

Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors

<https://doi.org/10.1038/s41586-020-2223-y>

Received: 9 February 2020

Accepted: 1 April 2020

Published online: 9 April 2020

 Check for updates

Zhenming Jin^{1,2,10}, Xiaoyu Du^{2,10}, Yechun Xu^{3,10}, Yongqiang Deng^{4,10}, Meiqin Liu^{5,10}, Yao Zhao¹, Bing Zhang¹, Xiaofeng Li⁴, Leike Zhang⁵, Chao Peng⁶, Yinkai Duan¹, Jing Yu¹, Lin Wang¹, Kailin Yang⁷, Fengjiang Liu¹, Rendu Jiang⁵, Xinglou Yang⁵, Tian You¹, Xiaoce Liu¹, Xiuna Yang¹, Fang Bai¹, Hong Liu³, Xiang Liu⁸, Luke W. Guddat⁹, Wenqing Xu^{1,6}, Gengfu Xiao⁵, Chengfeng Qin⁴, Zhengli Shi⁵, Hualiang Jiang^{1,3,10}, Zihao Rao^{1,2,8,10} & Haitao Yang^{1,10}

A new coronavirus, known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is the aetiological agent responsible for the 2019–2020 viral pneumonia outbreak of coronavirus disease 2019 (COVID-19)^{1–4}. Currently, there are no targeted therapeutic agents for the treatment of this disease, and effective treatment options remain very limited. Here we describe the results of a programme that aimed to rapidly discover lead compounds for clinical use, by combining structure-assisted drug design, virtual drug screening and high-throughput screening. This programme focused on identifying drug leads that target **main protease (M^{pro}) of SARS-CoV-2**: M^{pro} is a key enzyme of coronaviruses and has a pivotal role in mediating viral replication and transcription, making it an attractive drug target for SARS-CoV-2^{5,6}. We identified a mechanism-based inhibitor (N3) by computer-aided drug design, and then determined the crystal structure of M^{pro} of SARS-CoV-2 in complex with this compound. Through a combination of structure-based virtual and high-throughput screening, we assayed more than 10,000 compounds—including approved drugs, drug candidates in clinical trials and other pharmacologically active compounds—as inhibitors of M^{pro}. Six of these compounds inhibited M^{pro}, showing **half-maximal inhibitory concentration values that ranged from 0.67 to 21.4 µM**. One of these compounds (ebesen) also exhibited promising antiviral activity in cell-based assays. Our results demonstrate the efficacy of our screening strategy, which can lead to the rapid discovery of drug leads with clinical potential in response to new infectious diseases for which no specific drugs or vaccines are available.

Coronaviruses infect humans and other animals and cause a variety of highly prevalent and severe diseases, including severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS)⁷. The SARS-CoV-2 genome comprises about 30,000 nucleotides: the replicase gene of SARS-CoV-2 encodes two overlapping polyproteins—pp1a and pp1ab—that are required for viral replication and transcription^{3,4}. The functional polypeptides are released from the polyproteins by extensive proteolytic processing, predominantly by the 33.8-kDa M^{pro} (also known as 3C-like protease). M^{pro} digests the polyprotein at at least 11 conserved sites, starting with the autolytic cleavage of this enzyme itself from pp1a and pp1ab⁸. The functional importance of M^{pro} in the viral life cycle, combined with the absence of closely related

homologues in humans, identify M^{pro} as an attractive target for the design of antiviral drugs⁹.

To facilitate the rapid discovery of antiviral compounds with clinical potential, we developed a strategy that combines structure-assisted drug design, virtual drug screening and high-throughput screening to repurpose existing drugs to target SARS-CoV-2 M^{pro}.

Establishing a high-throughput activity assay

Recombinant SARS-CoV-2 M^{pro} with native N and C termini was expressed in *Escherichia coli*, and subsequently purified (Extended Data Fig. 1a, b). The molecular mass of SARS-CoV-2 M^{pro} as determined

¹Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai, China. ²Laboratory of Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing, China. ³Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. ⁴Department of Virology, State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Beijing, China. ⁵CAS Key Laboratory of Special Pathogens, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan, China. ⁶National Facility for Protein Science in Shanghai, Zhangjiang Lab, Shanghai Advanced Research Institute, Chinese Academy of Science, Shanghai, China. ⁷Taussig Cancer Center, Cleveland Clinic, Cleveland, OH, USA. ⁸State Key Laboratory of Medicinal Chemical Biology, Frontiers Science Center for Cell Response, College of Life Sciences, College of Pharmacy, Nankai University, Tianjin, China. ⁹School of Chemistry and Molecular Biosciences, the University of Queensland, Brisbane, Queensland, Australia. ¹⁰These authors contributed equally: Zhenming Jin, Xiaoyu Du, Yechun Xu, Yongqiang Deng, Meiqin Liu. ✉e-mail: hljiang@sim.ac.cn; raozh@mail.tsinghua.edu.cn; yanght@shanghaitech.edu.cn

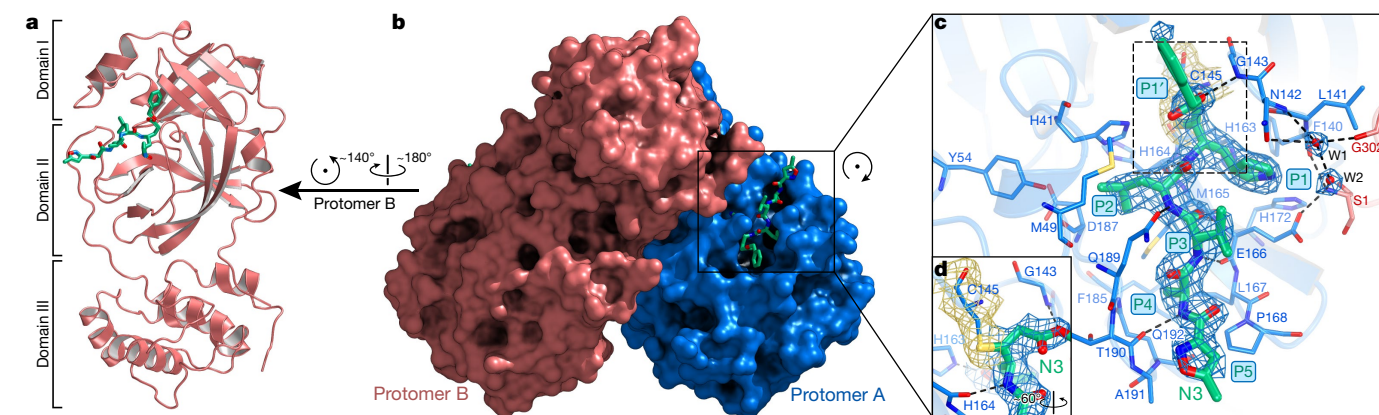


Fig. 1 | The crystal structure of SARS-CoV-2 M^{pro} in complex with N3.

a, Cartoon representation of one protomer of the dimeric M^{pro}–inhibitor complex. **b**, Surface representation of the homodimer of M^{pro}. Protomer A is in blue, protomer B is in salmon, N3 is presented as green sticks. **c**, An enlarged view of the substrate-binding pocket. The key residues that form the binding pocket are shown in sticks; the two water molecules (W1 and W2) are shown as

red spheres. The P1, P1', P2, P3, P4 and P5 sites of N3 are indicated. Hydrogen bonds that help to lock the inhibitor are shown in black dashed lines. The $2F_o - F_c$ density map contoured at 1.2σ is shown around N3 molecule (blue mesh), C145 of protomer A (yellow mesh) and the two waters (blue mesh). **d**, The C–S covalent bond.

by mass spectroscopy is 33797.0 Da, consistent with its theoretical molecular mass of 33796.8 Da. To characterize the enzymatic activity of SARS-CoV-2 M^{pro} and to carry out high-throughput screening of inhibitors, we developed a fluorescence resonance energy transfer assay. To do this, we designed and synthesized the fluorescently labelled substrate Mca–AVLQ↓SGFRK(Dnp)K, derived from the N-terminal autocleavage sequence of the viral protease, for time-dependent kinetic analysis (Extended Data Fig. 1e). The catalytic efficiency (k_{cat}/K_m) for SARS-CoV-2 M^{pro} was measured to be $28,500 \text{ M}^{-1} \text{ s}^{-1}$, which is slightly higher than that for SARS-CoV M^{pro} ($k_{cat}/K_m = 26,500 \text{ M}^{-1} \text{ s}^{-1}$)¹⁰ and more than 30-fold higher than that of human rhinovirus 3C protease ($k_{cat}/K_m = 920 \text{ M}^{-1} \text{ s}^{-1}$)¹¹.

N3 is a potent inhibitor of SARS-CoV-2 M^{pro}

In a previous study¹², a Michael acceptor inhibitor—known as N3—was developed using computer-aided drug design (Extended Data Fig. 1c). N3 can specifically inhibit M^{pro} from multiple coronaviruses, including SARS-CoV and MERS-CoV^{12–15}, and has displayed potent antiviral activity against infectious bronchitis virus in an animal model¹³. The 50% cytotoxicity concentration of N3 is $>133 \mu\text{M}$ (Extended Data Fig. 1f). Next, we constructed a homology model for SARS-CoV-2 M^{pro}, and used molecular docking to see whether N3 could target this M^{pro}. A docking pose showed that N3 could fit inside the substrate-binding pocket. To assess the efficacy of N3 for SARS-CoV-2 M^{pro}, we performed kinetic analysis. A progress curve showed that it is a time-dependent irreversible inhibitor of this enzyme. Further, the shape of this curve supports the mechanism of two-step irreversible inactivation. The inhibitor first associates with SARS-CoV-2 M^{pro} with a dissociation constant K_i , and then a stable covalent bond is formed between N3 and M^{pro}. The evaluation of this time-dependent inhibition requires both the equilibrium-binding constant (K_i , designated as k_2/k_1) and the inactivation-rate constant for covalent bond formation, k_3 . However, N3 exhibits a very potent inhibition of SARS-CoV-2 M^{pro}, such that the measurement of K_i and k_3 was not feasible (Extended Data Fig. 1d, e). When very rapid inactivation occurs, $k_{obs}/[I]$ was used to evaluate the inhibition as an approximation of the pseudo-second-order rate constant (k_3/K_i)¹². We determined the value of $k_{obs}/[I]$ of N3 for SARS-CoV-2 M^{pro} as $11,300 \pm 880 \text{ M}^{-1} \text{ s}^{-1}$, which suggests that this Michael acceptor was markedly inhibited.

