

## Structure of SARS-CoV-2 main protease in the apo state

Xuelan Zhou<sup>1†</sup>, Fanglin Zhong<sup>2,3†</sup>, Cheng Lin<sup>4,5</sup>, Xiaohui Hu<sup>1</sup>, Yan Zhang<sup>6</sup>, Bing Xiong<sup>7</sup>,  
Xiushan Yin<sup>8,9</sup>, Jinheng Fu<sup>10</sup>, Wei He<sup>2</sup>, Jingjing Duan<sup>11</sup>, Yang Fu<sup>12</sup>, Huan Zhou<sup>13</sup>,  
Peter J. McCormick<sup>14</sup>, Qisheng Wang<sup>13\*</sup>, Jian Li<sup>2,3\*</sup> & Jin Zhang<sup>1\*</sup>

<sup>1</sup>School of Basic Medical Sciences, Nanchang University, Nanchang 330031, China;

<sup>2</sup>College of Pharmaceutical Sciences, Gannan Medical University, Ganzhou 341000, China;

<sup>3</sup>Laboratory of Prevention and Treatment of Cardiovascular and Cerebrovascular Diseases, Ministry of Education, Gannan Medical University, Ganzhou 341000, China;

<sup>4</sup>Shenzhen Crystal Biopharmaceutical Co., Ltd, Shenzhen 518118, China;

<sup>5</sup>Jiangxi Jmery Biopharmaceutical Co., Ltd, Ganzhou 341000, China;

<sup>6</sup>The Second Affiliated Hospital of Nanchang University, Nanchang 330031, China;

<sup>7</sup>Department of Medicinal Chemistry, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China;

<sup>8</sup>Applied Biology Laboratory, Shenyang University of Chemical Technology, Shenyang 110142, China;

<sup>9</sup>Biotech & Biomedicine Science (Jiangxi) Co. Ltd, Ganzhou 341000, China;

<sup>10</sup>Jiangxi-OAI Joint Research Institution, Nanchang University, Nanchang 330047, China;

<sup>11</sup>Human Aging Research Institute (HARI), School of Life Sciences, Nanchang University, Nanchang 330031, China;

<sup>12</sup>School of Medicine, Southern University of Science and Technology, Shenzhen 518055, China;

<sup>13</sup>Shanghai Synchrotron Radiation Facility, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201204, China;

<sup>14</sup>Centre for Endocrinology, William Harvey Research Institute, Bart's and the London School of Medicine and Dentistry, Queen Mary, University of London, London EC1M 6BQ, UK

Received June 28, 2020; accepted August 1, 2020; published online September 1, 2020

**Citation:** Zhou, X., Zhong, F., Lin, C., Hu, X., Zhang, Y., Xiong, B., Yin, X., Fu, J., He, W., Duan, J., et al. (2021). Structure of SARS-CoV-2 main protease in the apo state. *Sci China Life Sci* 64, 656–659. <https://doi.org/10.1007/s11427-020-1791-3>

Dear Editor,

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been a global pandemic that severely threatens global health with concordant economic damage. However, there is currently no clinically approved vaccines or drugs against COVID-19 (Lu et al., 2020). SARS-CoV-2 main protease ( $M^{\text{pro}}$ , also called 3C-like protease, 3CL<sup>pro</sup>) mediates the proteolytic processing of large replicase polyprotein 1a (pp1a) and pp1ab into non-structural proteins (NSPs) at

eleven conservative sites (Zumla et al., 2016). Thus,  $M^{\text{pro}}$  is of considerable interest as a drug target in the treatment of COVID-19 since the proteolytic activity of this viral protease is essential for viral replication. Mutational and structural studies have identified substrate binding site and active site of  $M^{\text{pro}}$  that confers specificity for the Gln-P1 substrate residue in the active conformation (Jin et al., 2020a; Hilgenfeld, 2014; Yang et al., 2003). Recent structures of  $M^{\text{pro}}$  for SARS-CoV-2 have been solved in complexes with natural products and novel inhibitors very recently (Dai et al., 2020; Jin et al., 2020b). However, a structural description of these sites in the apo state has remained elusive.

