Chapter 8

Title: The Role of Structural Biology Task Force: Validation of the Binding Mode of Repurposed Drugs Against SARS-CoV-2 Protein Targets

Focus on SARS-CoV-2 main protease (Mpro): a promising target for COVID-19 treatment

Authors:

Stefano Morasso, Protein Facility - Structural Biology Lab - Elettra Sincrotrone Trieste S.C.p.A.

Elisa Costanzi, Protein Facility - Structural Biology Lab - Elettra Sincrotrone Trieste S.C.p.A.

Nicola Demitri, XRD2 beamline - Elettra Sincrotrone Trieste S.C.p.A.

Barbara Giabbai, Protein Facility - Structural Biology Lab - Elettra Sincrotrone Trieste S.C.p.A.

Paola Storici,* Protein Facility - Structural Biology Lab - Elettra Sincrotrone Trieste S.C.p.A.

Abstract

The main protease (Mpro) of SARS-CoV-2, a cysteine protease that plays a key role in generating the active proteins essential for coronavirus replication, is a validated drug target for treating COVID-19. The structure of Mpro has been elucidated by macromolecular crystallography, but owing to its conformational flexibility, finding effective inhibitory ligands was challenging. Screening libraries of ligands as part of EXaSCale smArt pLatform Against paThogEns (ExScalate4CoV) yielded several potential drug molecules that inhibit SARS-CoV-2 replication *in vitro.* We solved the crystal structures of Mpro in complex with repurposed drugs like myricetin, a natural flavonoid, and MG-132, a synthetic peptide aldehyde. We found that both inhibitors covalently bind the catalytic cysteine. Notably, myricetin has an unexpected binding mode, showing an inverted orientation with respect to that of the flavonoid baicalein. Moreover, the crystallographic model validates the docking pose suggested by molecular dynamics experiments. The mechanism of MG-132 activity against SARS-CoV-2 Mpro was elucidated by comparison of apo and inhibitor-bound crystals, showing that regardless of the redox state of the environment and the crystalline symmetry, this inhibitor binds covalently to Cys145 with a well-preserved binding pose that extends along the whole substrate binding site. MG-132 also fits well into the catalytic pocket of human cathepsin L, as shown by computational docking, suggesting that it might represent a good start to develop dual-targeting drugs against COVID-19.

Mpro as a drug target: structural properties

The main protease (Mpro) of SARS-CoV-2, also referred to as 3-chymotrypsin-like protease (3CLpro) or nonstructural-protein 5 (nsp5), is a cysteine protease that is part of the polyproteins (pp1a and pp1ab) encoded by the viral RNA genome. It catalyzes its own excision from pp1a/pp1ab and that of 15 other mature nonstructural proteins (nsps 1-16) [1]. Mpro activity is essential to the viral replication cycle and RNA transcription. The protein is fully functional as a dimer composed of two 33.8 kDa monomers that share a large dimerization interface and arrange perpendicularly to one another, forming a characteristic heart-shaped particle. The monomer structure is composed of three domains: domain I (residues 8-101) and II (residues 102-184) arranged in a β-barrel, and domain III (residues 185-303), which folds into a five–α-helix bundle (Figure 1a, b). The active site is located on the surface, at the interface between domains I and II, and contains the noncanonical Cys145-His41 catalytic dyad (Figure 1c). A conserved water molecule located in proximity to His41 is also important for catalysis [2]. Domain III largely contributes to dimer formation and is critical for enzyme activity. The dimer formation is functional to activation, whereas the single monomers are mostly inactive [3,4]. These structural features are shared with the main proteases of other coronaviruses [5]. The highly conserved structure and the low homology with human proteins make it a recognized target for the development of therapeutics against COVID-19, as well as against other coronavirus infections.