Crystal structure of SARS-CoV-2 M^{pro}–N3

To elucidate the inhibitory mechanism of N3, we determined the crystal structure of SARS-CoV-2 M^{pro} in complex with N3 to a resolution of 2.1 \AA . The asymmetric unit contains only one polypeptide (Extended Data Table 1). However, two of these polypeptides (designated **protomer A and B**) associate to form a **dimer** by a crystallographic two-fold axis of symmetry (Fig. 1b). All of the residues 1–306 are visible in electron density maps. Each protomer is composed of three domains (Fig. 1a). Domain I (residues 8–101) and domain II (residues 102–184) have an antiparallel β -barrel structure. Domain III (residues 201–303) contains five α -helices arranged into a largely antiparallel globular cluster, and it is connected to domain II by a long loop region (residues 185–200). SARS-CoV-2 M^{pro} has a Cys–His catalytic dyad, and the substrate-binding site is located in a cleft between domain I and domain II. These features are similar to previously reported M^{pro} from other coronaviruses^{5,6,13–15}. The electron density map shows that N3 binds in the substrate-binding pocket in an extended conformation (Fig. 1c, Extended Data Fig. 2); the inhibitor backbone atoms form an antiparallel sheet with residues 164–168 of the long strand (residues 155–168) on one side, and with residues 189–191 of the loop that links domain II to domain III on the other.

Here we detail the specific interactions of N3 with M^{pro} (Fig. 1c, d). The electron density shows that the **Sy atom of C145 of protomer A forms a covalent bond (1.8 \AA) with the C β atom of the vinyl group**, confirming that the **Michael addition** has occurred. **The S1 subsite has an absolute requirement for Gln at the P1 position**. The side chains of F140, N142, E166, H163 and H172 of protomer A, and S1 of protomer B—as well as the main chains of F140 and L141 of protomer A—are involved in the formation of the S1 subsite, which also includes **two ordered water molecules** (which we refer to as W1 and W2). The lactam at P1 inserts into the S1 subsite and forms a hydrogen bond with H163 of protomer A. The side chain of Leu at the P2 site inserts deeply into the hydrophobic S2 subsite, which consists of the side chains of H41, M49, Y54 and M165, as well as the alkyl portion of the side chain of D187, of protomer A. The side chain of Val at P3 is solvent-exposed, which indicates that this site can tolerate a wide range of functional groups. The side chain of Ala at the P4 side is surrounded by the side chains of M165, L167, F185, Q192 of protomer A and the main chain of Q189 of protomer A, which together form a small hydrophobic pocket. P5 makes van der Waals contacts with P168 of protomer A, and with the backbone of

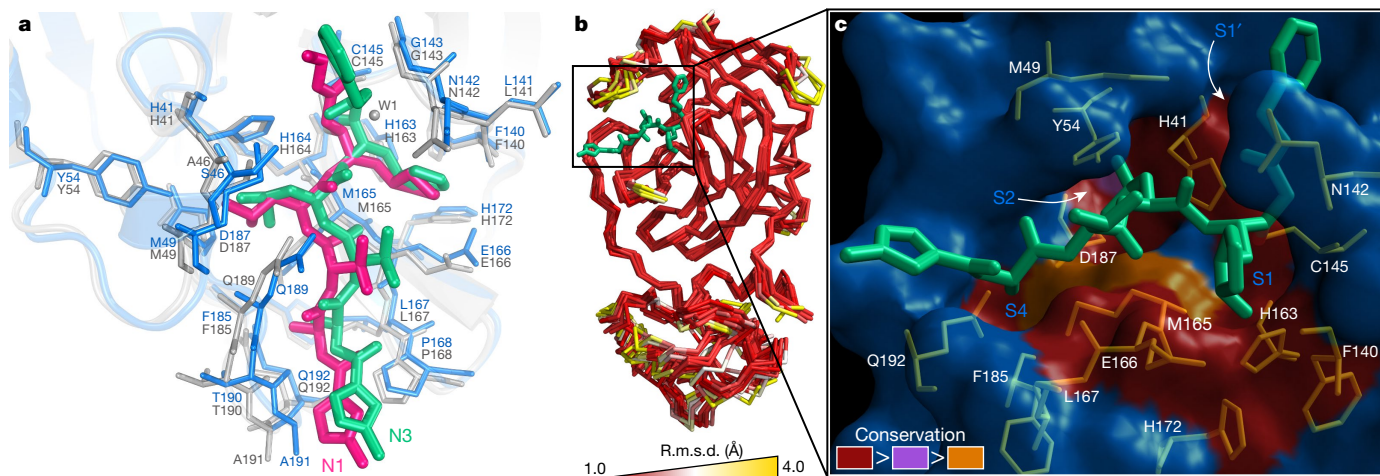


Fig. 2 | The substrate-binding pockets of M^{pro} across different species of coronavirus. **a**, Comparison of inhibitor binding mode between the structures of SARS-CoV-2 M^{pro} -N3 and SARS-CoV M^{pro} -N1. SARS-CoV-2 M^{pro} is shown in marine cartoon; SARS-CoV M^{pro} in grey; N3 in green sticks; and N1 in hot pink. **b**, Superposition of crystal structures of M^{pro} (C α 1–300) from 12 coronaviruses, including SARS-CoV-2, SARS-CoV, MERS-CoV, human coronavirus (HCoV)-HKU1, bat coronavirus HKU4, mouse hepatitis virus A59, porcine epidemic diarrhoea virus, feline infectious peritonitis virus, transmissible

gastroenteritis virus, HCoV-NL63, HCoV-229E and infectious bronchitis virus. The colour spectrum represents the root-mean-square deviation (r.m.s.d.) of the aligned C α atoms. **c**, Surface presentation of conserved substrate-binding pockets of M^{pro} from 12 coronaviruses. Red, residues are entirely identical among M^{pro} in all 12 viruses; violet, conserved substitution in M^{pro} of one of the coronaviruses; orange, conserved substitution in M^{pro} of more than one of the coronaviruses. S1, S2, S4 and S1' subsites are indicated.

residues 190–191. The bulky benzyl group extends into the S1' site, forming van der Waals interactions with T24 and T25 of protomer A. N3 forms multiple hydrogen bonds with the main chain of the residues in the substrate-binding pocket, which also helps to lock the inhibitor inside the substrate-binding pocket.

An overlay of the structures of SARS-CoV-2 M^{pro} in complex with N3 and SARS-CoV M^{pro} in complex with N1 (ref. ¹²), an alternative Michael acceptor inhibitor, shows that N3 and N1 bind to M^{pro} in a similar mode (Fig. 2a, Extended Data Fig. 3). The major difference lies in the P1' site. Compared with the benzyl ester portion of N3 in the SARS-CoV-2 M^{pro} -N3 structure, the ethyl ester portion in N1 in SARS-CoV M^{pro} -N1 adopts a slightly different conformation. This can be attributed to an ordered water (W1) in the SARS-CoV M^{pro} -N1 structure, which makes a long-distance hydrogen bond to the carboxylate oxygen of the ester and also forms two hydrogen bonds from the backbone NH of G143 and the side chain of N142. A previous study proposed that M^{pro} has substrate-recognition pocket that is highly conserved among all coronaviruses, and that this pocket could serve as a drug target for the design of broad-spectrum inhibitors¹². The recent discovery of new coronaviruses, and the accumulation of structural data for M^{pro} from coronaviruses of various species, provided the opportunity to further examine this hypothesis. Superposition of the 12 crystal structures of M^{pro} (refs. ^{12–21}) shows that the most variable regions are the helical domain III and surface loops, and that the substrate-binding pocket (located in a cleft between domain I and domain II) is highly conserved among M^{pro} in all coronaviruses; this suggests that antiviral inhibitors targeting this pocket should have wide-spectrum activity against coronaviruses (Fig. 2b, c).

Virtual screening

The structure of SARS-CoV-2 M^{pro} in complex with N3 provides a model for identifying lead inhibitors to target SARS-CoV-2 M^{pro} using in silico screening. To achieve this, we docked an in-house database of potential binding compounds using Glide (v.8.2)²². The results show that cinanserin fits snugly into the substrate-binding pocket, through cation- π interactions with H41 and E166 of M^{pro} . Subsequently, we determined

this compound has a half-maximal inhibitory concentration (IC_{50}) value of 125 μ M for M^{pro} . Moreover, cinanserin is a well-characterized serotonin antagonist, which underwent preliminary clinical testing in humans in the 1960s²³ and has previously been shown to inhibit SARS-CoV M^{pro} (ref. ²⁴). The 50% cytotoxicity concentration of cinanserin is >200 μ M (Extended Data Fig. 4); thus, it has potential for optimization as an antiviral drug lead.