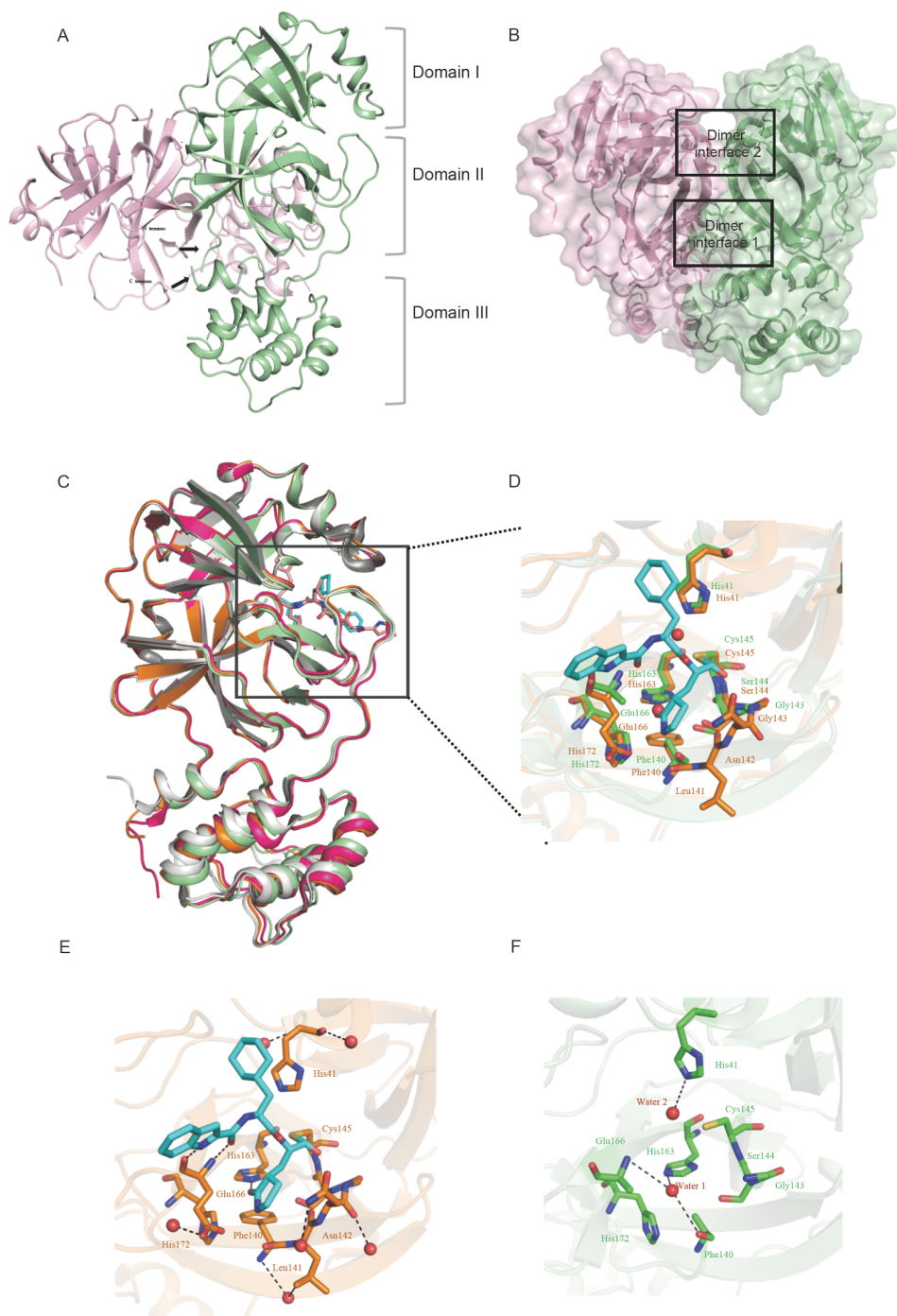
Here we report the first structure of  $M^{\text{pro}}$  for SARS-CoV-2 in the apo state under conditions close to the physiological state (pH 7.5) to an overall resolution of 1.9 Å (Table S1 in

