Electronic Supplementary Information

Allosteric Inhibition of SARS-CoV-2 3CL Protease by Colloidal Bismuth Subcitrate

Xuan Tao,[a] Lu Zhang,[b] Liubing Du,[c] Ruyan Liao,[b] Huiling Cai,[b] Kai Lu,[a] Zhennan Zhao,[a] Yanxuan Xie,[a] Pei-Hui Wang,[d] Ji-An Pan,[c] Yuebin Zhang,*[e] Guohui Li,*[e] Jun Dai,*[b] Zong-Wan Mao*[a] and Wei Xia*[a]

a. MOE Key Laboratory of Bioinorganic and Synthetic Chemistry School of Chemistry, Sun Yat-Sen University, Guangzhou, 510275, China

b. Guangzhou Customs District Technology Center, No. 66 Huacheng Avenue, Zhujiang New Town, Tianhe district, Guangzhou, 510700

c. The Center for Infection and Immunity Study, School of Medicine, Sun Yat-sen University, Guangming Science City, Shenzhen, 518107, China

d. Key Laboratory for Experimental Teratology of Ministry of Education and Advanced Medical Research Institute, Cheeloo College of Medicine, Shandong University, Jinan, 250012, China

e. State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, 116023, China

Cloning, expression and purification of SARS-CoV-2 3CLpro.

Construction of SARS-CoV-2 3CLpro expression plasmid was the same as described previously.[1] In brief, 3CLpro gene was PCR-amplified using the codon-optimized SARS-CoV-2 *nsp5* gene as a template. The primers used were listed in Table S1. The 12 nucleotides encoding amino acids AVLQ were added before 3CLpro Ser1 to generate an autocleavage site of 3CLpro. The 24 nucleotides encoding amino acids GPH_6 were added at the C-terminus of 3CLpro sequence to form a PreScission cleavage site. The PCR product was double-digested with BamHI and XhoI and ligated into pGEX-6p-1 plasmid that was digested with the same restriction enzymes. The generated plasmid pGEX-3CLpro was then transformed into BL21(DE3) cells. 5 mL overnight culture of such BL21(DE3) cells were transferred into 1 L LB medium supplemented with 50 μg/mL carbenicillin. When OD_{600} reached 0.6-0.8, IPTG was added to a final concentration of 0.3 mM to induce protein expression. Bacteria were further cultured at 20 °C for another 16 hours and collected by centrifugation at 4,000 ×*g*. Cell pellet were resuspended with Ni-NTA buffer A1 (40 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM phenylmethanesulfonyl fluoride (PMSF) and 10 mM imidazole) and lysed by high pressure homogenizer (Union-Biotech, Shanghai). After centrifugation at 20,000 ×*g* for 30 minutes, the supernatant was filtered and loaded onto a Ni-NTA column (Qiagen) pre-equilibrated with Ni-NTA buffer A1. Ni buffer A2 (40 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM PMSF and 50 mM imidazole) was used to remove nonspecific binding proteins and Ni buffer B (40 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine (TCEP), 1 mM PMSF and 500 mM imidazole) was used to elute the bound 3CLpro. The eluted protein was dialyzed against the PreScission digestion buffer (20 mM Tris-HCl, pH 8.0, 250 mM NaCl, 1 mM TCEP) and PreScission protease was added to remove His-tag with a molar ratio of 1:100 at 4 °C overnight. The dialyzed sample was pooled and loaded onto a Superdex Increase 200 10/300 column (GE Healthcare) pre-equilibrated with GF buffer (20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM TCEP). The purified 3CLpro protease was flash frozen by liquid nitrogen and stored at -80 °C. Plasmids for SARS-CoV-2 3CLpro mutants were generated via site-directed mutagenesis using pGEX-3CLpro as a

template with primers listed in Table S1. Expression and purification of 3CLpro mutants were similar to wild type.

