Electronic Supplementary Material (ESI) for ChemComm.

SARS-CoV-2 M^{pro} inhibition by zinc ion: structural features and hints for drug design

Deborah Grifagni,[a] Vito Calderone,[a,b,c] , Stefano Giuntini,[a] Francesca Cantini,[a,b,c] Marco Fragai[a,b,c] and Lucia Banci [a,b,c]

[a] Magnetic Resonance Center (CERM), University of Florence, via Sacconi 6, Sesto Fiorentino, 50019 Italy

[b] Department of Chemistry "Ugo Schiff", University of Florence, via della Lastruccia 3, Sesto Fiorentino, 50019 Italy

[c] Consorzio Interuniversitario Risonanze Magnetiche MetalloProteine (CIRMMP), via Sacconi 6, Sesto Fiorentino, 50019 Italy

Table of Contents

Protein expression and purification

The expression vector PGEX-6P-1 encoding a full-length SARS-CoV-2 M^{pro} (ORF1ab polyprotein residues 3264-3569, GenBank code: MN908947) with an additional N-terminal AVLQ and C-terminal GPHHHHHH was acquired from MRC PPU. The gene is codon optimized for expression in Escherichia Coli and it has been cloned between the BamHI and NotI restriction sites of plasmid pGEX-6P-1.

The gene construct was designed to have at the N-terminus the cleavage sequence of M^{pro} (SAVLQ↓SGFRK; the arrow indicates the cleavage site) such that auto-cleavage of M^{pro} itself occurs in the cell during gene expression.

Native C-terminal can be obtained by treatment with PreScission which recognizes and cuts the following sequence at the arrow (SGVTFQ↓GPHHHH)

Protein expression was carried out in E. coli BL21 (DE3). Transformed clones were pre-cultured at 37°C in 100 mL LB medium and grown ON. 10 mL of preculture were inoculated in 2 L of 1 x YT media supplemented with 0.05 mg/mL ampicillin. Cells were grown at 37 °C until they reached an OD600 ~1, the temperature was reduced to 16 °C, and protein expression was induced with addition of IPTG to 1 mM. After 5 h, cells were harvested by centrifugation at 5000 x g, 4°C for 15 min. Protein purification was performed following the protocol previously reported [1]. The final yield was 8 mg/1 L of culture (Fig. S1). The molar mass of monomeric SARS-CoV-2 native M^{pro} is 33796.8 Da. Analytical size exclusion chromatography was used to determine the oligomeric state that resulted to be dimeric. 200 ul of 200 uM M^{pro}, and standard mixture (Sigma-Aldrich) were injected onto a Superdex® 200 Increase 10/300 GL. The running buffer consisted of 25 mM MES, 150 mM NaCl, pH 6.5 at room temperature (Fig. S1).

Crystallization, Data Collection and Structure Solution

Crystals of apo SARS-CoV-2 M^{pro} were obtained in sitting drop by adding an aliquot of 2 µL of protein solution (20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.8) to 2 µL of reservoir buffer (20 mM ammonium acetate, 20% PEG 3350 pH 7) and stored at 20 °C. The protein concentration in the sample was 5 mg/mL.

The native crystals of apo SARS-CoV-2 M^{pro} were afterwards soaked in a reservoir solution containing $ZnCl_2$ having a ten-fold zinc concentration with respect to the protein for 5 days. Both the native and the Zn-adduct datasets were collected in-house, using a BRUKER D8 Venture diffractometer equipped with a PHOTON III detector, at 100 K; the crystals used for data collection were cryo-cooled using 25% ethylene glycol in the mother liquor. The crystals diffracted up to 1.8 Å resolution: they belong to space group C2 with one molecule in the asymmetric unit, a solvent content of about 50%, and a mosaicity of 0.4°. The data were processed using the program XDS [2], reduced and scaled using XSCALE [2] and amplitudes were calculated using XDSCONV [2]. The structure of the apo protein and of the zinc adduct were solved using the molecular replacement technique; in the first case, the model used was 6YB7 whereas in the second case the model used was the apo structure obtained in-house. The successful orientation hand translation of the molecule within the crystallographic unit cell was determined with MOLREP [3]. The refinement was carried out using PHENIX [4], applying TLS restraints. In between the refinement cycles, the model was subjected to manual rebuilding using COOT [5]. Water molecules have been added by using the standard procedures within the