†Contributed equally to this work

\*Corresponding authors (Qisheng Wang, email: [wangqisheng@zjlab.org.cn](mailto:wangqisheng@zjlab.org.cn); Jian Li, email: [rmsl\\_2040@163.com](mailto:rmsl_2040@163.com); Jin Zhang, email: [zhangxiaokong@hotmail.com](mailto:zhangxiaokong@hotmail.com))

Supporting Information). This structure highlights several key findings which will help guide drug discovery and functional studies (Figure 1A). Specifically, we found the

$M^{\text{pro}}$  forms a dimer in the crystal and has two distinct dimer interfaces, which are located in the N-terminal domain (residues 1–11) and the oxyanion loop (residues 137–145)



**Figure 1** The apo state structure of  $M^{\text{pro}}$  of SARS-CoV-2. A, The structure of the  $M^{\text{pro}}$  dimer is shown in stereo. Individual protomers are shown in red and green. N-finger (residues 1–7) is a loop located in the dimer interface and involved in the N-terminal auto cleavage. The domain I (residues 8–101) is comprised of three small  $\alpha$ -helices and six  $\beta$ -strands. The domain II (residues 102–184) consists of six  $\beta$ -strands. The domain III is composed of five  $\alpha$ -helices, which are closely related to the proteolytic activity. B, Two dimer interfaces of  $M^{\text{pro}}$  of SARS-CoV-2. The dimer interface of 1 and 2 are located in the oxyanion hole and N-terminal domain, respectively. C, Comparison of  $M^{\text{pro}}$  structures in the apo state of (green) in SARS-CoV-2, with inhibitor N3 in SARS-CoV-2 (red, PDB: 6LU7), with inhibitor 11b in SARS-CoV-2 (orange, PDB: 6LZE) and in the apo state of SARS (gray, PDB:1UJ1). N3 and 11b are shown in pink and cyan, respectively. D, Comparison of substrate binding site and active site in the apo state (green) and in the ligand-bound state in  $M^{\text{pro}}$  of SARS-CoV-2. 11b are shown in cyan. E, Structure of  $M^{\text{pro}}$  bound with 11b in an active conformation. Hydrogen bonds are shown as black dashed lines. F, Structure of  $M^{\text{pro}}$  in the apo state. Water 1 and 2 are shown in red spheres. Hydrogen bonds are indicated by dashed black lines.

(Figure 1B). Oxyanion holes, like the one found here, are classically found in proteases and are often formed by amides or positively charged residues (in this case predominantly made of histidine). They are crucial for stabilizing transition states during enzyme catalysis, thus facilitating enzyme activity. Comparison of our M<sup>pro</sup> structure in the apo state to the previously reported M<sup>pro</sup> structure in complex with an inhibitor revealed a backbone (C $\alpha$ ) RMSD of 0.92 Å showing a similar overall structure (Figure 1C) (Yang et al., 2003; Dai et al., 2020; Jin et al., 2020b). As in ligand-bound M<sup>pro</sup> structures, the protein consists of N-finger and three other domains that bind an inhibitor at the cleft between domains I and II (Figure 1A).

There were, however, several notable local differences between the apo and ligand-bound structures. Electron density of the N-finger (residues 1–2), oxyanion loop (residues 141–142), C-terminal domain (residues 299–306) were insufficient for backbone tracing, suggesting the flexibility of this region in the apo state. In addition, electron densities of side chains Phe140 and Glu166, which are key residues involved in the substrate binding are missing at this high resolution that may reflect different conformation of the apo state (Figure 1D).

