A Novel Y-Shaped, S-O-N-O-S-Bridged Crosslink between Three Residues C22, C44, and K61 Is a Redox Switch of the SARS-CoV-2 Main Protease

Kai S. Yang[a], Syuan-Ting Alex Kuo[b], Lauren R. Blankenship[a], Yan J. Sheng[b], Banumathi Sankaran[c], Pingwei Li[d], Carol A. Fierke[e], David H. Russell[b], Xin Yan[b], Shiqing Xu[a, *], and Wenshe Ray Liu[a, d, f, g, *]

[a]Texas A&M Drug Discovery Laboratory, Department of Chemistry, Texas A&M University, College Station, TX 77843, USA

[b]Department of Chemistry, Texas A&M University, College Station, TX 77843, USA

[c]Molecular Biophysics and Integrated Bioimaging, Berkeley Center for Structural Biology, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[d]Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA

[e]Department of Biochemistry, Brandeis University, Waltham, MA 02453, USA

[f]Institute of Biosciences and Technology and Department of Translational Medical Sciences, College of Medicine, Houston, TX 77030, USA

[g]Department of Molecular and Cellular Medicine, Texas A&M University, College Station, TX 77843, USA

*Correspondence should be addressed to Shiqing Xu: shiqing.xu@tamu.edu and Wenshe Ray Liu: wsliu2007@tamu.edu

ABSTRACT

As the COVID-19 pathogen, SARS-CoV-2 relies on its main protease (M^{Pro}) for pathogenesis and replication. During the crystallographic analyses of M^{Pro} crystals that were exposed to the air, a uniquely Y-shaped, S-O-N-O-S-bridged posttranslational crosslink that connects three residues C22, C44, and K61 at their side chains was frequently observed. As a novel posttranslational modification, this crosslink serves as a redox switch to regulate the catalytic activity of M^{Pro}, a demonstrated drug target of COVID-19. The formation of this linkage leads to a much more opened active site that can be potentially targeted for the development of novel SARS-CoV-2 antivirals. The inactivation of M^{Pro} by this crosslink indicates that small molecules that lock M^{Pro} in the crosslinked form can be potentially used with other active site-targeting molecules such as paxlovid for synergistic effects in inhibiting the SARS-CoV-2 viral replication. Therefore, this new finding reveals a unique aspect of the SARS-CoV-2 pathogenesis and is potentially paradigm-shifting in our current understanding of the function of M^{Pro} and the development of its inhibitors as COVID-19 antivirals.

SARS-CoV-2 is the viral pathogen of COVID-19 that has ravaged the whole world for more than two years. Effective vaccines that target the membrane Spike protein of SARS-CoV-2 have been developed and widely adopted for human immunization.[1] However, the continuous emergence of new SARS-CoV-2 variants with mutations at Spike has led to viral evasion of vaccines and consequently infection surges.[2] The situation has called for the search for SARS-CoV-2 antivirals and led to the recent success of the development of Pfizer's paxlovid.[3,4] SARS-CoV-2 is an RNA virus with a positive-sense RNA genome. It contains a big 20-kb open reading frame ORF1ab that is translated alternatively to form two large polypeptides pp1a and pp1ab in infected human cells.[5,6] These two polypeptides need to be hydrolyzed to form 16 nonstructural proteins (nsps) that are key to viral biology including the formation of the RNA-dependent RNA polymerase complex for the replication of the viral genome and subgenomic RNAs. The maturation of pp1a and pp1ab is catalyzed by two internal nsp fragments. One of them is main protease (M^{Pro}) that processes 13 out of the total 16 nsps.[7] Paxlovid is a combination therapy with two chemical components. One component is nirmatrelvir that potently inhibits M^{Pro} to prevent viral replication in infected human cells.[4]

As an essential enzyme for SARS-CoV-2, M^{Pro} is an established target for the development of SARS-CoV-2 antivirals.[8-12] Many academic research groups and a number of pharmaceutical companies have been working on the development of M^{Pro} inhibitors.[13-24] A large international COVID Moonshot project for the development of M^{Pro} inhibitors has also been organized.[25] As a key tool for the structure-based drug discovery, X-ray protein crystallography has been frequently employed to determine structures of M^{Pro}-inhibitor complexes to assist further inhibitor optimization. This is also the approach that we have adopted for the development of M^{Pro} inhibitors.[26] To facilitate a quick determination of M^{Pro}-inhibitor complexes, we have been crystalizing the apo form of M^{Pro} and then soaking the crystals with different inhibitors for the ensuing X-ray protein crystallographic analysis.[26-28] Our obtained apo-M^{Pro} crystals have a C121 or I121 space group that contains an M^{Pro} dimer or monomer, respectively, in an asymmetric unit. In the I121 space group, two monomers from two asymmetric units form a tightly bound dimer. In both C121 and I121 space groups, a tightly bound dimer interacts with other M^{Pro} dimers within the crystals at two regions, aa53-60 and aa216-222. The aa216-222 region is located

in the noncatalytic *C*-terminal domain. Although the aa53-60 region is within the catalytic *N*-terminal domain, it is distant to the active site (Figure 1A).[26] The M^{Pro} dimer in these crystals has an open active site whose structure rearrangements are not expected to be limited by the protein packaging in crystals (Supplementary Figure 1). Therefore, it is optimal for a soaking-based structural determination of M^{Pro}-inhibitor complexes. This relatively open form of M^{Pro} is considered more desired than a closed form in representing M^{Pro} in solutions as well since M^{Pro} dimers in solutions are not expected to interact with each other. Therefore, we have been using these apo-M^{Pro} crystals and soaking them with different inhibitors for the determination of crystal structures of more than 30 M^{Pro}-inhibitor complexes. One unique observation that we have made for almost all of our determined M^{Pro}-inhibitor structures was a poorly defined conformation for the aa46-50 region. This region is part of the M^{Pro} active site, serving as a cap of the S2 site to bind the P2 residue in a peptide substrate (Figure 1A).[26,29,30] The electron density of this region was very weak in all determined M^{Pro}-inhibitor structures indicating a highly flexible conformation (Figure 1B and Supplementary Figure 2). There have been a number of M^{Pro}-inhibitor complexes whose crystal structures have been determined and deposited into the
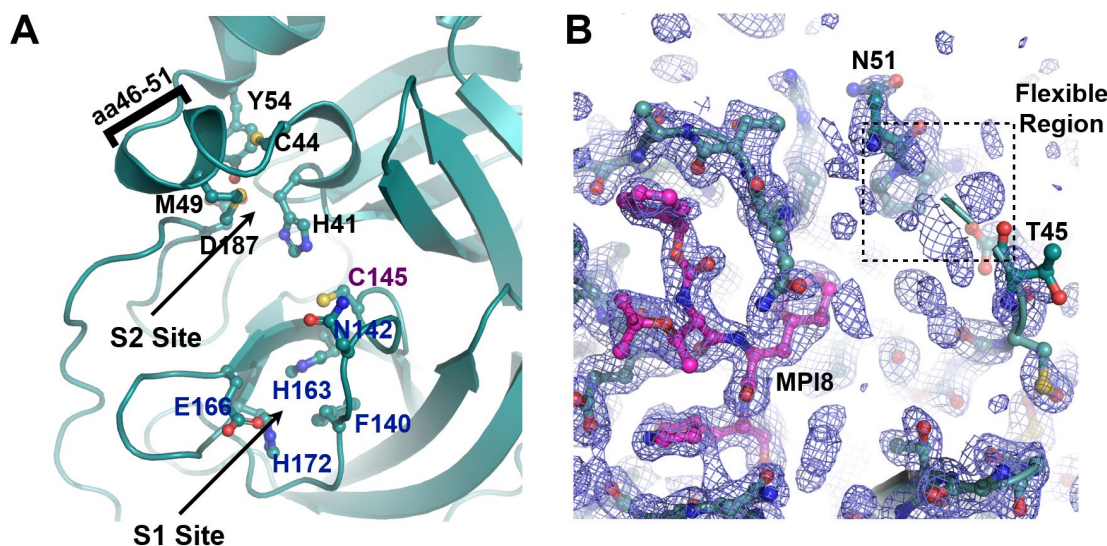