Within the EXSCALATE4CoV (E4C) consortium, the structural biology group of Elettra Synchrotron (www.elettra.eu) committed to validating the binding modes of the compounds selected through virtual and repurposing screenings [6,7]. In this framework, we first determined the crystal structure of the apoprotein in different space groups. Initial crystals were obtained from the PACT commercial screening, reproducing the conditions published by Zhang et al. [3]. Thus, we solved the structures of Mpro in the apo form at 1.65 Å resolution in the space groups C 2 (Protein Data Bank [PDB] identifier [ID]: 7ALH) and P 2_1 (PDB ID: 7ALI). These space groups are very frequent among the Mpro structures deposited in the PDB. In these space groups, the main protease always organizes as a dimer, with the 2-fold axis being either crystallographic (C 2) or noncrystallographic (P 2_1). Subsequently, we obtained apo crystals by seeding techniques in space group P $2_12_12_1$, which contains the whole dimer in the asymmetric unit (PDB ID: 7BB2). Among the three apo structures, the fold of the dimer is conserved, with minor differences in local regions that adopt slightly diverse conformations. In particular, the two loops ASL1 and ASL2 (residues 44-53 and 184-194, respectively) delimiting the active site show differences between the three structures, suggesting that Mpro has high plasticity for adapting to different substrates, as reported in the literature by us and others [2,6,8]. Using the seeding technique, we co-crystallized Mpro in a complex with various ligands. The main effort was dedicated to clarifying the binding mode of repurposed drugs selected within the E4C consortium.

Between February 2020, when the first SARS-CoV-2 Mpro structure was released in the PDB [9], and November 2022, almost 650 structures were solved by x-ray crystallography and made available to the scientific community, even before publication in several cases. This momentous work that incorporated co-crystals with repurposed drugs, as well as other small molecules

including natural compounds and chemical fragments, has largely contributed to the computer-aided drug design campaigns [10] of E4C.

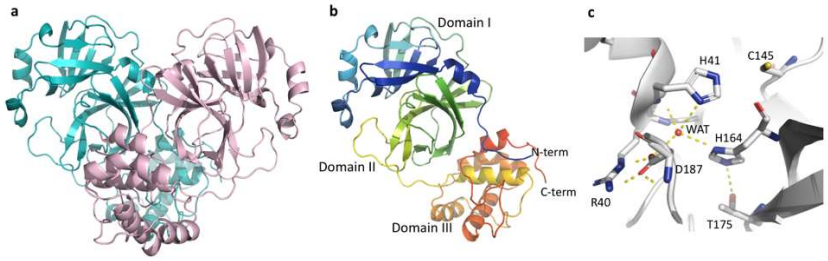

Figure 1. Structure of SARS-CoV-2 Mpro (PDB ID: 7BB2). (a) The enzyme is organized as a homodimer (A and B chains that are colored cyan and pink, respectively). (b) Each protomer consists of three domains (I−III): The chymotrypsin-like and picornavirus 3C protease−like domains I and II (in blue and green, respectively) harbor the active site. Domain III (in yellow-red) forms a five-helix bundle and is involved in the dimerization of the protein. (c) Close-up of the active site and of the hydrogen bond network. Atoms are in stick representation colored according to atom type, while hydrogen bonds are shown as dashed lines.

Known inhibitors of Mpro bind into the active site

Druggability of SARS-CoV-2 Mpro is confirmed by the thousands of small-molecule inhibitors that have been identified throughout screening campaigns performed as a global effort to fight COVID-19 [11,12]. Joint programs of virtual screening together with biochemical-based high-throughput screening have evaluated broad lists of chemicals, including natural compounds, repurposed drugs, and novel entities.

The first actions for fast development of drugs against SARS-CoV-2 pointed to the screening of inhibitors derived from previous research on the main protease from other coronaviruses such as SARS-CoV and MERS [5,13]. The resulting compounds had limited potency in enzymatic assays, despite the high sequence homology among the three viruses. Notably, the sequences of Mpro from SARS-CoV2 and SARS-CoV differ by only 12 amino acids, and only 1 in the active site. This difference suggests that these residues, though distant from the binding site, contribute to enzyme plasticity and ligand binding via allosteric regulation [6].

These findings led to the hypothesis that the complexity and variety of Mpro conformational changes and interactions with ligands would make fast drug identification challenging [14]. In this respect, macromolecular crystallography has largely contributed to protein-ligand models useful for accelerating the drug-discovery process.

Interestingly, although some allosteric binding sites have been identified, including a few that sit at the dimer interface [11,15], most of the inhibitors crystallized with Mpro bind into the active site [11,16]. The enzyme's active-site cavity reveals a high degree of malleability, allowing a variety of different chemical moieties to bind and inhibit SARS-CoV-2 3CL Mpro. The

compounds identified are covalent or noncovalent inhibitors. Among the covalent, a substantial number are peptidomimetics or peptidomimetic-derived molecules that mount different warheads that react with Cys145. Indeed, the first Mpro inhibitor authorized in many countries for COVID-19 treatment was PF-07321332 (nirmatrelvir) [17]. It is an orally available peptidomimetic developed by Pfizer from a series of molecules active against pan-corona Mpro. Nirmatrelvir was formulated in combination with ritonavir and branded as Paxlovid. Beyond nirmatrelvir, other potent compounds targeting Mpro are under development that derive from chemical scaffolds or natural origin [18-22].