***In vitro* protease activity assay of SARS-CoV-2 3CLpro.**

The protease activities of SARS-CoV-2 3CLpro and mutants were determined using the fluorogenic peptides MCA-AVLQSGFR-Lys(Dnp)-Lys-NH_2 as substrates.[2] The excitation and emission wavelengths were set at 320 nm (excitation) and 405 nm (emission). All the florescence spectra were recorded in a Cytation 3 multi-mode plate reader.

For inhibitory metallodrug screening, 0.5 μM 3CLpro was pre-incubated with 50 μM metallodrug (final concentration) in reaction buffer (20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM TCEP) on ice for 20 minutes. Subsequently, fluorogenic peptide substrate was added into the reaction mixture to a final concentration of 10 μM. The substrate cleavage reaction was carried out in a black 96-well plate at 30 °C for 30 min and the fluorescence intensities were recorded.

For IC_{50} calculation, 10 μM peptide substrates and 0-4 μM colloidal bismuth subcitrate (CBS) or bismuth gluconate or 0-2 μM GC376 were mixed with 0.5 μM 3CLpro. The relative protease activities were plotted against the log value of CBS concentrations. Dose-response curves for half-maximum inhibitory concentration (IC50) values were determined by nonlinear regression.

For protease kinetic analysis, the reaction mixture contained consisted of 0.625-40 μM 3CLpro peptide substrate in reaction buffer with a total volume of 200 μL. After addition of 0.5 μM 3CLpro with or without 1 μM CBS, the increase of fluorescence at 520 nm (excited at 329 nm) was continuously monitored at 30 °C. Fluorescence intensity was converted into the amount of the hydrolytic product through a standard curve measured by MCA-AVLQ peptides with known concentrations. The increase in fluorescence was linear for at least 10 min, and thus the slope of the line represented the initial velocity (v). The steady-state kinetic parameters of the enzyme were determined by fitting the Michaelis-Menten equation to the initial velocity data and peptide substrate concentrations.

***In cellulo* assay of 3CLpro protease activity.**

The details of 3CLpro protease activity reporter system (UB-CS-FL) are described in another manuscript (in preparation). In brief, the reporter gene, firefly luciferase (FL) is fused with polyubiquitin (UB) which leads to the quick degradation of luciferase in proteasome. A 3CL protease cleavage site was inserted between FL and UB. The cleavage on this site by 3CLpro set free the luciferase, the activity of which could be detected using luciferase activity assay.

HEK293T cells were co-transfected with the plasmid encoding 3CLpro, the corresponding UB-CS-FL and pRL-TK (Renilla luciferase expression plasmid) as an internal control. HEK293T cells co-transfected with a blank plasmid, UB-CS-FL and pRL-TK were used as a negative control. 24 h post-transfection, the cells were left untreated or treated with varying concentrations of CBS or GC376. 24 h post-treatment, the cells were collected and subjected to dual-luciferase reporter assays (Dual-Glo® Luciferase Assay System) following the manufacturer's instruction and the luciferase activities were quantified on Synergy H1 Hybrid Multi-Mode Reader (BIOTEK). Experiments were performed in triplicate.

Metal content determination.

The protein-bound metal contents were determined using a Thermo Scientific iCAP Q inductively coupled plasma mass spectrometry (ICP-MS). Each sample was quantified three times and the average value was used. For 3CLpro and different mutants, 5 molar equivalents of CBS were incubated with 10 μM protein in Tris-HCl buffer (20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM TCEP) on ice for 20 min. After removing the excess amount of Bi(III) by Hitrap Desalting column, the bound Bi(III) contents were determined by ICP-MS coupled with BCA assay as described above.

Protein oligomerization state analysis.

Size-exclusion chromatography (SEC) was performed with a Tricorn Superdex 200 10/300 GL analytical column (GE Healthcare) at 4 ^{o}C. The column was pre-equilibrated with Tris-HCl buffer (20 mM Tris-HCl, 300 mM NaCl, 1 mM TCEP, pH 7.4).

Approximate 10 μM protein sample (3CLpro or 3CLpro^{C300S}) was pre-incubated with indicated molar equivalents of CBS for 20 min at 4 ^{o}C before loading to the column at a flow rate of 0.5 ml/min. For analysis of 3CLpro^{E290A} and 3CLpro^{R298A}, about 10 μM protein was directly loaded onto the column at a flow rate of 0.5 ml/min. The Tricorn Superdex Increase 200 10/300 GL column was calibrated with GE HMW calibration kit.