High-throughput screening

Next, we used our fluorescence resonance energy transfer assay to screen a library of about 10,000 compounds, consisting of approved drugs, clinical-trial drug candidates and natural products. The primary hits were seven compounds, including approved drugs (disulfiram and carmofur) as well as preclinical or clinical-trial drug candidates (ebelsen, shikonin, tideglusib, PX-12 and TDZD-8). We then determined the IC_{50} values of these seven compounds, which range from 0.67 to 21.4 μ M (Fig. 3). Ebelsen has the strongest inhibition of M^{pro} activity, with an IC_{50} of 0.67 μ M. Using a previously described detergent-based assay²⁵, we found that TDZD-8 is an aggregate-based inhibitor that might nonspecifically inhibit M^{pro} (Extended Data Fig. 5); it was, therefore, not considered for further investigation. Next, we set out to identify the potential covalent inhibitors among these compounds through tandem mass spectrometry analysis. The tandem mass spectrometry data show that ebelsen, PX-12 and carmofur are all able to covalently bind to C145 of the catalytic dyad in SARS-CoV-2 M^{pro} . However, PX-12 and carmofur completely modified M^{pro} , whereas ebelsen could only partially modify this cysteine of the viral protease (Extended Data Fig. 6). As ebelsen has a stronger inhibitory effect than the other compounds, there is a possibility that ebelsen could also inhibit M^{pro} through noncovalent binding. It is likely that a portion of the hits identified by screening are covalently bonded to the catalytic cysteine of M^{pro} through their sulfhydryl groups. In general, such molecules are expected to be promiscuous binders and therefore—as they stand—may have limited potential as drug leads. As our structural data are based on N3, we investigated whether molecular docking could predict how disulfiram, tideglusib and shikonin bind to

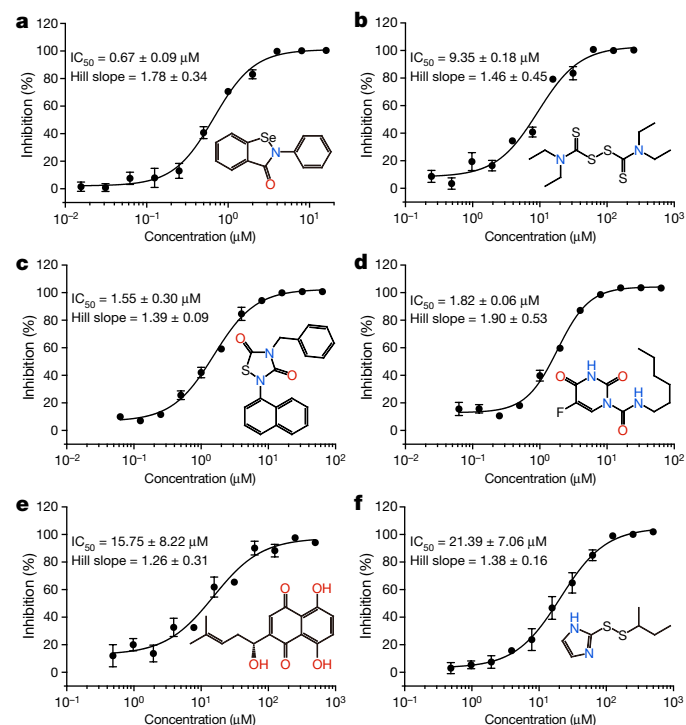


Fig. 3 | Drug leads inhibit the activity of SARS-CoV-2 M^{pro}. **a–f**, The hydrolytic activity of SARS-CoV-2 M^{pro} was measured in the presence of increasing concentrations of the drug candidates. **a**, Ebselen. **b**, Disulfiram. **c**, Tideglusib. **d**, Carmofur. **e**, Shikonin. **f**, PX-12. Dose–response curves for IC_{50} values were determined by nonlinear regression. All data are shown as mean \pm s.e.m., $n = 3$ biological replicates.

this protein. In all cases, reasonable docking poses were found, which demonstrates that they could fit inside the substrate-binding pocket (Extended Data Fig. 7).

Antiviral activity assay

To further substantiate the enzymatic inhibition results *in vitro*, we evaluated whether these compounds could prevent viral replication in cell-based assays. As shown in Fig. 4a, quantitative real-time RT–PCR (qRT–PCR) demonstrated that, among these compounds, ebselen and N3 showed the strongest antiviral effects at a concentration of 10 μM treatment in SARS-CoV-2-infected Vero cells. We performed a plaque-reduction assay (Extended Data Fig. 8) to further assess the efficacy of these two compounds in protecting cells. Ebselen and N3 displayed inhibition against SARS-CoV-2 with individual half-maximal effective concentration (EC_{50}) values of 4.67 μM and 16.77 μM , respectively (Fig. 4b, c). The dose–response curves suggest that both of these compounds may be able to penetrate the cellular membrane to access their targets. Ebselen is an organoselenium compound with anti-inflammatory, anti-oxidant and cytoprotective properties. This compound has previously been investigated for the treatment of multiple diseases, including bipolar disorders²⁶ and hearing loss^{27,28}. **Ebselen has extremely low cytotoxicity** (the median lethal dose in rats is $>4,600 \text{ mg kg}^{-1}$, when taken orally)²⁹, and its **safety in humans** has been evaluated in a number of clinical trials^{27,28,30}. These data strongly suggest the clinical potential of ebselen for the treatment of coronaviruses. It is also interesting to note that **cinanserin displayed moderate inhibition against SARS-CoV-2 with an EC_{50} value of 20.61 μM** , as shown from qRT–PCR analysis (Extended Data Fig. 4). This value is superior to that in the enzymatic inhibition assay, which suggests that cinanserin might have multidrug targets in preventing viral infection. In further

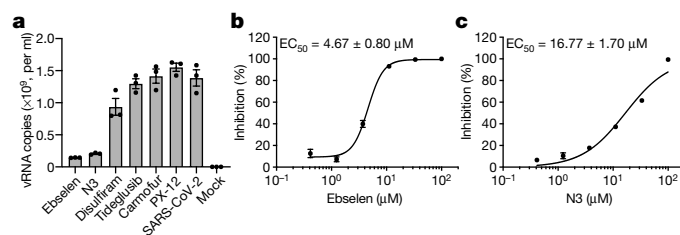


Fig. 4 | Antiviral activities of the drug leads against SARS-CoV-2. **a**, The quantification of absolute viral RNA (vRNA) copies (per ml) in the supernatant at 72 h after infection, determined by qRT–PCR analysis. Data are mean \pm s.e.m., $n = 3$ biological replicates. **b, c**, Dose–response curves for ebselen (**b**) and N3 (**c**) in the plaque-reduction assay. All data are shown as mean \pm s.e.m., $n = 4$ biological replicates.

studies, the selection and characterization of drug-resistant mutants will help to clarify the mode of action of cinanserin.

Discussion

Our crystal structural and docking data show that **the drug leads we identified can bind to the substrate-binding pocket of SARS-CoV-2 M^{pro}**, which is highly conserved among all coronaviruses. This strongly supports our hypothesis that the development of a single antiviral agent targeting M^{pro}, or such an agent used in combination with other potential therapies, could provide an effective first line of defence against all diseases associated with coronaviruses.

In the last 20 years, new infectious agents—such as SARS and MERS—have emerged and caused epidemics⁷. The timely development of effective antiviral agents for clinical use is extremely challenging, because conventional drug development approaches normally take years of investigation and cost billions of dollars. **The repurposing of approved pharmaceutical drugs and drug candidates** provides an alternative approach that allows for the rapid identification of potential drug leads to manage rapidly emerging viral infections. Cell-based phenotypic screening has proven to be valuable³¹, but the complexity of this approach is not readily compatible with high-throughput pipelines, and it cannot identify the molecular target or mechanism of action³². In this study, the convergence of structure-based *ab initio* drug design, virtual screening and high-throughput screening proved to be an efficient strategy to find antiviral leads against SARS-CoV-2. The methods presented here can greatly assist in the rapid discovery of drug leads with clinical potential in response to new emerging infectious diseases that currently lack specific drugs and vaccines.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2223-y>.

- Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
- Li, Q. et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
- Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
- Wu, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
- Anand, K. et al. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra α -helical domain. *EMBO J.* **21**, 3213–3224 (2002).
- Yang, H. et al. 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 (2003).