Myricetin binds covalently with Cys145 in the MPro active site

Flavonoids account for a large and important group of natural products widely observed in plants. They are polyphenolic secondary metabolites, and, owing to a combination of biochemical and antioxidant effects, they are considered beneficial for various diseases such as cancer, Alzheimer disease, and atherosclerosis. Interestingly, flavonoids have also been reported to have antiviral activity [23]. Specifically, flavonoids of natural or synthetic origins have been proposed to target SARS-CoV-2 [24,25].

Within the E4C initiative, a repurposing biochemical screening was performed that involved 8700 compounds containing marketed drugs, clinical and preclinical candidates, and small molecules regarded as safe in humans [7]. Among the 256 hits, some flavonoids were confirmed to inhibit Mpro with IC50s ranging from 3.6 to 0.18 μM. In this context, myricetin showed an IC50 of 220 nM in the biochemical assay; interestingly, it was also predicted as a nanomolar-level binder from a docking-based virtual screening performed on a collection of 30,000 compounds [6]. Given the interest in this molecule, which is largely present in several edible plants and a key ingredient of various foods and beverages, we determined the x-ray co-crystal structure of myricetin in complex with Mpro at 1.77Å resolution (Figure 2a, b; PDB ID: 7B3E) [7]. Unexpectedly, myricetin formed a covalent bond between the Cys145 sulfur and the 2′ position of the flavonoid, leading to unprecedented binding for a flavonoid scaffold. At that time, the only x-ray structure of SARS-CoV-2 Mpro in complex with a flavonoid had been obtained with baicalein (PDB ID: 6M2N), which was modeled in reverse orientation, with the chromone moiety occupying the S1 subpocket and no possibility to form a covalent adduct with Cys145 [26].

Consistent with the 7B3E crystallographic structure, the *in silico* noncovalent docking calculations led myricetin to adopt a similar pose (Figure 2d) [6]. Moreover, the electronic map of the x-ray structure showed that the Mpro binding pocket is only partially occupied by myricetin and that voids are filled by solvent molecules (ethylene glycol and water), suggesting an opportunity for future structure-based drug design efforts. The same conclusions have been subsequently reported by Su and colleagues [27], who proposed pyrogallol as a convenient warhead in designing new flavone-based covalent inhibitors of Mpro. Indeed, as suggested by Kuzikov et al. [7], the pyrogallol reactivity requires an oxidative step for the sulfur addition to the 2′ position of myricetin, as shown in Figure 2c. The unexpected binding of myricetin to Cys145 opens new routes for the development of more potent covalent ligands that are of great interest

for therapeutics and biochemical tools.