Calculation of solvent accessible area of 3CLpro.

The AREAIMOL software in the CCP4i Structure Analysis package was used to calculate the solvent accessible area of each residue in SARS-CoV-2 3CLpro. The SARS-CoV-2 3CLpro dimer structure (PDBID: 6M2Q) was used as an input structure model and the software was run with default radius of probe solvent molecule (1.4 Å). The scores for each Cys residue of 3CLpro have been listed in Table S2.

Molecular dynamics simulations.

The initial conformation of SARS-CoV-2 3CLpro was obtained from the crystal structure of 3CLpro dimer (PDB id: 6M2Q). The simulation systems were set up using Amber Tools 19 and the AMBER ff14SB force field[3] was used for the protein. The TIP4PEW water model was used to solvate the systems and Na(I) counter ions were used to neutralize the charges. The 12-6 Lennard-Jones parameter of Bi(III) was obtained from the Universal force field (UFF)[4] and a stoichiometric ratio of 20 Bi(III) : 1 3CLpro dimer was used in the MD simulations to enhance the diffusion of Bi(III) ions, which were randomly inserted around the 3CLpro dimer initially. The systems were refined using 500 steepest descent steps before switching to conjugate gradient energy minimization and gradually heated to 300K within 2ns. The positional restraints were exerted on the backbone atoms of the 3CLpro dimer with a weight of 10 kcal/mol/Å^2 during the energy minimization and heating process. Then the restraints were released gradually within six equilibration steps in the NPT ensemble. The hydrogen mass repartitioning was set to 4 amu to enable an integration step of 4fs for the production runs. The Langevin Integrator was used with a collision frequency of

1.0 ps^{-1} to couple the system temperature at 300 K. The Particle-Mesh Ewald (PME)[5] and the dispersion correction algorithm were exploited to estimate the contribution of long-range non-bonded interactions that beyond the cutoff of 12 Å. The MM/GBSA analysis were performed using MMPBSA.py module in AMBER package.[6]

Evaluation of antiviral activity of CBS.

To evaluate the antiviral efficacy of CBS, Vero E6 cells were cultured overnight in 96-well cell-culture plate with a density of 2×10^4 cells/well. Cells were pre-treated with the different doses of the indicated CBS (0-2000 μM) for 1 h. The SARS-CoV-2 virus (MOI of 0.05) was subsequently added to allow infection for 1 h. Virus-CBS mixtures were then replaced with fresh medium that only contains CBS. After 24 hours, the plate was fixed with 4% PFA. The SARS-CoV-2 nucleocapsid protein (N protein) in infected cell was detected with a polyclonal rabbit antibody against SARS-CoV-2-N protein followed by a goat-anti rabbit IgG antibody (Alexa Fluor®488, ab150077). Cell nuclei were labelled with the 4, 6-diamidino-2-phenylindole (DAPI) nucleic acid stain. The fluorescence images were acquired with a Celigo Imaging Cytometer.

Cell viability assay.

The cell viability was measured by Cell Counting Kit-8 (CCK-8) (Sangon, Shanghai). Vero E6 cells were cultured overnight in 96-well cell-culture plates with 2×10^4 cells/well. CBS at varying concentrations from 0-2000 μM were used to treat Vero E6 cells. After 24 hours, 10 μL CCK-8 reagent was added into each well and the incubation time was 2 hours. The absorption values at 450 nm were acquired by a Cytation 3 multi-mode plate reader.