Figure 1. (A) The M^{Pro} active site based on the PDB entry 7JPY. C145 is the catalytic cysteine. Residues that form the binding sites for P1 and P2 residues in a substrate are labeled. **(B)** The low electron density map around aa46-50 in a representative M^{Pro}-inhibitor complex. The ligand MPI8 that is covalently bound to C145 is shown in a color code in which carbon atoms are shown in red. For the protein, carbon atoms are shown in light teal. The aa46-50 flexible region is indicated in a dashed square. The electron density was contoured at 1σ.

Protein Data Bank.[8-24] Many have a clearly defined conformation for aa46-50. However, crystals for most of these structures either had a closed form around the active site due to the protein packaging in crystals or were obtained by co-crystallization with ligands. We reason that in the closed form of the M^{Pro} crystals the protein packaging limits the structure rearrangement of aa46-50 and prevents it from adopting a flexible conformation. However, in solutions this limitation does not apply. This structural rearrangement is apparently related to the soaking process since the apo-M^{Pro} crystals that did not undergo soaking show a well-defined confirmation for aa46-50 and M^{Pro}-inhibitor complexes in co-crystalized crystals did not show this rearrangement either.[8-10,26]

M^{Pro}-MPI8 was an early determined M^{Pro}-inhibitor complex structure.[26] In this crystal structure, besides a structurally undefined aa46-50 region, we noticed the uniquely connected electron density at residues C22, C44 and K61. As shown in Figure 2A, 2Fo-Fc electron density contoured at 1σ around the ends of side chains from all three residues clearly merged indicating that the three residues were covalently connected to form a Y-shaped tri-residue crosslink. Directly connecting three residues at their terminal heteroatoms to form a Y-shaped S-N-S crosslink did not fit the calculated electron density. Although there was no literature report about a Y-shaped crosslink between one lysine and two cysteine residues in a protein, a recent *Nature* article by Wensien *et al.* described a novel posttranslational N-O-S-bridged lysine-cysteine crosslink as a redox switch for the transaldolase enzyme from *Neisseria gonorrhoeae*.[31] *N. gonorrhoeae* transaldolase undergoes oxidation to form an N-O-S crosslink in which an oxygen atom covalently bridges the K8 side chain amine and C38 side chain thiol. The consequential oxidized enzyme has a substantially lower activity than the reduced form. As a newly identified posttranslational modification, the N-O-S crosslink via oxidation serves as a regulatory mechanism for *N. gonorrhoeae* transaldolase. Since in M^{Pro}-MPI8, the K61 side chain amine connected with two side chain thiol groups from C22 and C44, respectively, to form a Y-shaped crosslink instead of a linear S-O-N crosslink as observe between K8 and C38 in *N. gonorrhoeae* transaldolase, we suspected that a S-O-N-O-S-bridged, Y-shaped tri-residue crosslink in which two oxygen atoms bridged two cysteine thiol groups separately with the K61 side chain amine existed in our M^{Pro}-MPI8 structure. Adding this crosslink

into the structure for undergoing the structure refinement fit the calculated electron density perfectly.
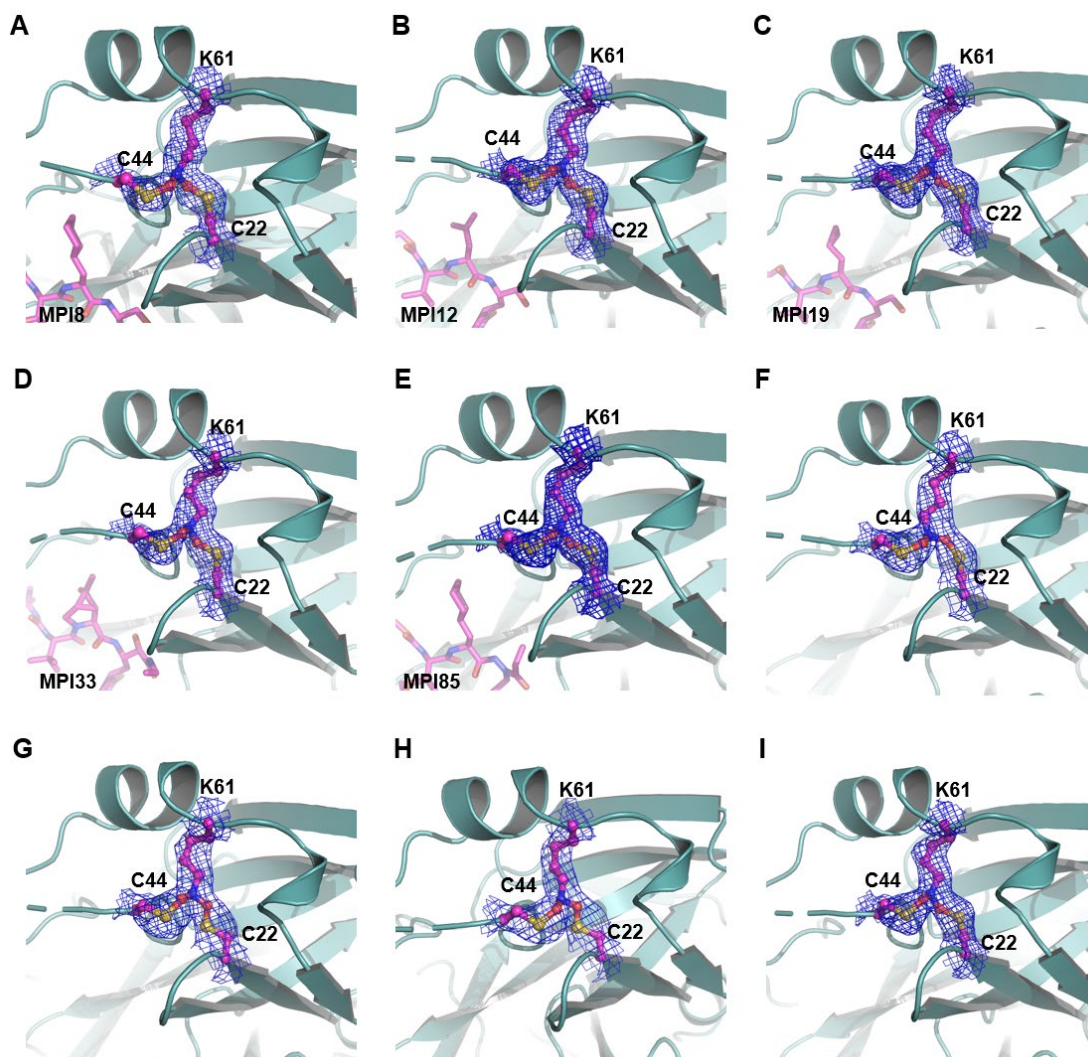
Figure 2. Five M^{Pro}-ligand complexes and four apo enzyme structures show a SONOS-bridged crosslink between three residues including C22, C44 and K61. The electron density was contoured at 1 σ. The structures information is provided in Table S1.