ARP/WARP [6] suite. The quality of the refined structures was assessed using the program MOLPROBITY [7]. Data processing and refinement statistics are shown in Table S1. Coordinates and structure factors have been deposited at the PDB under the accession code 7NXH for the apo and 7NWX for the zinc one.

NMR Spectroscopy

NMR spectra were acquired at 303K on Avance 900 Bruker spectrometer equipped with triple resonance cryoprobe. The binding of Zn^{2+} by SARS-CoV-2 M^{pro} was assessed by titrating the protein in 20mM HEPES, 50 mM NaCl, 0,5 mM TCEP at pH 6.5 with a solution of $ZnCl_2$ and following the process through ^{1}H-^{15}N TROSY-HSQC NMR spectra. A solution of M^{pro} (200 uL) at the concentration of 110 µM was titrated with stock solutions of $ZnCl_2$ at the concentration of 2,5 mM and 10 mM, respectively in order to add at each time volumes of 3-14 uL and to reach a final volume of 220 uL thus avoiding over-dilution. The titration was designed to record spectra at 0.3, 0.6, 1, 1.5 and 3 protein to Zn^{2+} ratio.

Each addition of zinc was carried out directly into the NMR tube which was then appropriately shaken to allow for a homogenous mixing and then placed back into the NMR spectrometer.

Activity assay

The effect of Zn^{2+} ion on the enzymatic activity of SARS-CoV-2 M^{pro} was determined by evaluating ability to inhibit the hydrolysis of fluorescence-quenched peptide substrate (Mca–AVLQ↓SGFR-K(Dnp)K) (Genescript). The catalytic efficiency (k_{cat}/K_m) for SARS-CoV-2 M^{pro} of the peptide Is reported to be 28,500 M^{-1} s^{-1} [8, 9].

The protein stock at 5 mg/mL concentration in 20 mM Tris-HCl 150mM NaCl 5 mM DTT was initially incubated in a measuring buffer (20 mM Tris-HCl 20% glycerol pH 7.2) at 30°C for 5 minutes at a final concentration of 0.2 uM. In order to exclude inhibitors possibly acting as aggregators, a detergent-based control was performed by adding 0.01% freshly made up Triton X-100 to the reaction at the same time [10]. The activity measurements were carried out in a 500 uL cuvette at 30° by adding the fluorescent peptide substrate at a final concentration of 4 uM and measuring the reaction kinetics for 1 minute at an excitation and emission wavelength of 320 and 405 nm, respectively. Then, the proteolytic activity of SARS-CoV-2 M^{pro} was measured in the presence of increasing concentrations of $ZnCl_2$ (0.2, 0.5, 0.75, 1, 1.25, 2, 4, 8 and 16 uM) (Fig. S4a). All experiments were performed in triplicate. Initial rates of substrate cleavage (V) were extracted as the slope of linear fits at 18 seconds of the fluorescence signal increase as a function of time (Fig. S4b).

The following equation was subsequently used to describe the effectiveness of zinc on hydrolysis rate.

$V= V_{max}+ (V_{min} - V_{max}) * x^n / (k_i^n + x^n)$

Here V is the experimentally measured rate of substrate cleavage calculated as described above and X is the concentration of Zn^{2+}. The Hill coefficient, n was fixed to 1, whereas the K_i, V_{max} and V_{min} were the fitting parameters in the non-linear fitting routine. All experimental data was analyzed using Originpro 2018.