7. de Wit, E., van Doremalen, N., Falzarano, D. & Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 (2016).
8. Hegyi, A. & Ziebuhr, J. Conservation of substrate specificities among coronavirus main proteases. *J. Gen. Virol.* **83**, 595–599 (2002).
9. Pillaiyar, T., Manickam, M., Namasivayam, V., Hayashi, Y. & Jung, S. H. An overview of severe acute respiratory syndrome-coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy. *J. Med. Chem.* **59**, 6595–6628 (2016).
10. Xue, X. et al. Production of authentic SARS-CoV M^{pro} with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. *J. Mol. Biol.* **366**, 965–975 (2007).
11. Wang, Q. M., Johnson, R. B., Cox, G. A., Villarreal, E. C. & Loncharich, R. J. A continuous colorimetric assay for rhinovirus-14 3C protease using peptide *p*-nitroanilides as substrates. *Anal. Biochem.* **252**, 238–245 (1997).
12. Yang, H. et al. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol.* **3**, e324 (2005).
13. Xue, X. et al. Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J. Virol.* **82**, 2515–2527 (2008).
14. Ren, Z. et al. The newly emerged SARS-like coronavirus HCoV-EMC also has an “Achilles’ heel”: current effective inhibitor targeting a 3C-like protease. *Protein Cell* **4**, 248–250 (2013).
15. Wang, F., Chen, C., Tan, W., Yang, K. & Yang, H. Structure of main protease from human coronavirus NL63: insights for wide spectrum anti-coronavirus drug design. *Sci. Rep.* **6**, 22677 (2016).
16. Zhao, Q. et al. Structure of the main protease from a global infectious human coronavirus, HCoV-HKU1. *J. Virol.* **82**, 8647–8655 (2008).
17. Lee, C.-C. et al. Structural basis of inhibition specificities of 3C and 3C-like proteases by zinc-coordinating and peptidomimetic compounds. *J. Biol. Chem.* **284**, 7646–7655 (2009).
18. St John, S. E., Tomar, S., Stauffer, S. R. & Mesecar, A. D. Targeting zoonotic viruses: structure-based inhibition of the 3C-like protease from bat coronavirus HKU4—the likely reservoir host to the human coronavirus that causes Middle East respiratory syndrome (MERS). *Bioorg. Med. Chem.* **23**, 6036–6048 (2015).
19. Wang, F. et al. Crystal structure of feline infectious peritonitis virus main protease in complex with synergetic dual inhibitors. *J. Virol.* **90**, 1910–1917 (2016).
20. Wang, F. et al. Michael acceptor-based peptidomimetic inhibitor of main protease from porcine epidemic diarrhea virus. *J. Med. Chem.* **60**, 3212–3216 (2017).
21. Cui, W. et al. The crystal structure of main protease from mouse hepatitis virus A59 in complex with an inhibitor. *Biochem. Biophys. Res. Commun.* **511**, 794–799 (2019).
22. Friesner, R. A. et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).
23. Rubin, B. & Waugh, M. H. Antiphlogistic effects of antiserotonin (SQ 10,643) and aminopyrine in rats versus endotoxin and other agents. *Proc. Soc. Exp. Biol. Med.* **119**, 438–443 (1965).
24. Chen, L. et al. Cinanserin is an inhibitor of the 3C-like proteinase of severe acute respiratory syndrome coronavirus and strongly reduces virus replication in vitro. *J. Virol.* **79**, 7095–7103 (2005).
25. Feng, B. Y. & Shoichet, B. K. A detergent-based assay for the detection of promiscuous inhibitors. *Nat. Protocols* **1**, 550–553 (2006).
26. Singh, N. et al. A safe lithium mimetic for bipolar disorder. *Nat. Commun.* **4**, 1332 (2013).
27. Lynch, E. & Kil, J. Development of ebselen, a glutathione peroxidase mimic, for the prevention and treatment of noise-induced hearing loss. *Semin. Hear.* **30**, 047–055 (2009).
28. Kil, J. et al. Safety and efficacy of ebselen for the prevention of noise-induced hearing loss: a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* **390**, 969–979 (2017).
29. Renson, M., Etschenberg, E. & Winkelmann, J. 2-Phenyl-1, 2-benzisoxaselenazol-3 (2H)-one containing pharmaceutical preparations and process for the treatment of rheumatic diseases. US patent **4**, 352, 799 (1982).
30. Masaki, C. et al. Effects of the potential lithium-mimetic, ebselen, on impulsivity and emotional processing. *Psychopharmacology (Berl.)* **233**, 2655–2661 (2016).
31. Xu, M. et al. Identification of small-molecule inhibitors of Zika virus infection and induced neural cell death via a drug repurposing screen. *Nat. Med.* **22**, 1101–1107 (2016).
32. Aulner, N., Danckaert, A., Ihm, J., Shum, D. & Shorte, S. L. Next-generation phenotypic screening in early drug discovery for infectious diseases. *Trends Parasitol.* **35**, 559–570 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Cloning, protein expression and purification of SARS-CoV-2 M^{pro}

The full-length gene that encodes SARS-CoV-2 M^{pro} (NC_045512) was optimized and synthesized for *E. coli* expression (Genewiz). The cloning strategy for producing authentic viral M^{pro} has previously been reported¹⁰. The expression plasmid was transformed into *E. coli* BL21 (DE3) cells and then cultured in Luria broth medium containing 100 µg/ml ampicillin at 37 °C. When the cells were grown to an optical density at 600 nm of 0.6–0.8, 0.5 mM IPTG was added to the cell culture to induce the expression at 16 °C. After 10 h, the cells were collected by centrifugation at 3,000g. The cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl), lysed by high-pressure homogenization, and then centrifuged at 25,000g for 40 min. The supernatant was loaded onto Ni-NTA affinity column (Qiagen), and washed in the resuspension buffer containing 20 mM imidazole. The His-tagged M^{pro} was eluted by cleavage buffer (50 mM Tris-HCl pH 7.0, 150 mM NaCl) including 300 mM imidazole. Human rhinovirus 3C protease was added to remove the C-terminal His tag. The M^{pro} was further purified by ion-exchange chromatography and size-exclusion chromatography. Coronavirus M^{pro} exists as a mixture of monomers and dimers in solution³³. The purified M^{pro} was stored in 50 mM Tris-HCl pH 7.3, 1 mM EDTA.

Crystallization, data collection and structure determination

SARS-CoV-2 M^{pro} was incubated with 10 mM N3 for 30 min and the complex (5 mg/ml) was crystallized by hanging drop vapour diffusion method at 20 °C. The best crystals were grown with well buffer containing 0.1 M MES pH 6.0, 2% polyethylene glycol (PEG) 6000, 3% DMSO, 1 mM DTT. The cryo-protectant solution contained 0.1 M MES pH 6.0, 30% PEG 400.

X-ray data were collected on beamline BL17U1 at Shanghai Synchrotron Radiation Facility (SSRF) at 100 K and at a wavelength of 1.07180 Å using an Eiger X 16M image plate detector. Data integration and scaling were performed using the program Xia2³⁴. The structure was determined by molecular replacement with the Phaser module³⁵ in CCP4³⁶ using the SARS-CoV M^{pro} (RCSP Protein Data Bank code (PDB) 2H2Z) as a search template. The output model from molecular replacement was subsequently subjected to iterative cycles of manual model adjustment with Coot³⁷ and refinement was finished with Phenix³⁸. The inhibitor N3 was built according to the omit map. The phasing and refinement statistics are summarized in Extended Data Table 1. The R_{work} and R_{free} values are 0.202 and 0.235, respectively. There are 97.3% of the residues in the most favoured regions of the Ramachandran plot, and no residues are found in disallowed regions.

Enzymatic activity and inhibition assays

The enzyme activity assays have previously been described¹⁰. In brief, the activity of SARS-CoV-2 M^{pro} was measured by a continuous kinetic assay, with the substrate Mca-AVLQ↓SGFR-K(Dnp)K (GL Biochem), using wavelengths of 320 nm and 405 nm for excitation and emission, respectively. The assay started by immediately mixing 0.2 µM SARS-CoV-2 M^{pro} with different concentrations of substrate (2.5–100 µM). Fluorescence intensity was monitored with an EnVision multimode plate reader (Perkin Elmer). Initial rates were obtained by fitting the linear portion of the curves to a straight line. The kinetic parameters K_m and k_{cat} were calculated from a double-reciprocal plot. As N3 is a mechanism-based irreversible inhibitor for SARS-CoV-2 M^{pro}, $k_{\text{obs}}/[I]$ was used as an approximation of the pseudo-second-order rate constant to evaluate the inhibition effect of the inhibitor N3 (ref.¹²). In this case, the measurement was carried out with 0.2 µM of enzyme, 20 µM of substrate and inhibitor at 6 different concentrations (0–1 µM).

Virtual screening

The virtual screening was performed using our in-house database via a workflow application of Glide (v.8.2)²² in Maestro (Schrödinger 2019-1a). All compounds in the database were considered to be at **pH 7.4 ± 0.2** to estimate their protonation state using the program EpiK³⁹. Their three-dimensional (3D) conformations were generated by the ligPrep module of Maestro. The structure of SARS-CoV-2 M^{pro} (PDB 6LU7) was used to generate the receptor grid for docking simulations. The centre of the active site of the grid was determined according to the position of N3 in the structure. The flexibility of the receptor hydroxyl and thiol groups in side chains of C145, S46 and Y54 were considered. At the very beginning, we performed a relatively fast but raw screening using the Glide standard precision model, and the top 20% of compounds were kept. Finally, the candidate molecules were picked by analysing the predicted binding modes and their scores.

High-throughput drug screening and IC₅₀ measurement

Potential inhibitors against SARS-CoV-2 M^{pro} were screened by an enzymatic inhibition assay. When the different compounds were added into the enzymatic reaction mixture, the change of initial rates was calculated to evaluate their inhibitory effect. Five drug libraries—the Approved Drug Library (Target Mol), Clinic Compound Library (Target Mol), FDA-approved Drug Library (Selleck), Natural Product Library (Selleck), and Anti-virus Drug Library (Shanghai Institute for Advanced Immunochemical Studies)—that together comprised about 10,000 compounds were used. The preliminary screening reaction mixture included 0.2 µM protein, 20 µM substrate and 50 µM compounds. The compounds of interest were defined as those with a percentage of inhibition over 60% compared with the reaction in the absence of inhibitor. IC₅₀ values of 7 drug leads were measured using 0.2 µM protein, 20 µM substrate and 11 different inhibitor concentrations. To exclude inhibitors possibly acting as **aggregators**, a detergent-based control was performed by adding 0.001% or 0.01% freshly made up **Triton X-100** to the reaction at the same time²⁵. All experimental data was analysed using GraphPad Prism. All experiments were performed in triplicate.

Molecular docking

To understand the binding interaction of these molecules with SARS-CoV-2 M^{pro}, two different molecular docking methods (**Glide** (v.8.2)²² and **iFitDock**⁴⁰) were used to predict their binding poses. Then, a 3D molecular similarity calculation method, SHAFTS⁴¹, was used for enumeration of the molecular alignment poses by matching the critical pharmacophore and volumetric overlay between the N3 molecule within the M^{pro} structure and the other drug candidates. Then, the obtained optimal superposition of these molecules was used to assess the reasonability of the predicted binding poses from the two docking methods, and only the binding orientations that were consistent among different methods were kept for constructing the initial complexes. Finally, these complexes were further optimized and re-scored using the MM-GBSA module⁴² of Schrödinger, and the residues within 5 Å around the ligand were refined.