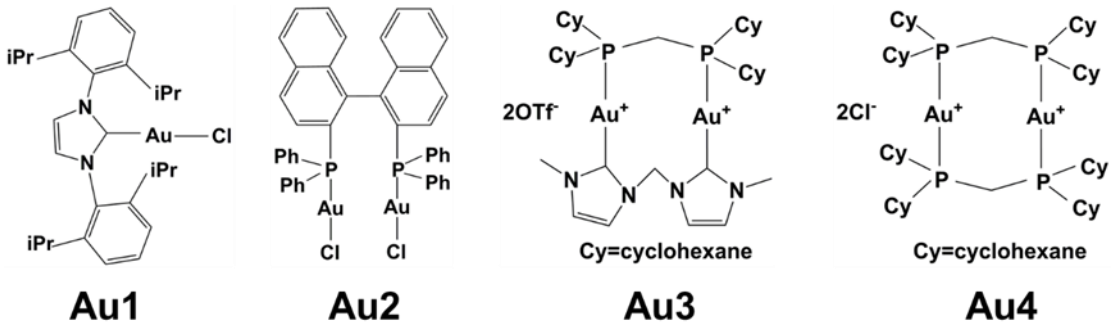


Figure S1. The structures of the four active Au(I) complexes used in *in vitro* protease inhibition assay**.**

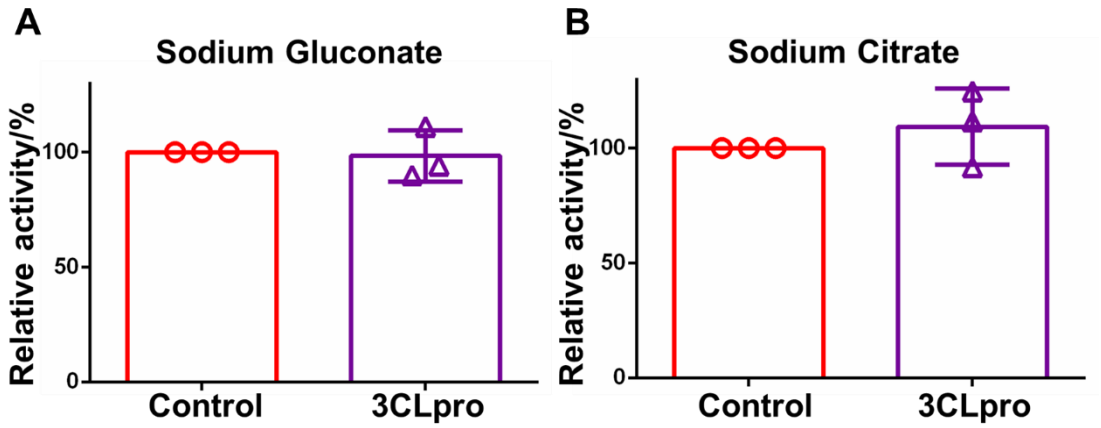


Figure S2. The relative protease activity of 3CLpro in the presence of 50 μM sodium gluconate (A) or sodium citrate (B). Each experiment were performed in triplicates. The data are shown in mean ± sd.

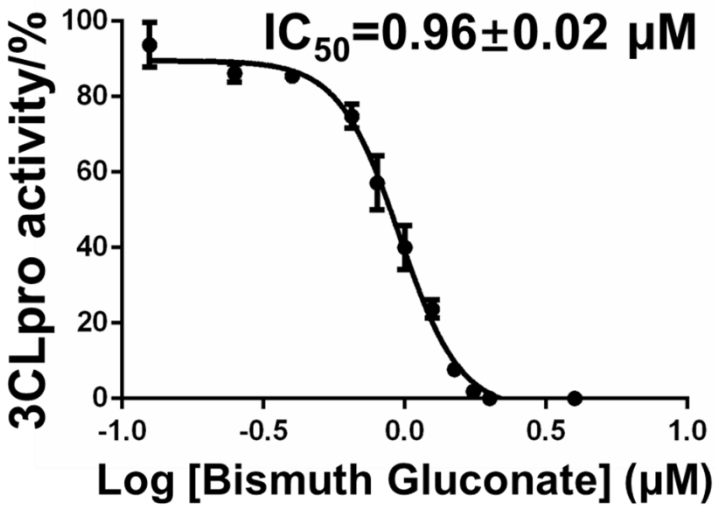


Figure S3. Inhibition of the protease activities of 3CLpro by bismuth gluconate at varying concentrations *in vitro*. Dose-response curves for half-maximum inhibitory concentration (IC_{50}) values were determined by nonlinear regression. Each experiment was performed in triplicates. The data are shown in mean ± sd.