Fig. S1 Analytical gel filtration of SARS-CoV-2 M^{pro} (red) and standard proteins (black). Running buffer: 25 mM MES, 150 mM NaCl pH 6.5.

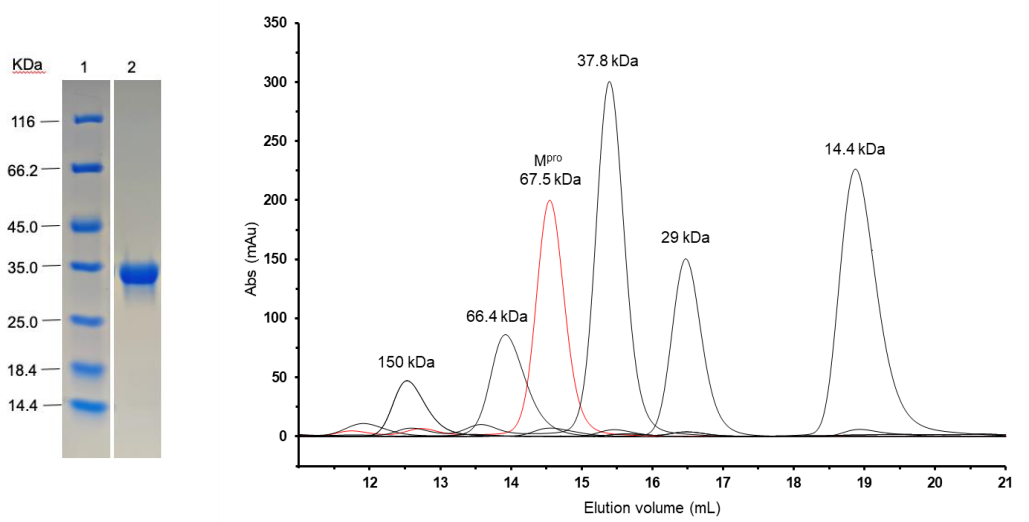


Fig. S2 X-ray structure of the homodimeric SARS-CoV-2 M^{pro}. The two monomers are shown in light blue and khaki respectively. His 41 and Cys 145 are shown as sticks. Zinc ion is shown as magenta sphere