Antiviral and cytotoxicity assays for compounds from high-throughput screening

The in vitro antiviral efficacy of the drug candidates on Vero cells was determined by qRT-PCR. About 1×10^4 Vero cells were seeded into a 96-well plate and incubated for 20–24 h at 37 °C. All the infection experiments were performed at biosafety level-3 (BSL-3). Cells were pretreated with the drug candidates (10 µM) for 1 h; SARS-CoV-2 (multiplicity of infection (MOI) of 0.01) was subsequently added to allow infection for 2 h. Then, the virus–drug mixture was removed and cells were further cultured with fresh drug-containing medium. At 72 h after infection, vRNA was extracted from the culture supernatant

using QIAamp viral RNA mini kit (Qiagen) according to the manufacturer's recommendation, and detected by qRT-PCR assay using the SARS-CoV-2-specific primers. Because shikonin showed cellular toxicity at the test concentration, its antiviral activity assay did not proceed further. vRNA copies per millilitre were determined using a synthetic RNA fragment to amplify the target region. The linearized plasmid containing the S gene of SARS-CoV-2 was subjected to in vitro transcription. The resulting RNA transcripts were purified and then quantified using spectrophotometry on Nanodrop 2000 (Thermo Fisher Scientific). The purified RNA was diluted tenfold serially using RNase-free water and was detected using qRT-PCR. Threshold cycle (C_t) values for the known concentrations of the RNA were plotted against the log of the number of genome-equivalent copies. The resultant standard curve was used to determine the number of genome equivalents of vRNA in the samples. The determination of the detection limit was based on the lowest level at which vRNA was detected and remained within the range of linearity of a standard curve (C_t value of 38). TaqMan primers for SARS-CoV-2 are 5'-TCCTGGTGATTCTTCTTCAGG-3' and 5'-TCTGAGAGAGGGTCAAGTGC-3' with SARS-CoV-2 probe 5'-FAM-AGCTGCAGCACCAGCTGTCCA-BHQ1-3'. The cytotoxicity of the tested drugs on Vero cell were determined by MTS cell proliferation assays (Promega). Ten thousand cells were seeded into a 96-well plate and incubated for 20–24 h at 37 °C. After that, the medium was removed, and 100 µl of medium containing decreasing concentrations of antiviral compounds was added to the wells. After 4 days incubation at 37 °C, MTS assays were performed according to manufacturer's protocols. All experiments were performed in triplicate. Vero cells were obtained from ATCC (American Type Culture Collection) with authentication service. All cell lines tested negative for mycoplasma contamination. No commonly misidentified cell lines were used.

Antiviral and cytotoxicity assays for cinanserin

For the antiviral assay, a clinical isolate of SARS-CoV-2³ was propagated in Vero E6 cells, and viral titre was determined as previously described⁴³. All of the infection experiments were performed at BSL-3. Preseeded Vero E6 cells (5×10^4 cells per well) were pretreated with the different concentrations of cinanserin for 1 h and the virus was subsequently added (MOI of 0.05) to allow infection for 2 h. Then, the virus–drug mixture was removed and cells were further cultured with fresh drug-containing medium. At 24 h after infection, the cell supernatant was collected and vRNA in supernatant was subjected to qRT-PCR analysis. For cytotoxicity assays, Vero E6 cells were suspended in growth medium in 96-well plates. The next day, appropriate concentrations of cinanserin were added to the medium. After 24 h, the relative numbers of surviving cells were measured by CCK8 (Beyotime) assay in accordance with the manufacturer's instructions. All experiments were performed in triplicate. Vero E6 cells were obtained from ATCC with authentication service. All cell lines tested negative for mycoplasma contamination. No commonly misidentified cell lines were used.

Plaque-reduction assays

One hundred thousand Vero E6 cells were seeded in a 24-well plate and treated with different doses of the inhibitors. All of the infection experiments were performed at BSL-3. Inhibitors with different dilution concentrations were mixed with SARS-CoV-2 (100 plaque-forming units), and 200 µl mixtures were inoculated onto monolayer Vero E6 cells for 1 h. After removing the supernatant, the plate was washed twice with DMEM medium, cells were incubated with 0.9% agarose containing appropriate concentrations of inhibitors. The overlay was discarded at 4 days after infection, and cells were fixed for 30 min in 4% polyoxymethylene and stained with crystal violet working solution. The plaque-forming units were determined. All experiments were performed in four biological replicates.

Intact protein analysis

In brief, 2.5 µl of compounds (10 mM in DMSO) was added into 50 µl of SARS-CoV-2 M^{pro} (10 mg/ml). The mixtures were kept at room temperature for 30 min. Liquid chromatography–mass spectrometry analyses were performed in positive-ion mode with a quadrupole-time-of-flight mass spectrometer (Agilent 6550) coupled with a high-performance liquid chromatograph (HPLC, Agilent 1260) for detecting the molecular weight of intact proteins. The samples were eluted from a Phenomenex Jupiter C4 300Å LC column (2×150 mm, 5 µm) over a 15-min gradient from 5% to 100% acetonitrile containing 0.1% formic acid at a flow rate of 0.5 ml/min. The acquisition method in positive-ion mode with Dual Agilent Jet Stream electrospray voltage used a capillary temperature of 250 °C, a fragmentor of 175 V and a capillary voltage of 3,000 V. Mass deconvolution was performed using Agilent MassHunter Qualitative Analysis B.06.00 software with BioConfirm Workflow.

Tandem mass spectrometry analysis

The samples were precipitated and redissolved by 8 M urea, and then digested for 16 h at 25 °C by chymotrypsin at an enzyme-to-substrate ratio of 1:50 (w/w). The digested peptides were desalted and loaded onto a homemade 30-cm-long pulled-tip analytical column (ReproSil-Pur C18 AQ 1.9-µm particle size, Dr Maisch, 75-µm inner diameter \times 360-µm outer diameter) connected to an Easy-nLC1200 UHPLC (Thermo Fisher Scientific) for mass spectrometry analysis. The elution gradient and mobile phase constitution used for peptide separation were as follows: 0–1 min, 4–8% B; 1–96 min, 8–35% B; 96–104 min, 35–60% B; 105–120 min, 60–100% B (mobile phase A: 0.1% formic acid in water; mobile phase B: 0.1% formic acid in 80% acetonitrile) at a flow rate of 300 nl/min. Peptides eluted from the liquid chromatography column were directly electro-sprayed into the mass spectrometer with the application of a distal 1.8-kV spray voltage. Survey full-scan mass spectra (from m/z 300–1,800) were acquired in the Orbitrap analyser (Q Exactive, Thermo Fisher Scientific) with resolution $r = 70,000$ at m/z 400. The top 20 tandem mass spectrometry (MS/MS) events were sequentially generated and selected from the full mass spectrum at a 30% normalized collision energy. The dynamic exclusion time was set to 10 s. One acquisition cycle includes one full-scan mass spectrum followed by top 20 MS/MS events, sequentially generated on the first to the twentieth most intense ions selected from the full mass spectrum at a 28% normalized collision energy. The acquired MS/MS data were analysed using the UniProtKB *E. coli* database (database released on 11 November 2016) and SARS-CoV-2 nsp5, using Protein Discoverer 2.1. To accurately estimate peptide probabilities and false-discovery rates, we used a decoy database containing the reversed sequences of all the proteins appended to the target database. The false-discovery rate was set to 0.01. Mass tolerance for precursor ions was set to 20 ppm. Chymotrypsin was defined as cleavage enzyme and the maximal number of missed cleavage sites was set to four. Protein N terminus acetylation, methionine oxidation and compounds covalent bindings were set as variable modifications. The modified peptides were manually checked and labelled.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The coordinates and structure factors for SARS-CoV-2 M^{pro} in complex with the inhibitor N3 have been deposited in the PDB with accession number [6LU7](#), deposited on the 26 January 2020 and released on the 5 February 2020. While this work was under review, we solved the complex structure at a higher resolution (1.7 Å); the relevant coordinates and structure factors have been deposited in the PDB with accession

number **7BQY**. Any other relevant data are available from the corresponding authors upon reasonable request.

33. Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R. **Coronavirus main proteinase (3CL^{pro}) structure: basis for design of anti-SARS drugs.** *Science* **300**, 1763–1767 (2003).
34. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
35. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
36. Potterton, L. et al. CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr. D* **74**, 68–84 (2018).
37. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
38. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
39. Greenwood, J. R., Calkins, D., Sullivan, A. P. & Shelley, J. C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. Aided Mol. Des.* **24**, 591–604 (2010).
40. Bai, F. et al. Free energy landscape for the binding process of huperzine A to acetylcholinesterase. *Proc. Natl Acad. Sci. USA* **110**, 4273–4278 (2013).
41. Liu, X., Jiang, H. & Li, H. SHAFTS: a hybrid approach for 3D molecular similarity calculation. 1. Method and assessment of virtual screening. *J. Chem. Inf. Model.* **51**, 2372–2385 (2011).
42. Guimarães, C. R. W. & Cardozo, M. MM-GB/SA rescoring of docking poses in structure-based lead optimization. *J. Chem. Inf. Model.* **48**, 958–970 (2008).
43. Wang, M. et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **30**, 269–271 (2020).

Acknowledgements We thank Y. Lei and J. Kong from High Throughput Platform and staff from Analytical Chemistry Platform at the Shanghai Institute for Advanced Immunochemical

Studies for their technical support; the National Centre for Protein Science Shanghai and The Molecular and Cell Biology Core Facility of the School of Life Science and Technology, ShanghaiTech University for use of their instrumentation and technical assistance; Z. Liu and H. Su for discussion; and the staff from beamlines BL17U1, BL18U1 and BL19U1 at the Shanghai Synchrotron Radiation Facility. This work was supported by grants from National Key R&D Program of China (grant no. 2017YFC0840300 to Z.R. and no. 2020YFA0707500), Project of International Cooperation and Exchanges NSFC (grant no. 81520108019 to Z.R.), Science and Technology Commission of Shanghai Municipality (grant no. 20431900200) and Department of Science and Technology of Guangxi Zhuang Autonomous Region (grant no. 2020AB40007).

