>9</sup> State Key Laboratory of Medicinal Chemical Biology, Frontiers Science Center for Cell Response, College of Life Sciences, Nankai University, and Tianjin Key Laboratory of Protein Sciences, Tianjin 300071, China
- ✉ Correspondence: madw@sioc.ac.cn (D. Ma), raozh@mail.tsinghua.edu.cn (Z. Rao), yanght@shanghaitech.edu.cn (H. Yang)

Accepted September 26, 2021

## OPEN ACCESS

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

## REFERENCES

- Boras B, Jones RM, Anson BJ, Arenson D, Aschenbrenner L, Bakowski MA, Beutler N, Binder J, Chen E, Eng H et al (2021) Discovery of a novel inhibitor of coronavirus 3CL protease for the potential treatment of COVID-19. *BbioRxiv* 1:2020–2029
- Cui W, Yang K, Yang H (2020) Recent progress in the drug development targeting SARS-CoV-2 main protease as treatment for COVID-19. *Front Mol Biosci* 7:616341



- Dai W, Zhang B, Jiang XM, Su H, Li J, Zhao Y, Xie X, Jin Z, Peng J, Liu F et al (2020) Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* 368:1331–1335
- Fu L, Ye F, Feng Y, Yu F, Wang Q, Wu Y, Zhao C, Sun H, Huang B, Niu P et al (2020) Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat Commun* 11:4417
- Hoffman RL, Kania RS, Brothers MA, Davies JF, Ferre RA, Gajiwala KS, He M, Hogan RJ, Kozminski K, Li LY et al (2020) Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19. *J Med Chem* 63:12725–12747
- Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C et al (2020) Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582:289–293
- Keehner J, Horton LE, Binkin NJ, Laurent LC, Pride D, Longhurst CA, Abeles SR, Torriani FJ (2021) Resurgence of SARS-CoV-2 infection in a highly vaccinated health system workforce. *N Engl J Med* 384:1774–1775
- Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N et al (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 395:565–574
- Manns MP, von Hahn T (2013) Novel therapies for hepatitis C - one pill fits all? *Nat Rev Drug Discov* 12:595–610
- Tregoning JS, Flight KE, Higham SL, Wang Z, Pierce BF (2021) Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. *Nat Rev Immunol* 21:626–636
- Vandyck K, Deval J (2021) Considerations for the discovery and development of 3-chymotrypsin-like cysteine protease inhibitors targeting SARS-CoV-2 infection. *Curr Opin Virol* 49:36–40
- Wang F, Chen C, Tan W, Yang K, Yang H (2016) Structure of main protease from human coronavirus NL63: insights for wide spectrum anti-coronavirus drug design. *Sci Rep* 6:22677
- Yang H, Xie W, Xue X, Yang K, Ma J, Liang W, Zhao Q, Zhou Z, Pei D, Ziebuhr J et al (2005) Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol* 3:e324
- Zhao Y, Du X, Duan Y, Pan X, Sun Y, You T, Han L, Jin Z, Shang W, Yu J et al (2021) High-throughput screening identifies established drugs as SARS-CoV-2 PLpro inhibitors. *Protein Cell*. <https://doi.org/10.1007/s13238-021-00836-9>
- Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X et al (2020) Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395:1054–1062

---

Yao Zhao, Chao Fang and Qi Zhang have contributed equally to this work.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s13238-021-00883-2>.