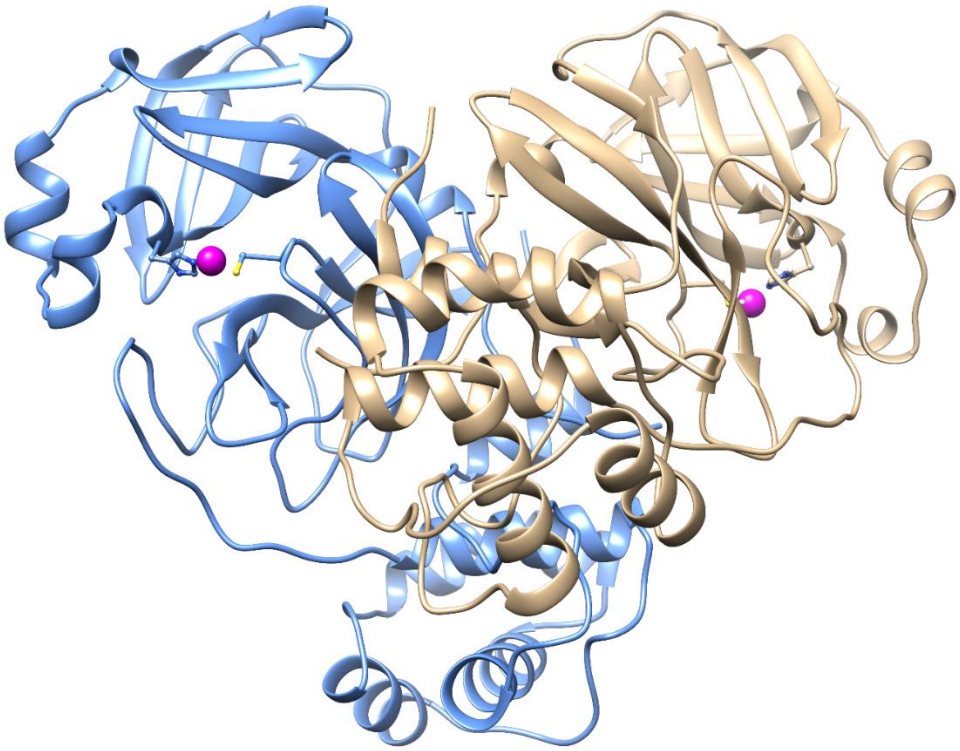


Fig. S3 Per-residue comparison of backbone RMSD (Root Mean Square Deviation) between apo and Zn-bound forms of SARS-CoV-2 M^{pro}

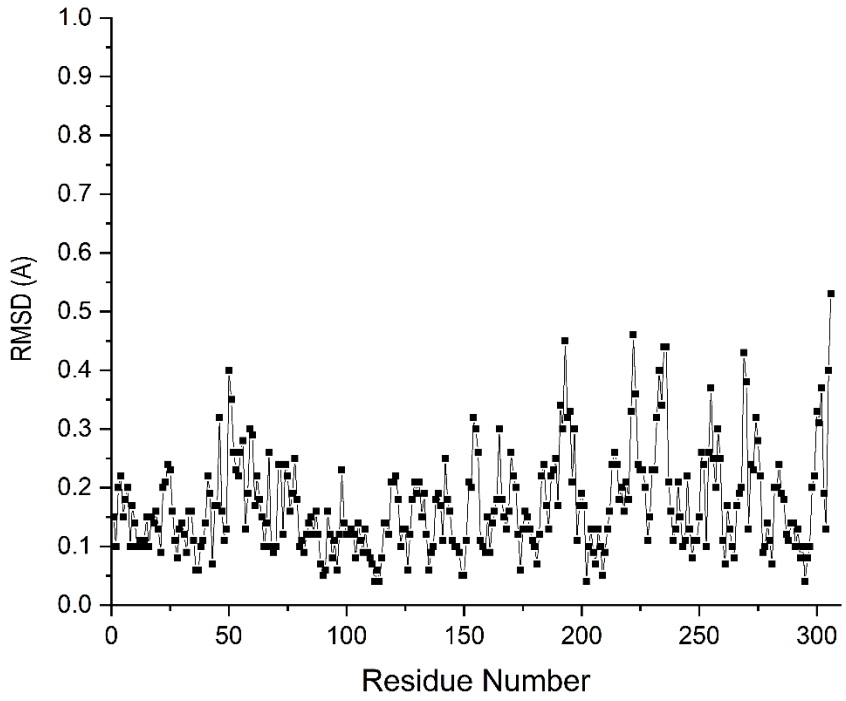
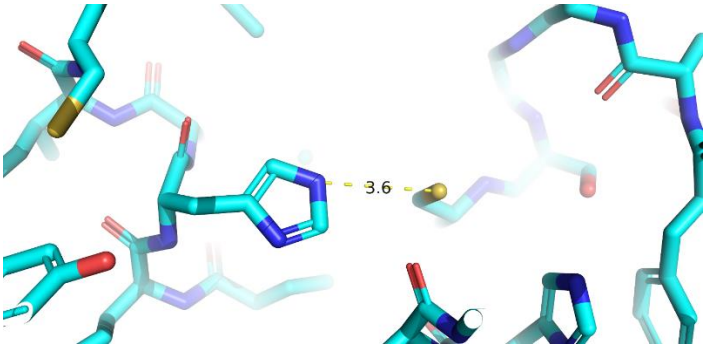


Fig. S4 Active site of the apo SARS- COV-2 M^{pro} (**a**) and zinc bound SARS-COV-2 M^{pro} (**b**).Distances between the sulfur atom of Cys145, the Nε atom of the imidazole ring of His41 and the zinc ion are shown.