Author contributions Z.R. and H.Y. conceived the project; Z.J., H.J., Z.R. and H.Y. designed the experiments; Z.J., X.D., Y. Duan, J.Y., T.Y., Xiaoe Liu and Xiuna Yang cloned, expressed, purified and crystallized proteins; Z.J., Y.Z., B.Z. and F.L. collected the diffraction data; B.Z. and Xiang Liu solved the crystal structure; Z.J., X.D., Y. Duan and J.Y. performed enzymatic activity and inhibition assays, high-throughput drug screening and IC₅₀ measurements; L.W. and F.B. performed virtual screening and molecular docking; Y.X., L.Z. and H.L. performed enzymatic inhibition, cell-based antiviral and cytotoxicity assays for cinanserin; Y. Deng and X.L. performed qRT-PCR analysis and cytotoxicity assay of N3; M.L., R.J. and Xinglou Yang performed plaque-reduction assays; C.P. performed intact protein and MS/MS analyses; Z.J., X.D., Y.X., Y. Deng, C.P., F.B., H.L., Xiang Liu, K.Y., L.G., W.X., G.X., C.Q., Z.S., H.J., Z.R. and H.Y. analysed and discussed the data; Z.J., X.D., F.B., Xiang Liu, L.G., G.X., C.Q., Z.S., H.J., Z.R. and H.Y. wrote the manuscript.

Competing interests The authors declare no competing interests.

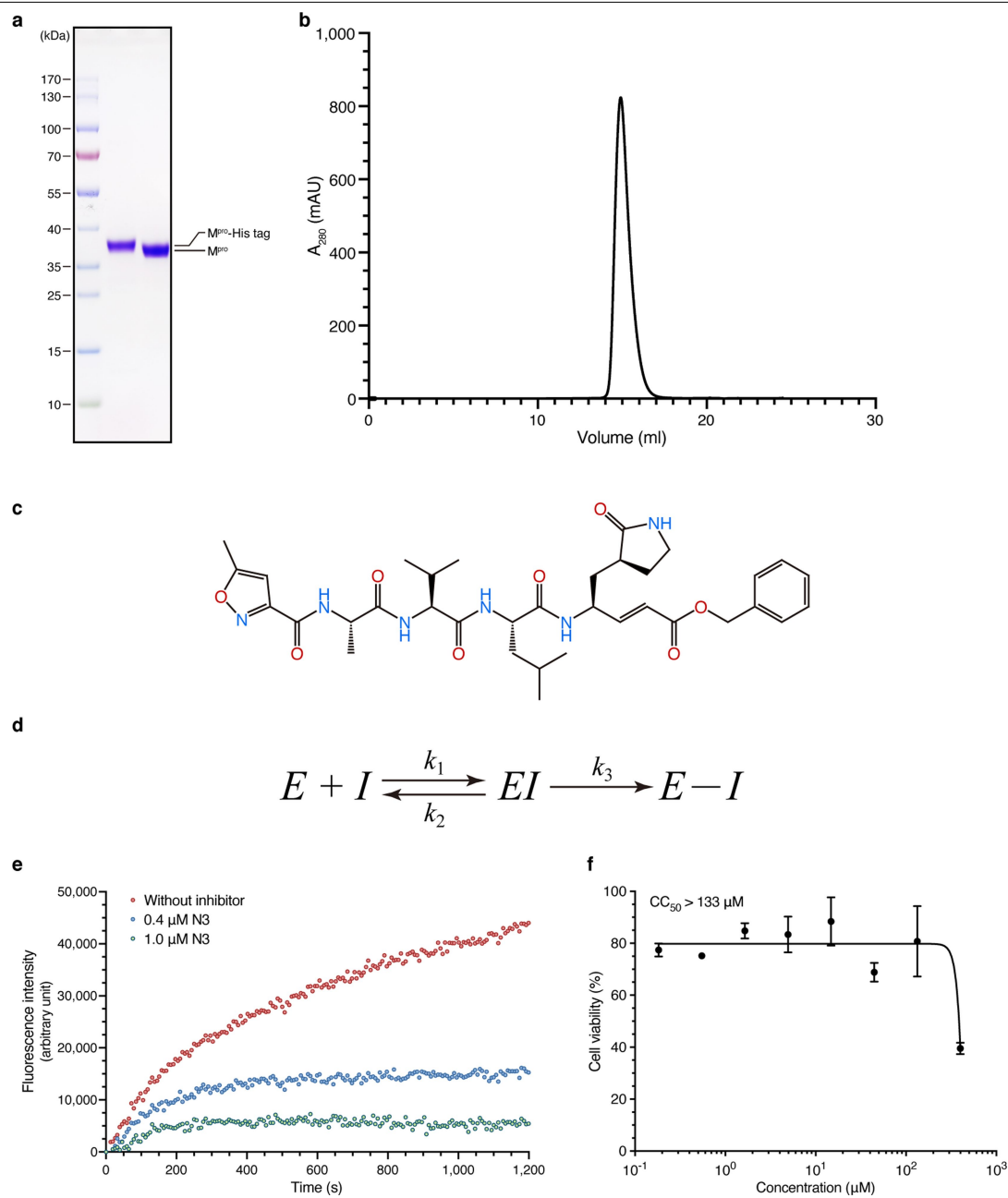
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-2223-y>.

Correspondence and requests for materials should be addressed to H.J., Z.R. or H.Y.

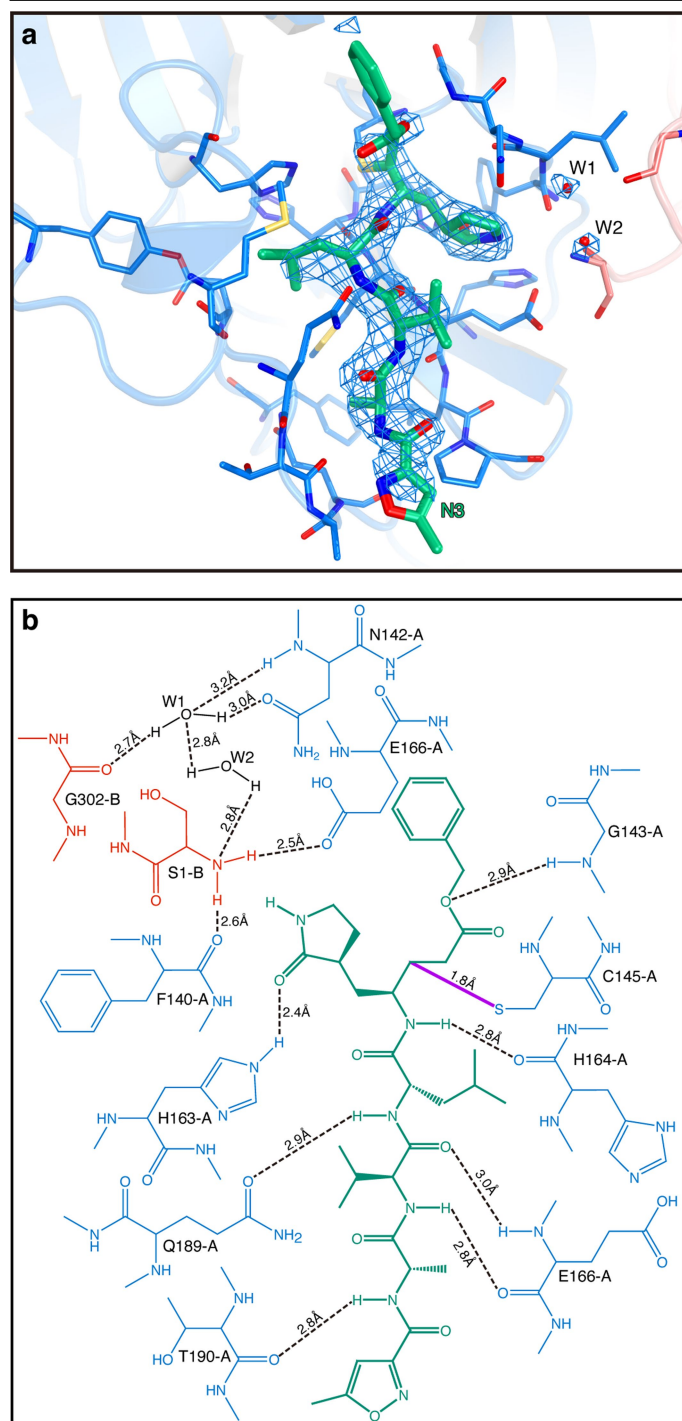
Peer review information *Nature* thanks Julien Lescar and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