Among about 10 early crystal structures of M^{Pro}-inhibitor complexes, M^{Pro}-MPI8 was the only one that contained the Y-shaped crosslink. However, this crosslink was subsequently also observed in several M^{Pro}-inhibitor complexes determined later. Figures 2B-E show crystal structures for four other M^{Pro} complexes with MPI12, MPI19, MPI33, and MPI85, respectively. They all show a well-defined Y-shaped, S-O-N-O-S-bridged

crosslink between K61 and two cysteine residues C22 and C44. We suspected that the inhibitor binding process might have triggered the formation of the crosslink, however, a careful search of our other determined structures revealed four apo-M^{Pro} structures that contained this crosslink as well. As shown in Figures 1F-H, all four structures showed clearly connected electron density at the ends of three residues C22, C44 and K61. These four apo-M^{Pro} structures were determined from apo-M^{Pro} crystals that were soaked with chemical fragments but no bound ligands were observed. So far, all structures that showed this crosslink were determined from crystals that were exposed to the air due to the soaking process.

We also determined X-ray crystal structures for several apo-M^{Pro} crystals that were not exposed to the air. They all had a structurally defined aa46-50 region and exhibited no crosslink between C22, C44 and K61. One of these structures was previously deposited into the Protein Data Bank (PDB entry: 7JPY).[26] As shown in Figure 3A, the superposition of 7JPY over an apo-M^{Pro} structure with the Y-shaped crosslink reveals a large structural rearrangement at the region aa43-52 between two structures. In 7JPY, the side chain of C44 points toward the active site. The small aa46-51 helix tucks C44 toward the active site and makes it a key component to form the S2 binding pocket for a substrate P2 residue. The thiol group of C44 is 9.1Å away from the K61 side chain amine (Figure 3A in lemon). This distance is too far to form any possible direct interaction between the two residues. In order for C44 to physically meet K61 for a covalent interaction, the C44 backbone α-carbon moved 4.1Å closer to K61 and rotated its side chain almost 180˚ to adopt a conformation to form a covalent adduct with the K61 side chain amine as shown in the crosslinked apo-M^{Pro} (Figure 3A in light teal). This structural rearrangement also pushed the whole aa43-52 region away from the M^{Pro} active site, making the active site much more open. In order for C44 to translocate and rotate, a large structural rearrangement for the whole aa43-52 region is required. This explains why the Y-shaped S-O-N-O-S crosslink was observed in our crystal structures but not in most other determined M^{Pro} structures since the protein packaging in our M^{Pro} crystals allowed the large structure rearrangement at aa43-52. In 7JPY, C22 and K61 are in a close distance to each other. This close distance will likely make them form an N-O-S-bridged crosslink first and then a structurally flipped C44 will be engaged to generate the second N-O-S bridge. Since all our M^{Pro} structures

that contained the Y-shaped posttranslational crosslink were determined from crystals that were exposed to the air, the molecular oxygen was most likely the reagent that generated the Y-shaped S-O-N-O-S bridge. Although how exactly the molecular oxygen oxidizes the three residues to form the crosslink needs to be further explored and confirmed, we propose a likely mechanism as shown in Figure 3B. In this mechanism, the C22 thiolate reacts with oxygen to form a peroxysulfane that then reacts with lysine to generate the first S-O-N crosslink. After C44 flips toward K61, its thiolate can react with oxygen to form the second peroxysulfane that then reacts with the first S-O-N crosslink to form the Y-shaped, S-O-N-O-S-bridged tri-residue crosslink.
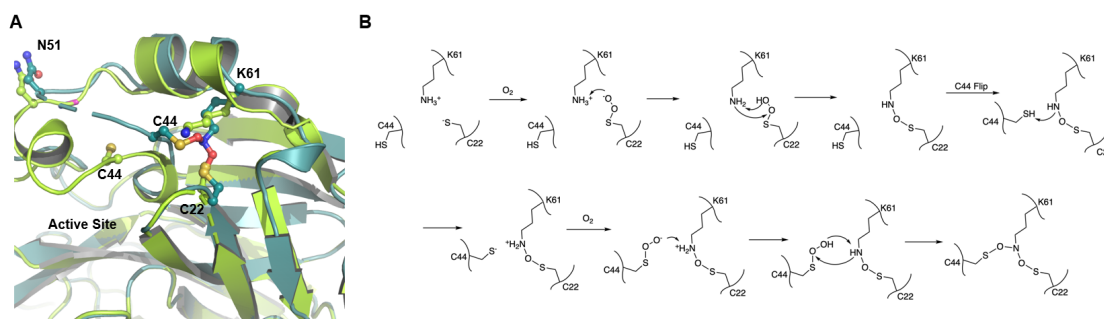

Figure 3. (**A**) The superposition of two different apo-M^{Pro} structures. One doesn't have the S-O-N-O-S-bridged triresidue crosslink (lemon) and the other one does (light teal). (**B**) A proposed scheme for the formation of the SONOS-bridged triresidue crosslink in M^{Pro}.