a



b

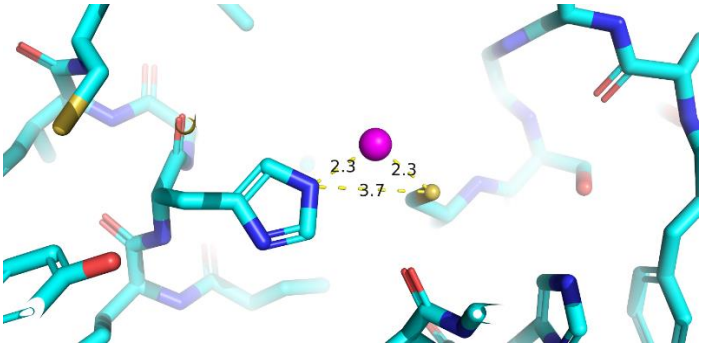


Fig. S5 Representation of activity based assay. a. SARS-CoV-2 M^{pro} activity at increasing concentration of $ZnCl_2$ b. Linear fit at 18 seconds of the of the fluorescence signal increase as a function of time.

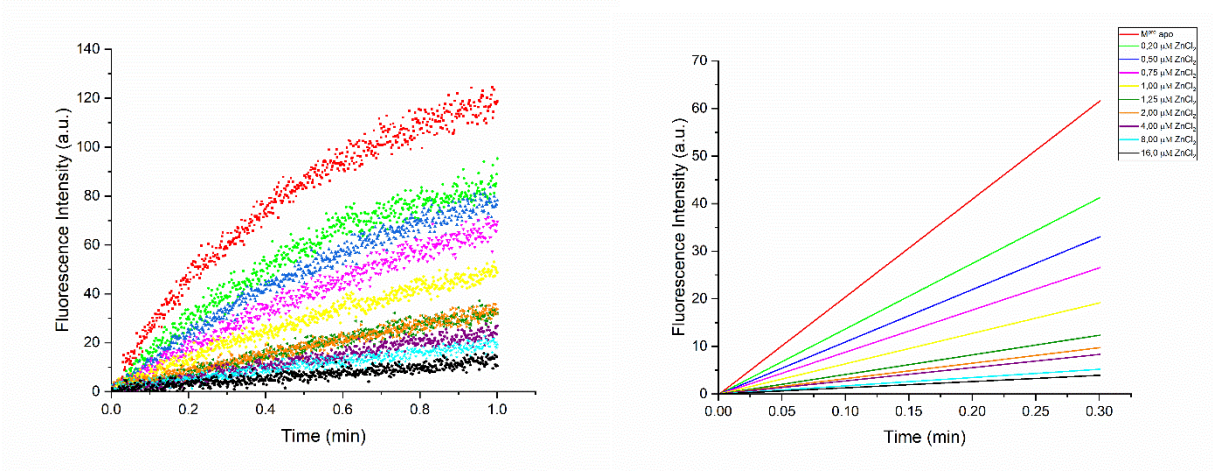


Fig. S6 Active site of SARS-CoV-2 M^{pro} bound to zinc pyrithione (PDB ID: 7B83). Zinc is shown as magenta sphere and pyrithione molecule is shown as stick