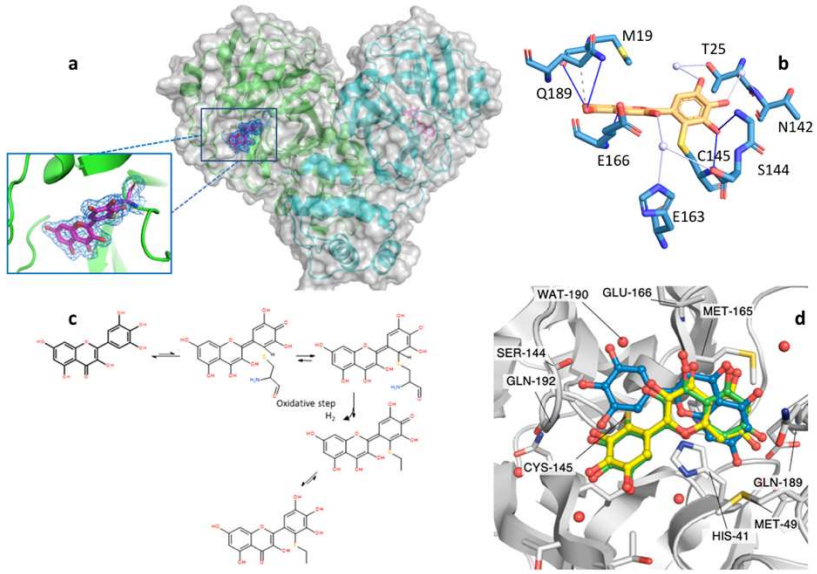

Figure 2. Myricetin binds to Cys145 in the Mpro active site (PDB ID: 7B3E). (a) Overall structure of Mpro dimer in complex with myricetin, which occupies both active sites: the dimer is shown in a cartoon model (chain A colored in light green, chain B colored in light blue) superimposed on the surface (white); the myricetin is shown in magenta stick-bone and superimposed by fo-fc map contoured at 1 sigma. (b) Interaction of myricetin (yellow) with active-site residues (blue) and water molecules (white). Hydrogen bonds are shown with blue lines, water bridges in light blue, hydrophobic interaction with a dotted gray line. (c) Mechanism of myricetin oxidation and Michael addition to Cys145. (d) Overlay of crystal structure (green), docked (blue, RMSD 3.14 Å), and refined (yellow, RMSD 0.46 Å) binding poses of myricetin. The image in panel c was adapted from Kuzikov et al [7], and the image in panel d was adapted from Gossen et al [6].

The peptidomimetic MG-132 acts as dual inhibitor of Mpro and cathepsin L

The same screening of Mpro ligands against SARS-CoV-2 performed as part of the E4C program and mentioned in the previous chapter led to the identification of MG-132 (Figure 3b). This synthetic peptidomimetic aldehyde, originally identified as a proteasome inhibitor and investigated as an antineoplastic drug, blocks SARS-CoV2 Mpro enzymatic activity with an IC50 of 7.4 μM and shows good antiviral activity, detected as reduction of viral RNA (EC50 of 0.1 μM) in cells infected with SARS-CoV-2.

We solved the crystal structure of Mpro in complex with MG-132 under different crystallization conditions, showing that in the presence and absence of reducing agents and independently from the space organization of the crystals, this compound attaches covalently to Cys145 through a Michael addition, and it has a well-defined binding mode that does not alter the overall

fold of the dimer [28]. The covalent stereoselective (S) hemithioacetal bond is nicely defined in the electron-density maps of our structures (Figure 3b). MG-132 extends along the S1–S4 subsites of the substrate binding pocket, interacting with residues through hydrogen bonds and hydrophobic interactions, in addition to the covalent bond with Cys145 (Figure 3c). An extensive biochemical analysis revealed that this bond is reversible, as expected, but K_m and V_{max} measured at different incubation times suggest a slow k_{off}, indicative of a long residency time.

Considering that MG-132 is known to inhibit other cysteine proteases and that, despite its rather poor inhibition of Mpro, it has sub-micromolar potency in antiviral cell assays [29,30], we investigated the role of MG-132 in inhibiting the human cathepsin L, a host protease that is important for viral entry. This lysosomal cysteine protease is proposed as a target for COVID-19, as it cleaves the viral S protein to promote entry of the virus into host cells [31,32]. MG-132 is known to inhibit cathepsin L in the nanomolar range [28,33]. Induced-fit docking and covalent docking models of MG-132 bound to cathepsin L show that the ligand can form a covalent linkage with Cys26 and embrace the active site by numerous hydrogen bonds and pi-staking interaction, as shown in Figure 3d [28]. This analysis provides new hints for the development of Mpro/cathepsin L dual inhibitors that may prove beneficial against COVID-19, increasing efficacy and reducing the threat of drug resistance.