Since the crosslink formation generates a much more opened active site in M^{Pro}, it presumably weakens the binding of a peptide substrate to M^{Pro} and consequently leads to a less active enzyme. To test this prospect, we freshly expressed M^{Pro} and purified it without adding a reducing reagent in lysis and purification buffers. 1 mM DTT was previously used to maintain *N. gonorrhoeae* transaldolase in its reduced state presumably due to its reduction of the N-O-S bond.[31] We added 1 mM DTT to the freshly purified M^{Pro} as well for 30 min to reduce its potentially generated Y-shaped, S-O-N-O-S-bridged tri-substrate crosslink. We then tested the activities of the enzyme in two conditions, one with the addition of 1 mM DTT and one without. The results, as shown in Figure 4A, clearly indicate that the addition of DTT leads to about a 60% increase in activity in M^{Pro} indicating that the S-O-N-O-S crosslink is potentially a redox switch. Since there are no information such as chemical stability about the S-O-N-O-S crosslink in the literature, we conducted a

native mass spectrometry analysis of the freshly purified M^{Pro} to keep the potential S-O-N-O-S crosslink intact. Our collected spectrum of M^{Pro} at 25°C clearly showed the M^{Pro} dimer, the M^{Pro} dimer with two additional oxygen atoms, and the M^{Pro} dimer with four additional oxygen atoms (Figure 4B). The M^{Pro} dimer with four additional oxygen atoms is equivalent to M^{Pro} with the Y-shaped crosslink. The M^{Pro} dimer with two additional oxygen atoms is likely an intermediate state that has either just one M^{Pro} monomer with the S-O-N-O-S crosslink or both M^{Pro} monomers with just the S-O-N crosslink. An interesting observation was also made during the native mass spectrometry analysis of M^{Pro} when we did the in-instrument heating of M^{Pro} by sequentially increasing the temperature by 5°C and waiting for 30 min before the native mass spectrometry analysis was conducted. When the temperature was increased to 40°C, the majority of the M^{Pro} dimer was converted to the one containing four additional oxygen atoms. One likely explanation is that the protein at 40°C has enhanced structural dynamics that assists the formation of the Y-shaped S-O-N-O-S crosslink. Therefore, both our activity and native mass spectrometry results support that the Y-shaped, S-O-N-O-S-bridged tri-residue crosslink potentially serves as a redox switch to regulate the SARS-CoV-2 M^{Pro} activity.
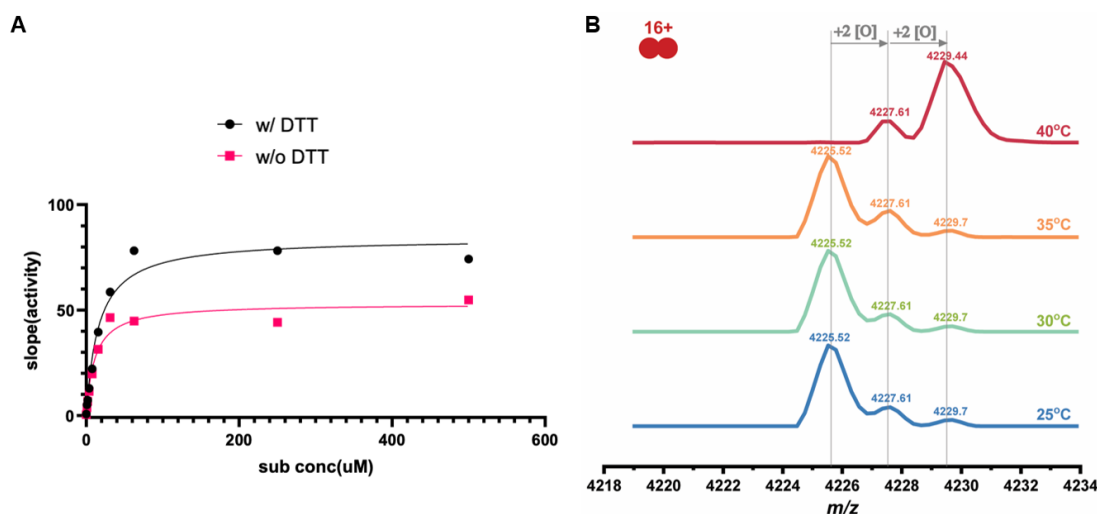
Figure 4. (A) The activity of M^{Pro} that was freshly purified without a reducing reagent in lysis and purification buffers measured in the presence (K_m: 11 ± 2 uM) and absence (K_m: 16 ± 3 uM) of 1 mM DTT. **(B)** Native mass spectra of M^{Pro} that was freshly purified without a reducing reagent in lysis and purification buffers and sequentially treated in the instrument by increasing the temperature by 5°C and waiting for 30 min.

In summary, we have observed a uniquely Y-shaped, S-O-N-O-S-bridged crosslink between three residues C22, C44 and K61 in the SARS-CoV-2 M[Pro]. As a redox switch and novel posttranslational modification, this crosslink regulates the SARS-CoV-2 M[Pro]. The discovery of this crosslink is significant in that the S-O-N-O-S crosslink is a novel posttranslational modification and this novel crosslink potentially provides a key information in understanding the SARS-CoV-2 pathogenesis. Many COVID-19 patients have experienced damage to their lungs that induces hypoxia.[32] A hypoxic condition that provides a reducing environment will favor more active M[Pro] that drives the SARS-CoV-2 reproduction in infected human lungs. Our discovery may partially explain the severe lung damage in COVID-19 patients since the more active M[Pro] enhances SARS-CoV-2 reproduction and therefore aggravates the symptom. SARS-CoV-2 patients will typically develop inflammation and redox imbalance with altered levels of metabolites and antioxidants.[33] There have been many proposals to use redox-based therapeutics to potentially treat COVID-19 patients. One typical suggestion is to take antioxidants. However, the intake of antioxidants could further activate M[Pro].[34] Our current discovery calls into question whether taking antioxidants will be definitely beneficial to COVID-19 patients. Our finding is also important in assisting the development of M[Pro] inhibitors as SARS-CoV-2 antivirals. The generation of the crosslink leaves a more open, larger active site allowing for the design and testing of many new inhibitors. One significant benefit of inhibitors that target this more open active site is their synergy in driving and stabilizing the crosslink formation. Since the crosslink leads to a less active enzyme, small molecules that could permanently lock C22, C44 and K61 in their crosslinked form in M[Pro] will lead to the demise of SARS-CoV-2. Exploring this direction will allow the development of a totally new group of SARS-CoV-2 antivirals that can be provided to COVID-19 patients together with active site-targeting M[Pro] inhibitors, including paxlovid, for potential synergistic effects in killing the virus.

ACKNOWLEDGMENT

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

REFERENCES

1   Martinez-Flores, D. *et al.* SARS-CoV-2 Vaccines Based on the Spike Glycoprotein and Implications of New Viral Variants. *Front Immunol* **12**, 701501, doi:10.3389/fimmu.2021.701501 (2021).

2   Mengist, H. M. *et al.* Mutations of SARS-CoV-2 spike protein: Implications on immune evasion and vaccine-induced immunity. *Semin Immunol* **55**, 101533, doi:10.1016/j.smim.2021.101533 (2021).

3   Mahase, E. Covid-19: Pfizer's Paxlovid is 89% Effective in Patients at Risk of Serious Illness, Company Reports. *theBMJ* **375** (2021).