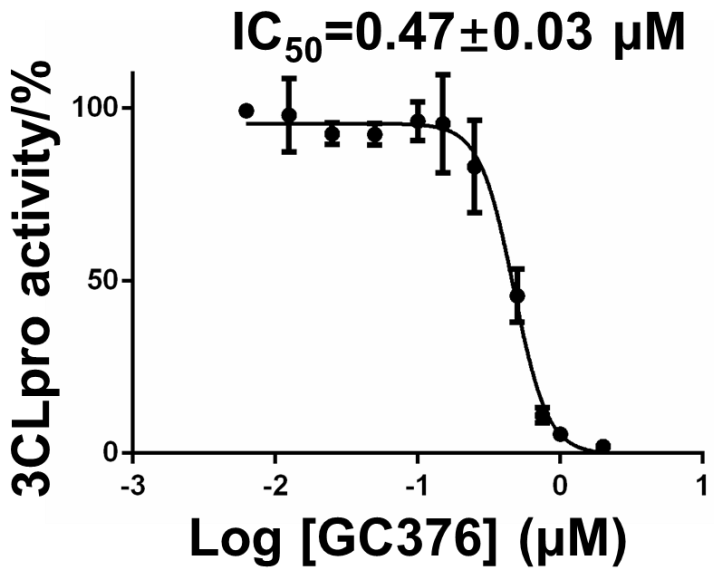
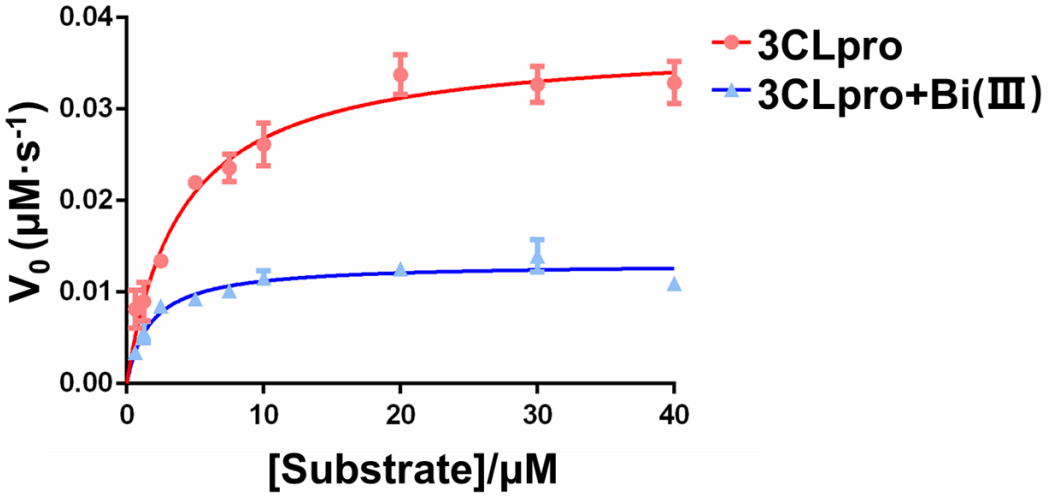


Figure S4. Inhibition of the protease activities of 3CLpro by GC376 at varying concentrations *in vitro*. Dose-response curves for half-maximum inhibitory concentration (IC_{50}) values were determined by nonlinear regression. Each experiment was performed in triplicates. The data are shown in mean ± sd.



	k_{cat} (10^{-2} s^{-1})	K_m (10^{-6} M)	k_{cat}/K_m (10^4 $s^{-1}\cdot M^{-1}$)
3CLpro	7.46±0.24	3.93±0.46	1.90±0.32
3CLpro+Bi(Ⅲ)	2.62±0.10	1.73±0.27	1.51±0.35

Figure S5. Kinetic analysis of SARS-CoV-2 3CLpro protease activities in the absence and presence of bismuth drug inhibitor. The calculated apparent Michaelis-Menten constant (K_m) and k_{cat} were listed.