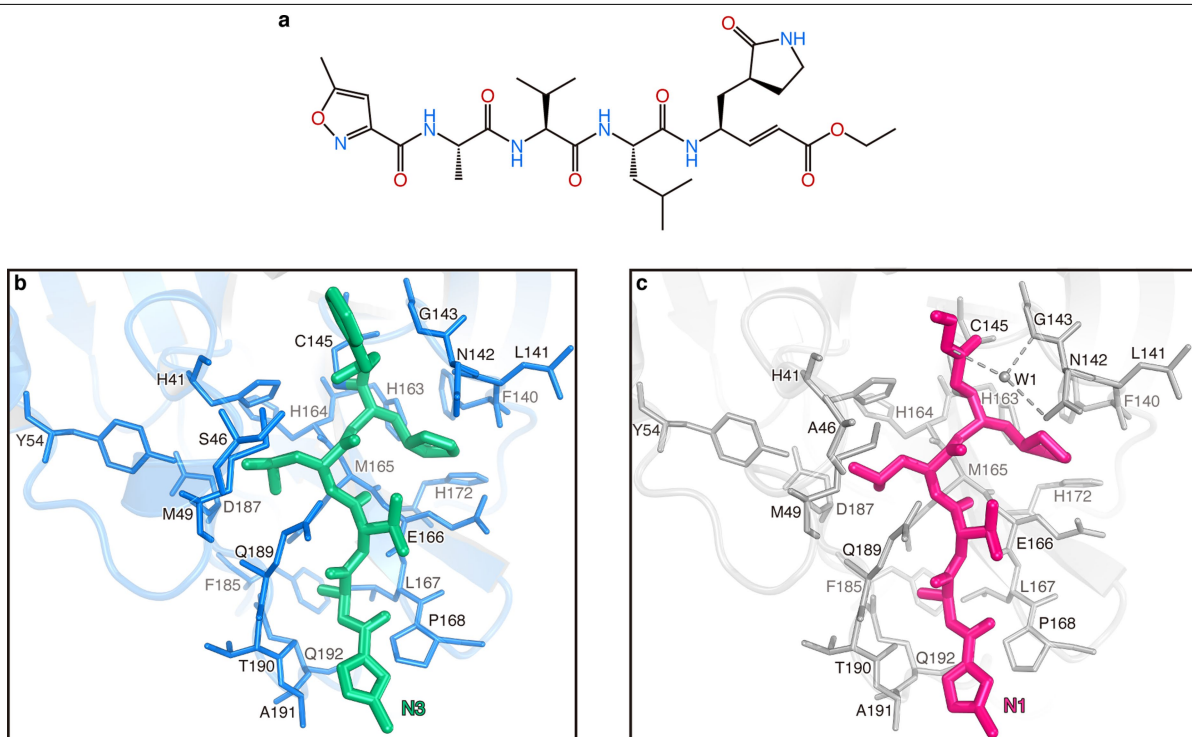
Extended Data Fig. 1 | The purification of SARS-CoV-2 M^{pro} and the inhibitory assay of the N3 compound. **a**, SDS-PAGE gel of SARS-CoV-2 M^{pro}. First lane, marker; second lane, M^{pro} before treating with rhinovirus 3C protease; third lane, M^{pro} after the cleavage of C-terminal His tag. For gel source data, see Supplementary Fig. 1. **b**, Size-exclusion chromatography profile of

M^{pro}. **c**, The chemical structure of the N3 inhibitor. **d**, Inhibition mechanism for N3. **e**, Typical inhibition curves for N3. **f**, Cytotoxicity assay of N3 on Vero cells. Data are shown as mean ± s.e.m., *n* = 3 biological replicates. The data in **a**, **b**, **e**, **f** are representative of three independent experiments with similar results.



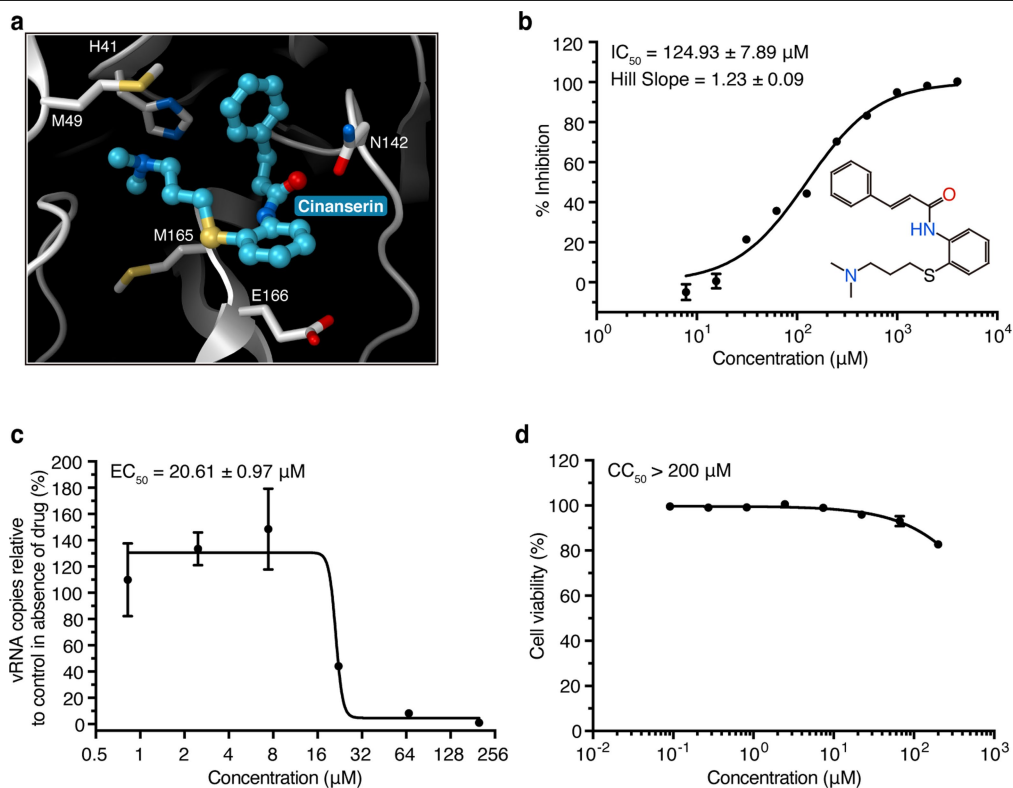
Extended Data Fig. 2 | The interactions between SARS-CoV-2 M^{pro} and N3.

a, The $F_o - F_c$ omit map (contour level of 3σ , shown as blue mesh). **b**, Detailed view of the interactions between the inhibitor N3 and SARS-CoV-2 M^{pro} . M^{pro} residues are shown in blue (protomer A) and salmon (protomer B); N3 is in green; and water is in black. The hydrogen bonds are shown as black dashed lines. The covalent bond between N3 and the C145 of protomer A is in purple.



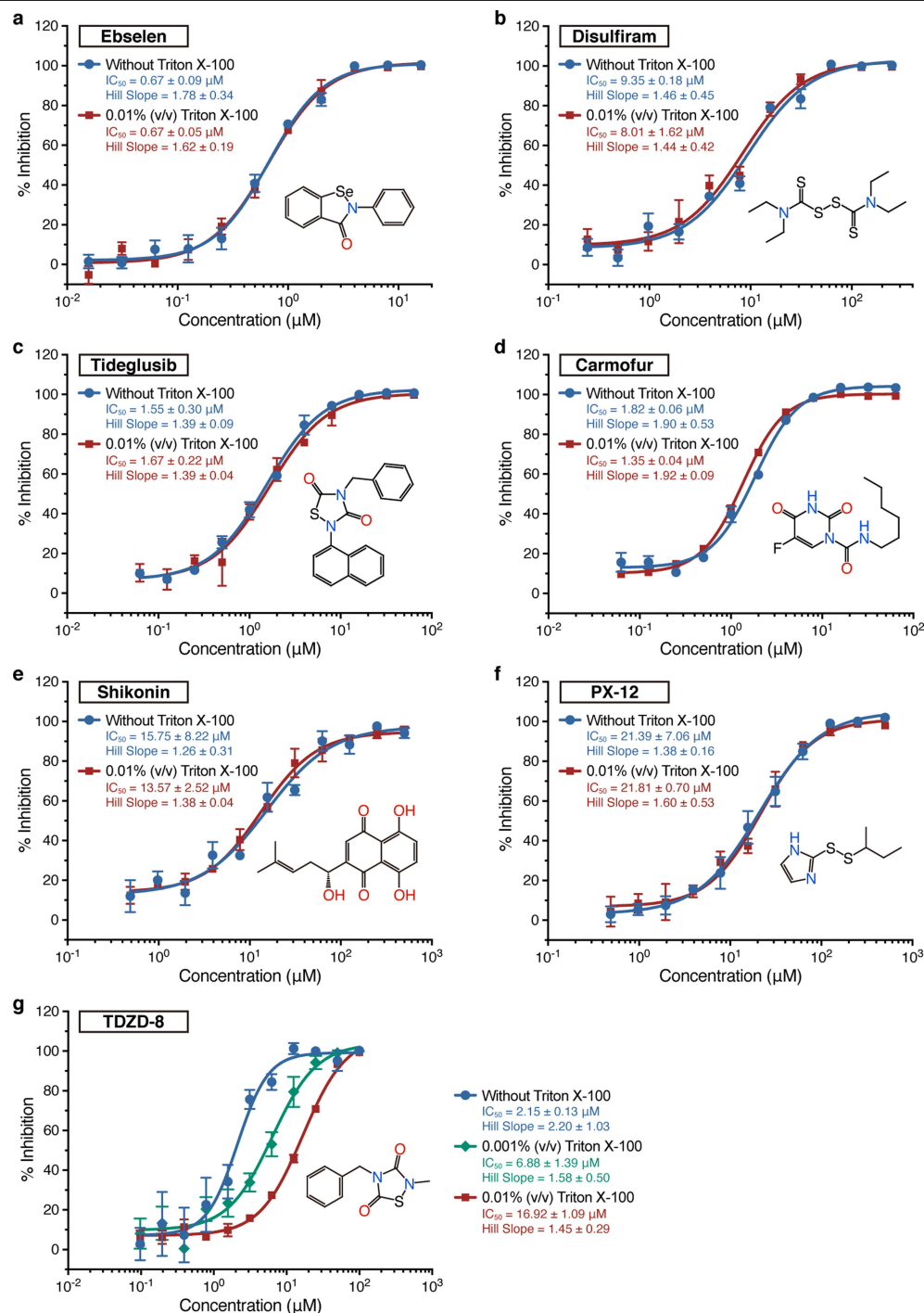
Extended Data Fig. 3 | Comparison of the binding modes between SARS-CoV-2 M^{pro}-N3 and SARS-CoV M^{pro}-N1. a, The chemical structure of the N1 inhibitor. **b**, The binding mode of SARS-CoV-2 M^{pro} with **N3**

(green sticks). **c**, The binding mode of SARS-CoV M^{pro} with **N1** (pink sticks). The hydrogen bonds formed by water (W1) are indicated by the dashed lines.