4   Owen, D. R. *et al.* An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. *Science* **374**, 1586-1593, doi:10.1126/science.abl4784 (2021).

5   Kim, D. *et al.* The Architecture of SARS-CoV-2 Transcriptome. *Cell* **181**, 914-921 e910, doi:10.1016/j.cell.2020.04.011 (2020).

6   Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273, doi:10.1038/s41586-020-2012-7 (2020).

7   Morse, J. S., Lalonde, T., Xu, S. & Liu, W. R. Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV. *Chembiochem* **21**, 730-738, doi:10.1002/cbic.202000047 (2020).

8   Dai, W. *et al.* Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **368**, 1331-1335, doi:10.1126/science.abb4489 (2020).

9   Jin, Z. *et al.* Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289-293, doi:10.1038/s41586-020-2223-y (2020).

10   Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* **368**, 409-412, doi:10.1126/science.abb3405 (2020).

11   Gunther, S. *et al.* X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. *Science* **372**, 642-646, doi:10.1126/science.abf7945 (2021).

12   Qiao, J. *et al.* SARS-CoV-2 M(pro) inhibitors with antiviral activity in a transgenic mouse model. *Science* **371**, 1374-1378, doi:10.1126/science.abf1611 (2021).

13   Dai, W. *et al.* Design, Synthesis, and Biological Evaluation of Peptidomimetic Aldehydes as Broad-Spectrum Inhibitors against Enterovirus and SARS-CoV-2. *J Med Chem* **65**, 2794-2808, doi:10.1021/acs.jmedchem.0c02258 (2022).

14   Ma, C. *et al.* Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res* **30**, 678-692, doi:10.1038/s41422-020-0356-z (2020).

15   Ghahremanpour, M. M. *et al.* Identification of 14 Known Drugs as Inhibitors of the Main Protease of SARS-CoV-2. *ACS Med Chem Lett* **11**, 2526-2533, doi:10.1021/acsmedchemlett.0c00521 (2020).

16   Zhu, W. *et al.* Identification of SARS-CoV-2 3CL Protease Inhibitors by a Quantitative High-Throughput Screening. *ACS Pharmacol Transl Sci* **3**, 1008-1016, doi:10.1021/acsptsci.0c00108 (2020).

17   Hoffman, R. L. *et al.* Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. *J Med Chem* **63**, 12725-12747, doi:10.1021/acs.jmedchem.0c01063 (2020).

18   Bai, B. *et al.* Peptidomimetic nitrile warheads as SARS-CoV-2 3CL protease inhibitors. *RSC Med Chem* **12**, 1722-1730, doi:10.1039/d1md00247c (2021).

19   Bai, B. *et al.* Peptidomimetic alpha-Acyloxymethylketone Warheads with Six-Membered Lactam P1 Glutamine Mimic: SARS-CoV-2 3CL Protease Inhibition, Coronavirus Antiviral Activity, and in Vitro Biological Stability. *J Med Chem* **65**, 2905-2925, doi:10.1021/acs.jmedchem.1c00616 (2022).

20   Breidenbach, J. *et al.* Targeting the Main Protease of SARS-CoV-2: From the Establishment of High Throughput Screening to the Design of Tailored Inhibitors. *Angew Chem Int Ed Engl* **60**, 10423-10429, doi:10.1002/anie.202016961 (2021).

21   Dampalla, C. S. *et al.* Postinfection treatment with a protease inhibitor increases survival of mice with a fatal SARS-CoV-2 infection. *Proc Natl Acad Sci U S A* **118**, e2101555118, doi:10.1073/pnas.2101555118 (2021).

22   Ghosh, A. K. *et al.* Indole Chloropyridinyl Ester-Derived SARS-CoV-2 3CLpro Inhibitors: Enzyme Inhibition, Antiviral Efficacy, Structure-Activity Relationship, and X-ray Structural Studies. *J Med Chem* **64**, 14702-14714, doi:10.1021/acs.jmedchem.1c01214 (2021).

23   Ma, C. *et al.* Discovery of Di- and Trihaloacetamides as Covalent SARS-CoV-2 Main Protease Inhibitors with High Target Specificity. *J Am Chem Soc* **143**, 20697-20709, doi:10.1021/jacs.1c08060 (2021).

24   Unoh, Y. *et al.* Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19. *J Med Chem*, e2024358118, doi:10.1021/acs.jmedchem.2c00117 (2022).

25   Kantsadi, A. L., Cattermole, E., Matsoukas, M. T., Spyroulias, G. A. & Vakonakis, I. A COVID moonshot: assessment of ligand binding to the SARS-CoV-2 main protease by saturation transfer difference NMR spectroscopy. *J Biomol NMR* **75**, 167-178, doi:10.1007/s10858-021-00365-x (2021).

26   Yang, K. S. *et al.* A Quick Route to Multiple Highly Potent SARS-CoV-2 Main Protease Inhibitors*. *ChemMedChem* **16**, 942-948, doi:10.1002/cmdc.202000924 (2021).

27   Li, L. *et al.* Self-Masked Aldehyde Inhibitors: A Novel Strategy for Inhibiting Cysteine Proteases. *J Med Chem* **64**, 11267-11287, doi:10.1021/acs.jmedchem.1c00628 (2021).

28   Li, S. G. *et al.* An Enhanced Hybrid Screening Approach to Identify Potent Inhibitors for the SARS-CoV-2 Main Protease From the NCI Compound Library. *Front Chem* **10**, 816576, doi:10.3389/fchem.2022.816576 (2022).

29   Ma, Y. *et al.* The P3 O - Tert -Butyl-Threonine is Key to High Cellular and Antiviral Potency for Aldehyde-Based SARS-CoV-2 Main Protease Inhibitors. *bioRxiv*, doi:10.1101/2021.12.18.473326 (2021).

30   Alugubelli, Y. R. *et al.* The N -Terminal Carbamate is Key to High Cellular and Antiviral Potency for Boceprevir-Based SARS-CoV-2 Main Protease Inhibitors. *bioRxiv*, doi:10.1101/2021.12.18.473330 (2021).

31   Wensien, M. *et al.* A lysine-cysteine redox switch with an NOS bridge regulates enzyme function. *Nature* **593**, 460-464, doi:10.1038/s41586-021-03513-3 (2021).

32   Rahman, A. *et al.* Silent hypoxia in COVID-19: pathomechanism and possible management strategy. *Mol Biol Rep* **48**, 3863-3869, doi:10.1007/s11033-021-06358-1 (2021).

33   Paul, B. D., Lemle, M. D., Komaroff, A. L. & Snyder, S. H. Redox imbalance links COVID-19 and myalgic encephalomyelitis/chronic fatigue syndrome. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2024358118 (2021).

34   Fratta Pasini, A. M., Stranieri, C., Cominacini, L. & Mozzini, C. Potential Role of Antioxidant and Anti-Inflammatory Therapies to Prevent Severe SARS-Cov-2 Complications. *Antioxidants (Basel)* **10**, 272, doi:10.3390/antiox10020272 (2021).