As mentioned, the oxyanion hole composed of backbone amides or positively charged residues is directly related to the enzyme activity and substrate binding. In ligand-bound structures of M<sup>pro</sup>, the oxyanion hole consists of loop (residues 140–145), negatively charged residues Glu166, positively charged residues His41, His163 and His172 remains in an active conformation (Figure 1E) (Dai et al., 2020). A  $\pi$ - $\pi$  stacking interaction (Phe140/His163) is found in the oxyanion hole. A hydrogen bond and salt bridge involving Glu166 with water and His172 at the domain II further stabilize the oxyanion hole. However, the oxyanion loop (residues 137–145) is less well ordered and the side chains of Glu166 and Phe140 cannot be fit well due to poor density in our apo state structure. The salt bridge and  $\pi$ - $\pi$  stacking interactions between Glu166/His172 and Phe140/His163 are broken, resulting in rearrangements in this region and further collapses of the oxyanion hole (Figure 1F).

The N-finger plays an important role in the formation of the active site and auto cleavage activity of M<sup>pro</sup> (Hilgenfeld, 2014). Gly2 has interactions with Gly143 in the oxyanion loop in the neighboring protomer, stabilizing the active site and dimer in the active conformation, while the electron density of Gly2 is completely missing in our structure. Interestingly, His163 forms hydrogen bonds with water molecular (Water 1) in our structure, which is not observed in the ligand-bound structures. Another unprecedented water molecular (Water 2) is found near Cys145-His41 catalytic dyad in the active site, working as bridge for proton transfer. We speculate that these water molecules may affect the negatively charged oxygen of the substrate or inhibitor, which

suffers from steric hindrance, making rational drug design more difficult (Figure 1D–F).

In summary, we determined the apo state structure of M<sup>pro</sup> for SARS-CoV-2 under conditions close to the physiological state. The comparisons of M<sup>pro</sup> in different states reveal that the substrate binding site and the active site are more flexible in the apo state than that in the ligand-bound structures. Our structure does reveal that for drug discovery campaigns, the water molecules imbedded in the oxyanion hole and the corresponding interactions should be taken into consideration. Two water molecules are present in the oxyanion hole in our apo state structure, whereas in the ligand-bound structure, water molecular is absence in the same region. The water molecules, which is found near His163 and His41 in the occluded pocket, stabilizes the positively charged His residues, increasing the steric hindrance that may affect the catalytic efficiency of the enzyme. Altogether, the apo state structure of M<sup>pro</sup> for SARS-CoV-2 is an important complementary to the available structures. This structure provides novel and important insights that have broad implications for understanding the structural basis underlying enzyme activity, and can facilitate rational and structure-based approaches for the design of specific SARS-CoV-2 ligands as new therapeutic agents.

**Compliance and ethics** The author(s) declare that they have no conflict of interest.

**Acknowledgements** This work was supported by the Thousand Young Talents Program of China, the National Natural Science Foundation of China (31770795 and 81974514), the Jiangxi Province Natural Science Foundation (20181ACB20014), the Open Project of Key Laboratory of Prevention and Treatment of Cardiovascular and Cerebrovascular Diseases, Ministry of Education (XN201904), Gannan Medical University (QD201910), Jiangxi “Double Thousand Plan”, the Foreign Talent project of Jiangxi Province, and Ganzhou COVID-19 Emergency Research Project.

## References

- Dai, W., Zhang, B., Jiang, X.M., 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.
- Hilgenfeld, R. (2014). From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *FEBS J* 281, 4085–4096.
- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., et al. (2020a). Structure of M<sup>pro</sup> from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582, 289–293.
- Jin, Z., Zhao, Y., Sun, Y., Zhang, B., Wang, H., Wu, Y., Zhu, Y., Zhu, C., Hu, T., Du, X., et al. (2020b). Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nat Struct Mol Biol* 27, 529–532.
- 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.
- Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L.,

Ye, S., Pang, H., et al. (2003). The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci USA* 100, 13190–13195.

Zumla, A., Chan, J.F.W., Azhar, E.I., Hui, D.S.C., and Yuen, K.Y. (2016). Coronaviruses—Drug discovery and therapeutic options. *Nat Rev Drug Discov* 15, 327–347.

## Supporting information

The supporting information is available online at <https://doi.org/10.1007/s11427-020-1791-3>. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.