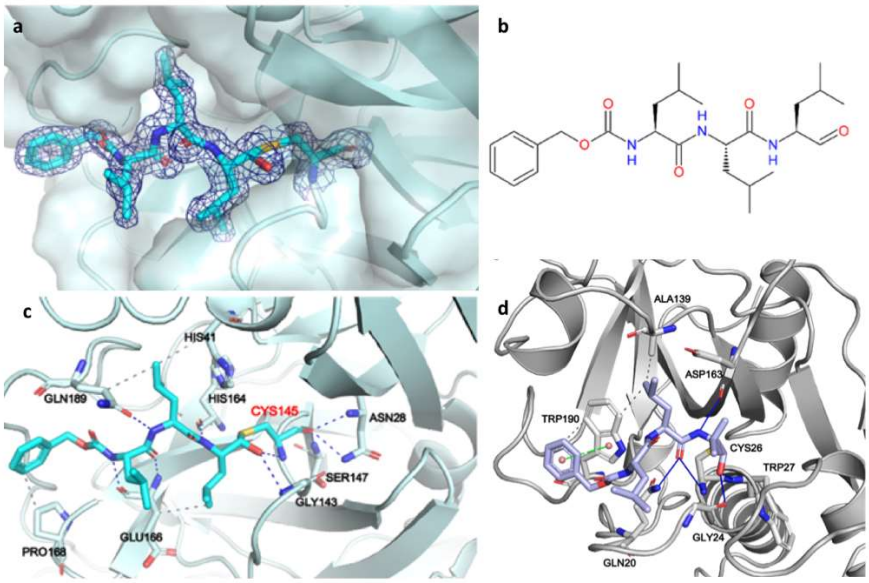

Figure 3. Binding mode of MG-132 in Mpro active site. (a) Electron-density map (fo-fc, contoured at 1 sigma) of MG-132 covalently bound to Cys145. (b) Chemical model of MG-132. (c) Main interactions of MG-132 bound to Cys145 with Mpro active-site residues (ligand and residues are represented as stick-bones, MG-132 and Cys145 are colored in cyan, and other residues and cartoon protein model are in light blue). d) Best scoring pose obtained from the covalent docking of MG-132 on cathepsin L. The main interactions of MG-132 are shown: hydrogen bonds (blue lines), hydrophobic interactions (gray dashed lines), and pi-stacking interactions (green dashed lines). Figure adapted from Costanzi et al [28].

References

1. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol*. 2021;19(3):155-170.

2. Kneller DW, Phillips G, O'Neill HM, Jedrzejczak R, Stols L, Langan P, et al. Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat Commun*. 2020;11(1):3202.

3. Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science*. 2020;368(6489):409-412.

4. Goyal B, Goyal D. Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy. *ACS Comb Sci*. 2020;22(6):297-305.

5. Amin SA, Banerjee S, Gayen S, Jha T. Protease targeted COVID-19 drug discovery: what we have learned from the past SARS-CoV inhibitors? *Eur J Med Chem*. 2021;215:113294.

6. Gossen J, Albani S, Hanke A, Joseph BP, Bergh C, Kuzikov M, et al. A blueprint for high affinity SARS-CoV-2 Mpro inhibitors from activity-based compound library screening guided by analysis of protein dynamics. *ACS Pharmacol Transl Sci*. 2021;4(3):1079-1095.

7. Kuzikov M, Costanzi E, Reinshagen J, Esposito F, Vangeel L, Wolf M, et al. Identification of inhibitors of SARS-CoV-2 3CL-pro enzymatic activity using a small molecule in vitro repurposing screen. *ACS Pharmacol Transl Sci*. 2021;4(3):1096-1110.

8. Ebrahim A, Riley BT, Kumaran D, Andi B, Fuchs MR, McSweeney S, et al. The temperature-dependent conformational ensemble of SARS-CoV-2 main protease (M^{pro}). *IUCrJ*. 2022;9(5).

9. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature*. 2020;582(7811):289-293.

10. Liu Y, Gan J, Wang R, Yang X, Xiao Z, Cao Y. DrugDevCovid19: an atlas of anti-COVID-19 compounds derived by computer-aided drug design. *Molecules.* 2022;27(3):683.

11. Macip G, Garcia-Segura P, Mestres-Truyol J, Saldivar-Espinoza B, Pujadas G, Garcia-Vallvé S. A review of the current landscape of SARS-CoV-2 main protease inhibitors: have we hit the bullseye yet? *Int J Mol Sci*. 2021;23(1):259.

12. Maghsoudi S, Taghavi Shahraki B, Rameh F, Nazarabi M, Fatahi Y, Akhavan O, et al. A review on computer-aided chemogenomics and drug repositioning for rational COVID-19 drug discovery. *Chem Biol Drug Des*. 2022;100(5):699-721.

13. Yang H, Yang J. A review of the latest research on M^{pro} targeting SARS-COV inhibitors. *RSC Med Chem*. 2021;12(7):1026.

14. Bzówka M, Mitusińska K, Raczyńska A, Samol A, Tuszyński JA, Góra A. Structural and evolutionary analysis indicate that the SARS-CoV-2 Mpro is a challenging target for small-molecule inhibitor design. *Int J Mol Sci*. 2020;21(9):3099.