35   Yang, K. S. *et al.* A Quick Route to Multiple Highly Potent SARS-CoV-2 Main Protease Inhibitors**. *Chemmedchem* **16**, 942-948, doi:10.1002/cmdc.202000924 (2021).

36   Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).

37   Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

38   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

39   McCabe, J. W. *et al.* Variable-Temperature Electrospray Ionization for Temperature-Dependent Folding/Refolding Reactions of Proteins and Ligand Binding. *Anal Chem* **93**, 6924-6931, doi:10.1021/acs.analchem.1c00870 (2021).

Supplemental Information

Figure S1. Eight dimers of M^Pro are shown to present the packing forms in crystals. Similar colors are given to the homogeneous monomer in one dimer. The catalytic dyad C145 and H41 are labeled as yellow. Region of aa46-50 is labeled as red.
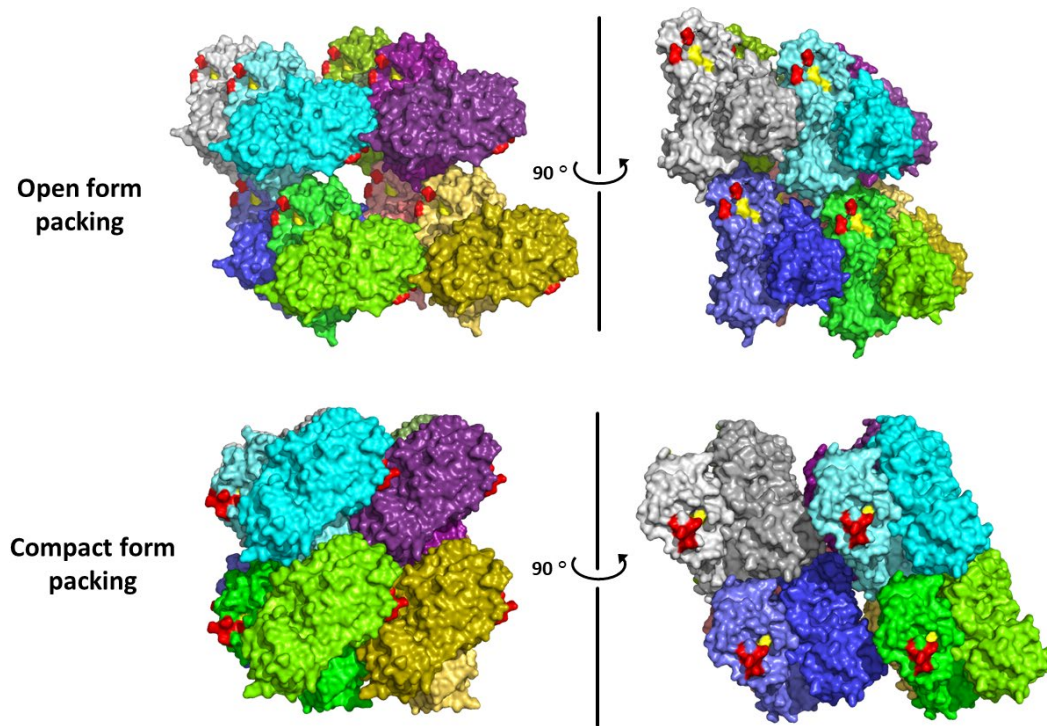
Figure S2. The low electron density map around aa46-50 in a representative M^{Pro}-inhibitor complexes.
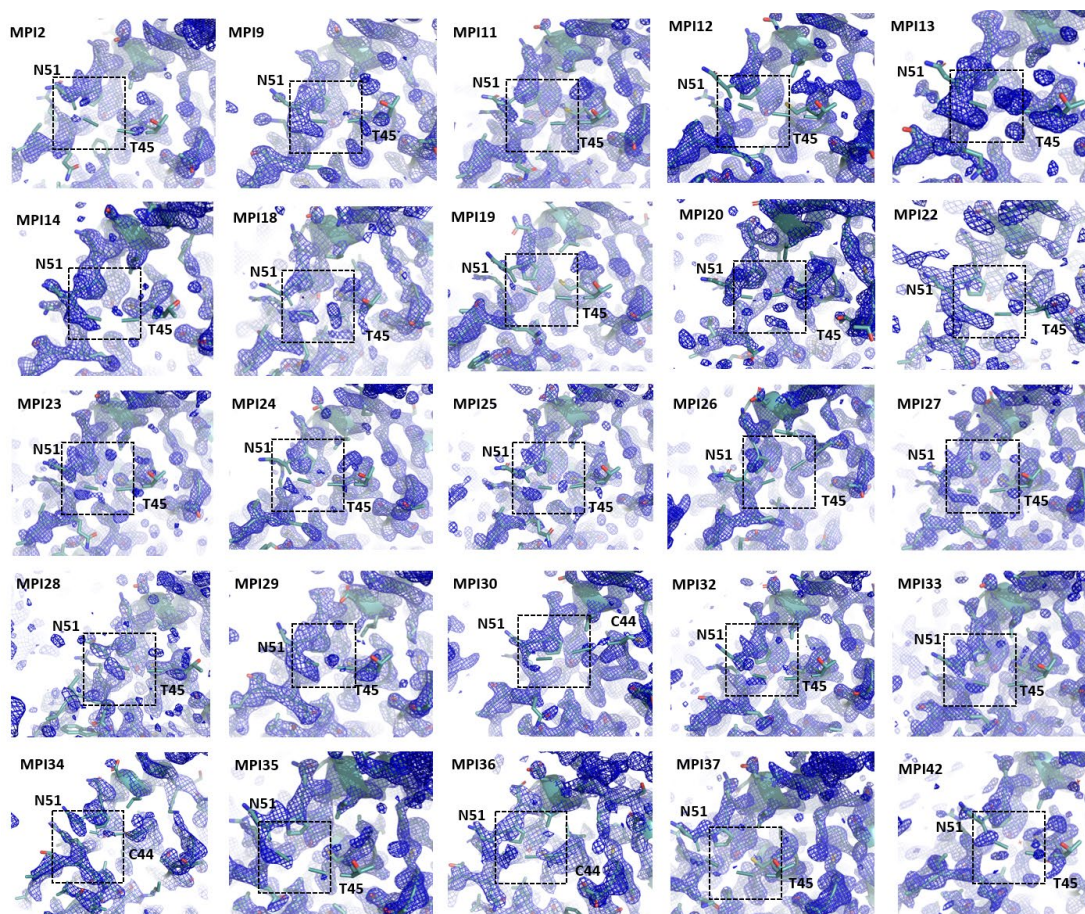
Figure S3. Full mass spectra of M^{Pro} at T = 25, 30, 35, and 40˚C. All the spectra were collected with the same borosilicate glass tip to reduce the possible variability from tips to tips. The charge states were calculated on the basis of theoretical mass of M^{Pro} (Molecular weight 33,792) and the resolved charge state envelop. Both monomeric (with 10+ as most abundant species) and dimeric (with 16+ as most abundant species) M^{Pro} existed given the concentration of 4 µM. The 2[O] mass difference shown in the main text was calculated by using M^{Pro} dimeric form with 16+ charges.
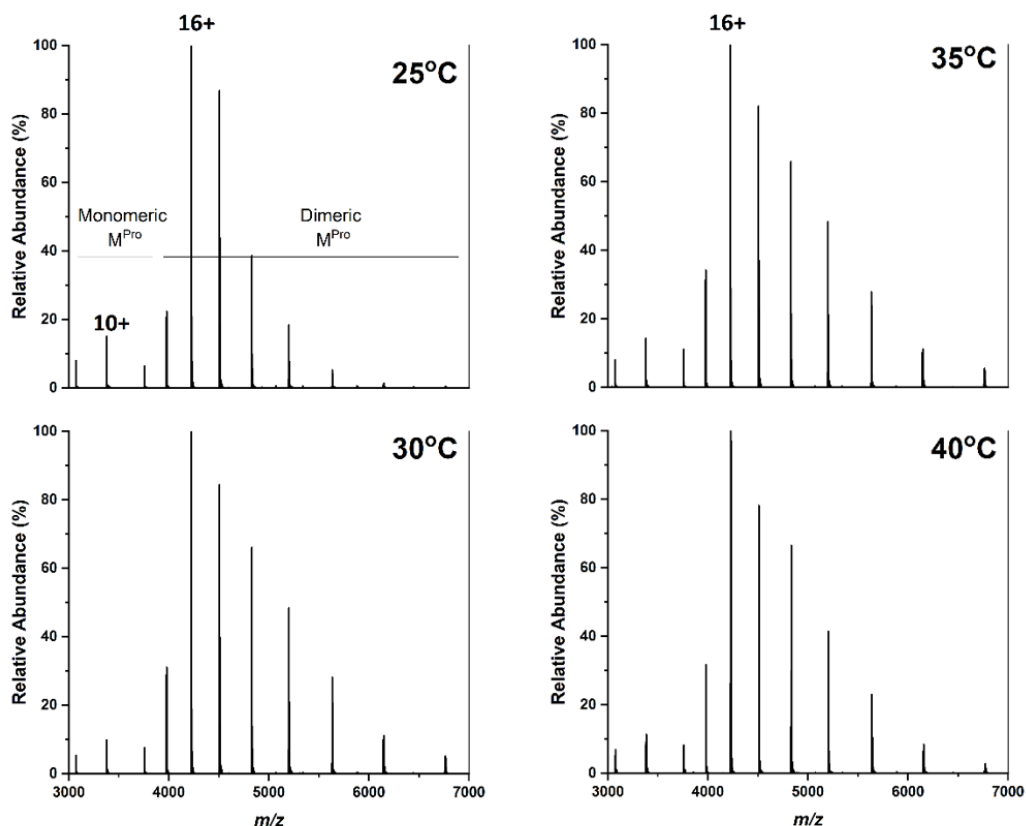
Table S1. Statistics of crystallographic analysis of M$^{\mathrm{Pro}}$ in apo form and in complexed with inhibitors.