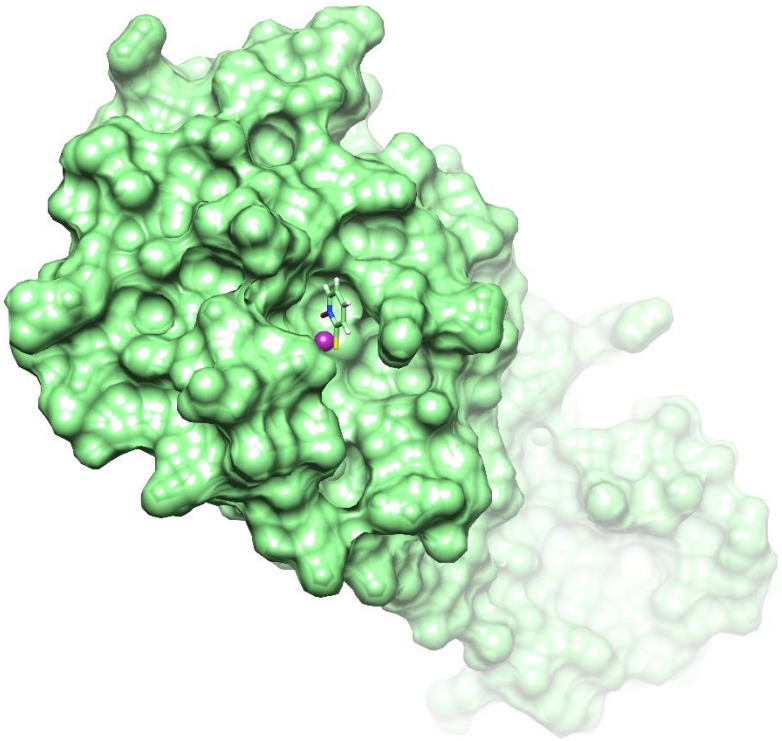


Table S1, Crystallographic data processing and refinement statistics.

	7NXH	7NWX
Diffraction source	BRUKER D8 Venture	BRUKER D8 Venture
Wavelength (Å)	1.541	1.541
Temperature (K)	100	100
Detector	BRUKER PHOTON III	BRUKER PHOTON III
Crystal-detector distance (mm)	60	60
Oscillation range (°)	0.5	0.5
Total rotation range (°)	360	360
Exposure time/image (s)	120	120
Space group	C2	C2
a, b, c (Å), β (°)	113.91, 53.43, 44.97, 101.852	114.62, 53.36, 44.65, 102.02
Mosaicity (°)	0.4	0.4
Resolution range (Å)	48.18 – 2.10 (2.23 – 2.10)	48.18 – 1.80 (1.91 – 1.80)
Total reflections	87305 (9584)	130973 (7326)
Unique reflections	14851 (1989)	20490 (1891)
Completeness (%)	95.4	92.1 (50.1)
CC1/2	0.99 (0.82)	99.7 (48.0)
I/(σI)	8.48 (1.30)	13.4 (1.4)
R_{merge} †	0.17 (0.99)	0.12 (0.95)
Wilson B factor ($Å^2$)	38.9	26.2
R_{cryst} / R_{free}‡ (%)	21.6/26.8	20.4/24.4
Protein atoms	2368	2437
Water molecules	70	80
Ions	-	1
RMSD bond lengths (Å)	0.002	0.007
RMSD bond angles (º)	0.538	0.938

References

1. L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox and R. Hilgenfeld, *Science*, 2020, **368**, 409-412.
2. W. Kabsch, *Acta Crystallogr. D. Biol. Crystallogr*, 2010, **66**, 125-132.
3. A. A. Vagin and A. Teplyakov, *Acta Crystallogr D Biol Crystallogr*, 2000, **56**, 1622-1624.
4. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 213-221.
5. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr*, 2010, **66**, 486-501.
6. G. Langer, S. X. Cohen, V. S. Lamzin and A. Perrakis, *Nat Protoc*, 2008, **3**, 1171-1179.
7. V. B. Chen, W. B. Arendall, III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, *Acta Crystallogr. D Biol. Crystallogr*, 2010, **66**, 12-21.
8. J. L. Lauer-Fields, M. Kele, G. D. Sui, H. Nagase, R. M. Leblanc and G. B. Fields, *Analytical Biochemistry*, 2003, **321**, 105-115.
9. Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao and H. Yang, *Nature*, 2020, **582**, 289-293.
10. B. Y. Feng and B. K. Shoichet, *Nat Protoc*, 2006, **1**, 550-553.