15. Gunther S, Reinke PYA, Fernandez-Garcia Y, Lieske J, Lane TJ, Ginn HM, et al. X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. *Science*. 2021;372(6542):642-646.

16. Nguyen DD, Gao K, Chen J, Wang R, Wei G-W. Unveiling the molecular mechanism of SARS-CoV-2 main protease inhibition from 137 crystal structures using algebraic topology and deep learning. *Chem Sci*. 2020;11(44):12036-12046.

17. Owen DR, Allerton CMN, Anderson AS, Aschenbrenner L, Avery M, Berritt S, et al. An oral SARS-CoV-2 M^{pro} inhibitor clinical candidate for the treatment of COVID-19. *Science*. 2021;374(6575):1586-1593.

18. Raman K, Rajagopal K, Islam F, Dhawan M, Mitra S, Aparna B, et al. Role of natural products towards the SARS-CoV-2: a critical review. *Ann Med Surg*. 2022:104062.

19. Nepali K, Sharma R, Sharma S, Thakur A, Liou J-P. Beyond the vaccines: a glance at the small molecule and peptide-based anti-COVID19 arsenal. *J Biomed Sci*. 2022;29(1):65.

20. Zhong L, Zhao Z, Peng X, Zou J, Yang S. Recent advances in small-molecular therapeutics for COVID-19. *Precis Clin Med*. 2022;5(4):pbac024.

21. Mousavi S, Zare S, Mirzaei M, Feizi A. Novel drug design for treatment of COVID-19: a systematic review of preclinical studies. *Can J Infect Dis Med Microbiol*. 2022;2022:2044282.

22. Ton A-T, Pandey M, Smith JR, Ban F, Fernandez M, Cherkasov A. Targeting SARS-CoV-2 papain-like protease in the post-vaccine era. *Trends Pharmacol Sci*. 2022;43(11):906-919.

23. Badshah SL, Faisal S, Muhammad A, Poulson BG, Emwas AH, Jaremko M. Antiviral activities of flavonoids. *Biomed Pharmacother*. 2021;140:111596.

24. Jo S, Kim S, Shin DH, Kim M-S. Inhibition of SARS-CoV 3CL protease by flavonoids. *J Enzyme Inhib Med Chem*. 2020;35(1):145-151.

25. Batool F, Mughal EU, Zia K, Sadiq A, Naeem N, Javid A, et al. Synthetic flavonoids as potential antiviral agents against SARS-CoV-2 main protease. *J Biomol Struct Dyn*. 2022;40(8):3777-3788.

26. Su H, Yao S, Zhao W, Li M, Liu J, Shang W, et al. Discovery of baicalin and baicalein as

novel, natural product inhibitors of SARS-CoV-2 3CL protease in vitro. *BioRxiv*. 2020.04.13.038687; doi: https://doi.org/10.1101/2020.04.13.038687.

27. Su H, Yao S, Zhao W, Zhang Y, Liu J, Shao Q, et al. Identification of pyrogallol as a warhead in design of covalent inhibitors for the SARS-CoV-2 3CL protease. *Nat Commun*. 2021;12(1):3623.

28. Costanzi E, Kuzikov M, Esposito F, Albani S, Demitri N, Giabbai B, et al. Structural and biochemical analysis of the dual inhibition of MG-132 against SARS-CoV-2 main protease (Mpro/3CLpro) and human cathepsin-L. *Int J Mol Sci*. 2021;22(21):11779.

29. Dittmar M, Lee JS, Whig K, Segrist E, Li M, Jurado K, et al. Drug repurposing screens reveal FDA approved drugs active against SARS-Cov-2. *Cell Rep.* 2021;35(1):108959.

30. Zaliani A, Vangeel L, Reinshagen J, Iaconis D, Kuzikov M, Keminer O, et al. Cytopathic SARS-CoV-2 screening on VERO-E6 cells in a large-scale repurposing effort. *Sci Data*. 2022;9(1):405.

31. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, et al. Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A*. 2020;117(21):11727-11734.

32. Zhao M-M, Yang W-L, Yang F-Y, Zhang L, Huang W-J, Hou W, et al. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduct Target Ther*. 2021;6(1):134.

33. Ito H, Watanabe M, Kim Y-T, Takahashi K. Inhibition of rat liver cathepsins B and L by the peptide aldehyde benzyloxycarbonyl-leucyl-leucyl-leucinal and its analogues. *J Enzyme Inhib Med Chem*. 2009;24(1):279-286.