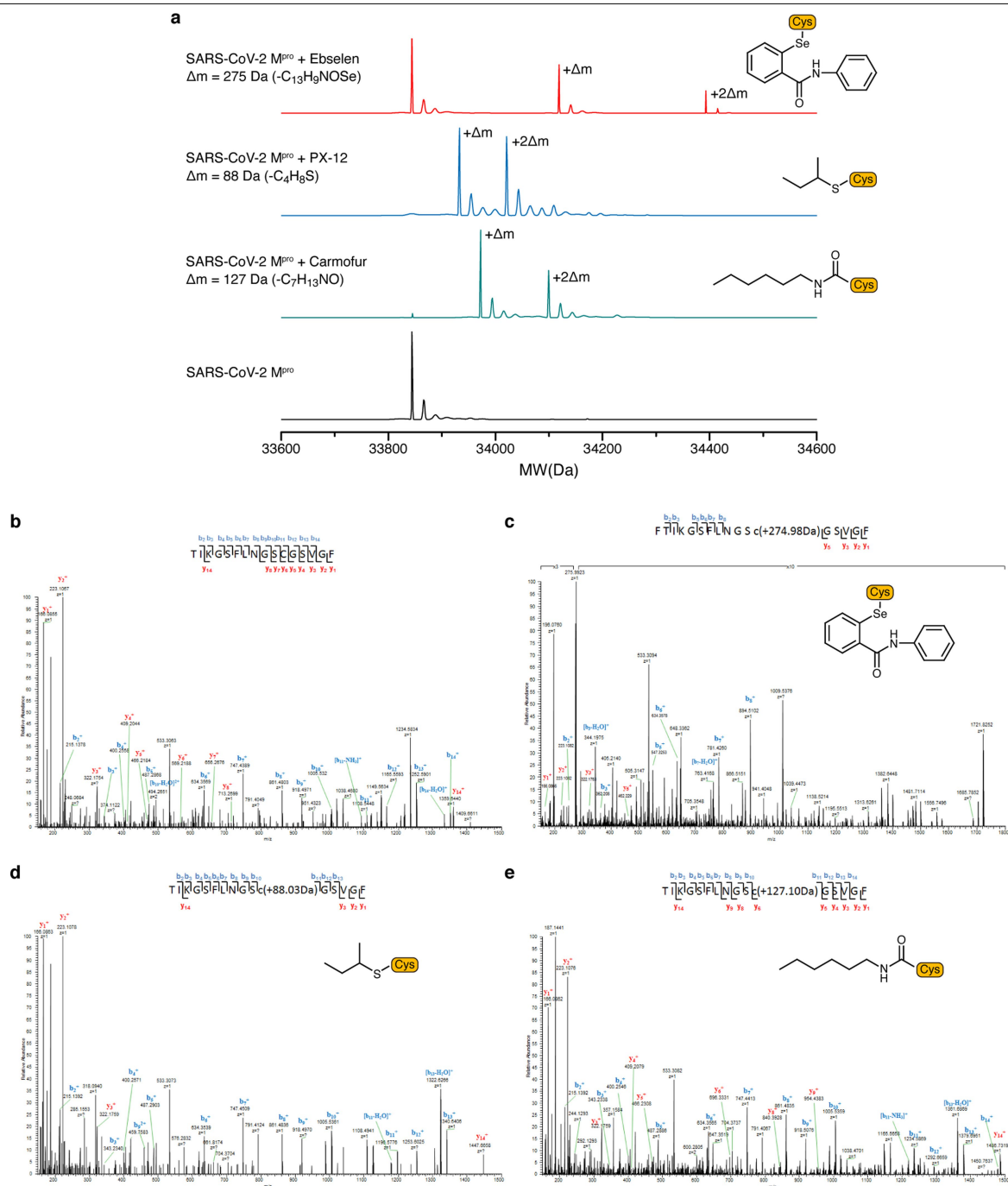
Extended Data Fig. 4 | Cinanserin is an inhibitor of SARS-CoV-2 M^{pro}. **a**, The docking result of cinanserin. The structure of SARS-CoV-2 M^{pro} is shown as a white cartoon, cinanserin is shown as cyan balls and sticks, and residues predicted to be interacting with cinanserin are shown as sticks. **b**, Inhibitory

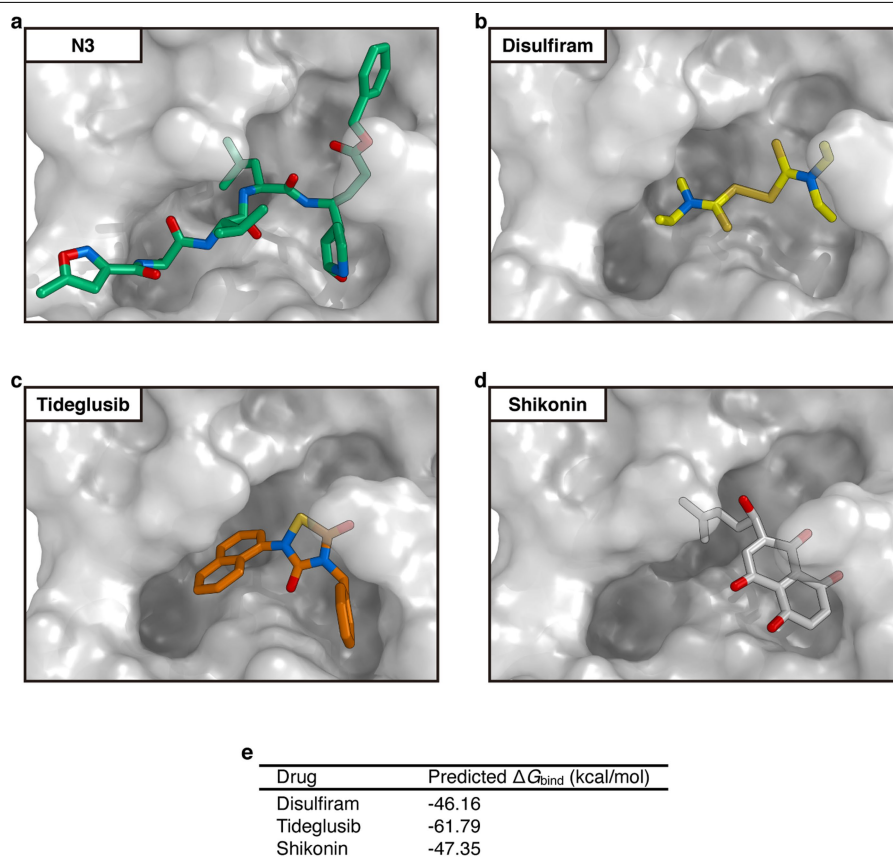
activity of cinanserin on M^{pro}. **c**, Antiviral activity of cinanserin determined by qRT-PCR. **d**, Cytotoxicity assay of cinanserin on Vero E6 cells. All data are shown as mean ± s.e.m., $n = 3$ biological replicates.



Extended Data Fig. 5 | The detergent-based assay for drug leads. a-f, The IC_{50} values determined in the presence or absence of 0.01% Triton X-100, which show that detergent did not affect the results. **g,** Different concentrations of

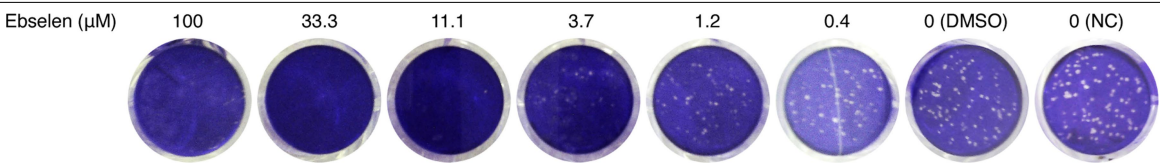
Triton X-100 notably affected IC_{50} curves for TDZD-8. All data are shown as mean \pm s.e.m., $n = 3$ biological replicates.





Extended Data Fig. 7 | Docking poses of different SARS-CoV-2 M^{pro} inhibitors. **a**, The crystal structure of the SARS-CoV-2 M^{pro}-N3 complex. **b–d**, The docking results of three drug leads. M^{pro} is shown as grey background, and inhibitors are shown in different colours. The inhibitors identified through

the high-throughput screening are likely to occupy the same pocket as N3. **e**, Predicted binding affinities for the drug leads to SARS-CoV-2 M^{pro} by using the MM-GBSA module, integrated in Schrödinger.



Extended Data Fig. 8 | Images for the plaque-reduction assay using ebselen. As the concentration of ebselen increases, there is a considerable reduction in the numbers of the plaques in comparison to the negative control (NC) and DMSO. Results are shown as representative of four biological replicates. For image source data, see Supplementary Fig. 2.

Extended Data Table 1 | Data collection and refinement statistics

| PDB code: 6LU7* | |
|---------------------------------------|-------------------------------------|
| Data collection | |
| Space group | C2 |
| Cell dimensions | |
| a, b, c (Å) | 97.931, 79.477, 51.803 |
| α, β, γ (°) | 90, 114.55, 90 |
| Resolution (Å) | 50.00-2.16 (2.22-2.16) [†] |
| R _{merge} | 18.9 (147.2) |
| I / σI | 6.3 (3.2) |
| Completeness (%) | 100.0 (100.0) |
| Redundancy | 6.6 (6.1) |
| Refinement | |
| Resolution (Å) | 50.00-2.16 |
| No. reflections | 19455 (1431) |
| R _{work} / R _{free} | 0.2020/0.2350 |
| No. atoms | |
| Protein | 2367 |
| Ligand/ion | 49 |
| Water | 84 |
| B-factors | |
| Protein | 42.7 |
| Ligand/ion | 46.3 |
| Water | 44.2 |
| R.m.s. deviations | |
| Bond lengths (Å) | 0.002 |
| Bond angles (°) | 0.474 |

*A single crystal was used for data collection and structure determination.
[†]Values in parentheses are for the highest-resolution shell.