Protein/Ligand (PDI entry)	apo1 (7UU6)	apo2 (7UU7)	apo3 (7UU8)	apo4 (7UU9)	MPI8 (7UUA)
Data Collection					
Space group	I 1 2 1	I 1 2 1	I 1 2 1	I 1 2 1	I 1 2 1
cell dimensions					
a, b, c (Å)	51.6293 82.4004 88.6871	51.9582 83.3449 90.1396	52.1505 83.6505 89.3834	51.5037 83.059 87.8969	54.2471 80.8179 85.7217
α, β, γ (°)	90 96.7536 90	90 96.8018 90	90 95.9493 90	90 96.1827 90	90 96.9239 90
Resolution (Å)	24.3 - 1.85 (1.916 - 1.85)	24.6 - 2.491 (2.58 - 2.491)	24.56 - 2.5 (2.589 - 2.5)	24.22 - 2.474 (2.563 - 2.474)	24.09 - 1.85 (1.916 - 1.85)
R_{merge}	0.117 (1.589)	0.078 (0.255)	0.118 (0.631)	0.112 (0.631)	0.100 (1.261)
$I/\sigma I$	14.1 (0.6)	19.4 (4.7)	13.3 (2.4)	14.2 (2.5)	13.8 (1.6)
Completeness (%)	95.30 (92.08)	97.72 (82.61)	99.32 (94.84)	99.01 (92.42)	99.39 (99.23)
Redundancy	17.4 (5.2)	9.5 (6.1)	9.4 (6.6)	9.7 (7.1)	11.3 (8.0)
Refinement					
No. Reflections	30060 (2917)	13126 (1107)	13211 (1267)	13089 (1219)	31202 (3106)
$R_{\mathrm{work}}/R_{\mathrm{free}}$	0.2239 /0.2526	0.1944 /0.2466	0.2021 /0.2584	0.2012 /0.2500	0.2435 /0.2731
No. atoms					
Protein	2360	2360	2360	2360	2327
Water	154	81	67	68	234
B factors					
Protein	36.02	33.85	41.37	41.35	28.2
Water	39.13	34.19	39.5	38.99	32.59
R.m.s deviations					
Bond lengths (Å)	0.011	0.009	0.015	0.01	0.01
Bond angles (°)	1.16	1.07	1.26	1.22	1.06

Protein/Ligand (PDI entry)	MPI12 (7UUB)	MPI19 (7UUC)	MPI33 (7UUD)	MPI85 (7UUE)
Data Collection				
Space group	I 1 2 1	I 1 2 1	I 1 2 1	I 1 2 1
cell dimensions				
a, b, c (Å)	51.92 81.18 90.0856	52.01 81.32 89.4029	54.2617 81.3062 85.6688	54.6222 83.3566 85.8644
α, β, γ (°)	90 96.2965 90	90 96.1217 90	90 97.4294 90	90 96.8577 90
Resolution (Å)	47 - 1.63 (1.688 - 1.63)	46.93 - 1.6 (1.657 - 1.6)	23.82 - 1.85 (1.916 - 1.85)	24.73 - 1.85 (1.916 - 1.85)
R_{merge}	0.066 (0.776)	0.050 (2.064)	0.071 (0.701)	0.084 (2.043)
$I/\sigma I$	12.5 (1.6)	11.3 (0.8)	19.4 (2.8)	16.6 (1.3)
Completeness (%)	95.46 (91.40)	99.56 (99.81)	95.77 (94.30)	93.47 (83.22)
Redundancy	5.9 (5.5)	4.8 (4.8)	11.8 (8.3)	10.4 (7.4)
Refinement				
No. Reflections	44234 (4239)	48554 (4840)	30206 (2961)	30516 (2698)
$R_{\mathrm{work}}/R_{\mathrm{free}}$	0.2093 /0.2281	0.2082 /0.2292	0.2089 /0.2356	0.2041 /0.2314
No. atoms				
Protein	2347	2347	2338	2338
Water	199	149	221	232
B factors				
Protein	34.53	42.18	26.81	28.03
Water	40.9	48.51	32.2	34.32
R.m.s deviations				
Bond lengths (Å)	0.01	0.013	0.01	0.01
Bond angles (°)	1.29	1.93	1.12	1.08

MATERIALS AND METHODS

Recombinant M^{Pro} protein expression and purification

The pET28a-His-SUMO-M^{Pro} expression and purification are according to our previous report.[35] The pET28a-His-SUMO-M^{Pro} construct was transformed into *E. coli* BL21(DE3) cells. Transformed cells were cultured at 37°C in 2xYT medium with kanamycin (50 g/mL) until OD_{600} reaching 0.6, then induced with 1 mM isopropyl-D-1-thiogalactoside (IPTG) at 37 °C. After 3 hours, cells were harvested and lysed in buffer A (20 mM Tris, 100 mM NaCl, 10 mM imidazole, pH 8.0). The supernatant was loaded onto a nickel-chelating column (GenScript) washed with buffer A, followed by elution with buffer B (20 mM Tris, 100 mM NaCl, 250 mM imidazole, pH 8.0). The eluted protein solution was desalted to buffer C (20 mM Tris, 10 mM NaCl, pH 8.0) by HiPrep 26/10 desalting column (GE Healthcare). The His-SUMO-M^{Pro} proteins were digested with SUMO protease overnight at 4°C. The digested protein was applied to nickel-chelating column again to remove the His-tagged SUMO protease, the His-SUMO tag, and the expressed protein with uncleaved His-SUMO tag. The tag-free M^{Pro} protein was loaded to the size exclusion column HiPrep 16/60 Sephacryl S-100 HR (GE Healthcare) pre-equilibrated with buffer D (20 mM Tris, 100 mM NaCl, 1 mM EDTA, pH 7.8). The eluted M^{Pro} protein was stored with buffer D in -80 °C for further use.

Km analysis of M^{Pro}

The assays were carried out with 50 nM M^{Pro} at 37°C with continuous shaking. The Sub3 substrate (DABCYL-Lys-Thr-Ser-Ala-Val-Leu-Gln-Ser-Gly-Phe-Arg-Lys-Met-Glu-EDANS) was purchased from BACHEM and stored as 10 mM solution in 100% DMSO at -80°C. Enzyme activity was monitored by detecting fluorescence with excitation at 336 nm and emission at 490 nm wavelength. The dilution buffer (used for enzyme and substrate dilution) is 10 mM $Na_xH_yPO_4$,10 mM NaCl, 0.5 mM EDTA, pH 7.6. To conduct the Km analysis, the enzyme was diluted to 100 nM using dilution buffer with or without 1 mM DTT and set on ice for 30 min. 25 µL of the 100 nM enzyme solution was then added to a 96-well plate and incubated at 37ºC for 30 min. When the incubation period was over, 25 µL of each substrate solution, final concentration from 4 µM to 500 µM with 5% DMSO,

was added to each well to initiate the assay. The first 0-200 seconds were analyzed by linear regression for initial slope analyses with GraphPad Prism 8.0.

X-Ray Crystallography Analysis

The production of crystals of apo M^{Pro} and M^{Pro}-inhibitor complexes was following the previous protocols with the crystal growing condition of 0.1 M Bis-tris, pH6.5, 16% w/v PEG10k.[35] The data of M^{Pro} with MPI12 was collected on a Rigaku R-AXIS IV++ image plate detector, the data of M^{Pro} with MPI19 was collected at the Advanced Light Source (ALS) beamline 5.0.2 using a Pilatus3 6M detector, all the other data were collected on a Bruker Photon II detector. The diffraction data were indexed, integrated and scaled with iMosflm or PROTEUM3[36]. The structure was determined by molecular replacement using the structure model of the free enzyme of the SARS-CoV-2 M^{pro} [Protein Data Bank (PDB) ID code 7JPY] as the search model using Phaser in the Phenix package [35,37]. *JLigand* and *Sketcher* from the CCP4 suite were employed for the generation of PDB and geometric restraints for the inhibitors. The inhibitors were built into the *F$_O$-F$_C$* density by using *Coot* [38]. Refinement of all the structures was performed with Real-space Refinement in Phenix [37]. Details of data quality and structure refinement are summarized in Table S1. All structural figures were generated with PyMOL (https://www.pymol.org).

Native Mass Spectrometry

Native mass spectrometry (nMS) analysis was performed on a Q Exactive UHMR Hybrid Quadruple-Orbitrap Mass Spectrometer (ThermoFisher) with m/z range was set from 1,000 to 10,000 and with resolution set to 12,500 (at m/z 400). 10 μL of the sample was loaded to a borosilicate glass capillary tip (Sutter, CA) with 1100 to 1500 V spray voltage supplied by an inserted platinum wire. Important parameters to reduce non-specific adducts include: capillary temperature 100oC, in-source trapping and activation -10 V, ion transfer high m/z, collision-induced dissociation (CID) 10 eV, and higher energy dissociation (HCD) 30 V. In variable-temperature electrospray ionization (vT-ESI) experiment, the temperature of solution was controlled as previously described.[39] The temperature was varied from 25oC to 40oC with 5oC increments and the equilibrium time at each temperature was 5 minutes.

REFERENCES

1.   Yang, K. S.; Ma, X. Y. R.; Ma, Y. Y.; Alugubelli, Y. R.; Scott, D. A.; Vatansever, E. C.; Drelich, A. K.; Sankaran, B.; Geng, Z. Z.; Blankenship, L. R.; Ward, H. E.; Sheng, Y. J.; Hsu, J. C.; Kratch, K. C.; Zhao, B. Y.; Hayatshahi, H. S.; Liu, J.; Li, P. W.; Fierke, C. A.; Tseng, C. T.; Xu, S. Q.; Liu, W. R., A Quick Route to Multiple Highly Potent SARS-CoV-2 Main Protease Inhibitors**. *Chemmedchem* **2021,** *16* (6), 942-948.

2.   Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 235-42.

3.   Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.

4.   Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 4), 486-501.

5.   McCabe, J. W.; Shirzadeh, M.; Walker, T. E.; Lin, C. W.; Jones, B. J.; Wysocki, V. H.; Barondeau, D. P.; Clemmer, D. E.; Laganowsky, A.; Russell, D. H., Variable-Temperature Electrospray Ionization for Temperature-Dependent Folding/Refolding Reactions of Proteins and Ligand Binding. *Anal Chem* **2021,** *93* (18), 6924-6931.