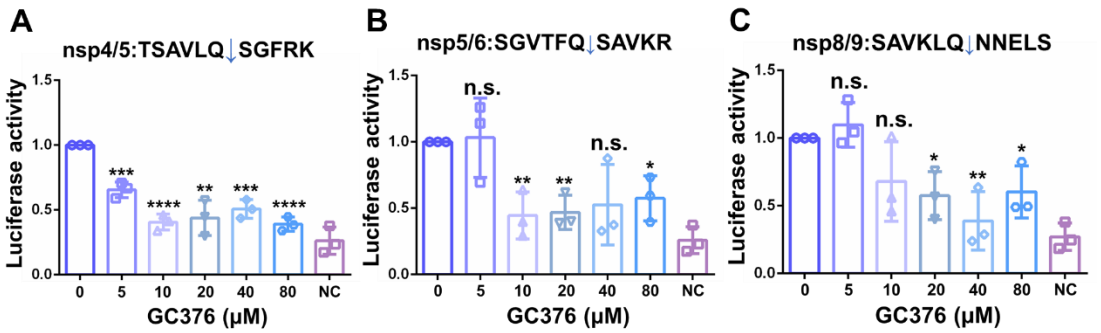


Figure S6. Inhibition of SARS-CoV-2 3CLpro by GC376 *in cellulo*. The HEK293T cells co-expressing 3CLpro and luciferase reporter with nsp4/5 (A), nsp5/6 (B) and nsp8/9 (C) cleavage sites are treated with varying concentrations of GC376. NC represents negtive control. Bar charts show the relative luciferase activity of cell lysate. Each experiment were performed in triplicates. The data are shown in mean ± sd. * p<0.05, ** p<0.01, *** p<0.001, ****p<0.0001.

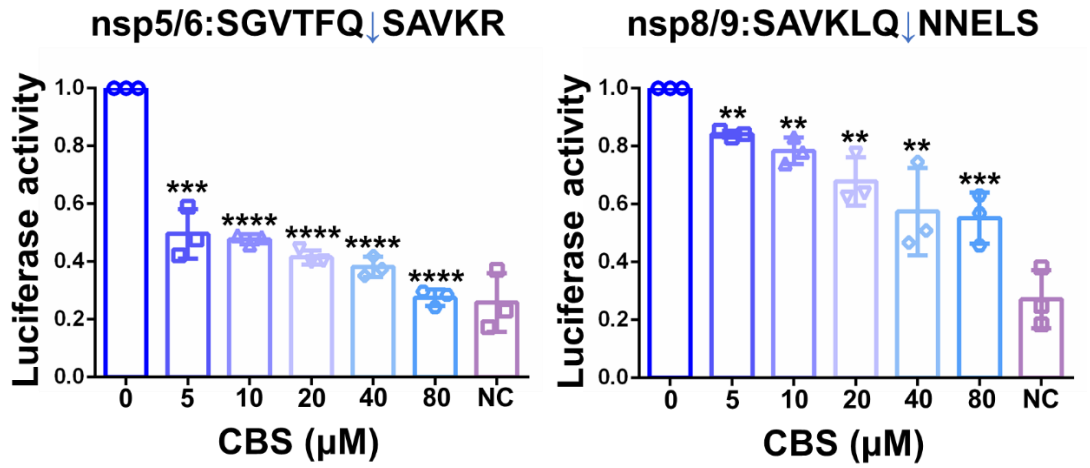


Figure S7. Inhibition of SARS-CoV-2 3CLpro by bismuth drug *in cellulo*. The 3CLpro and luciferase reporter with nsp5/6 (A) and nsp8/9 (B) cleavage sites are co-expressed in HEK293T cells. The cells are then treated with varying concentrations of colloidal bismuth subcitrate (CBS). NC represents negtive control. Bar charts show the relative luciferase activity of cell lysate. Each experiment were performed in triplicates. The data are shown in mean ± sd. ** p<0.01, *** p<0.001, ****p<0.0001.

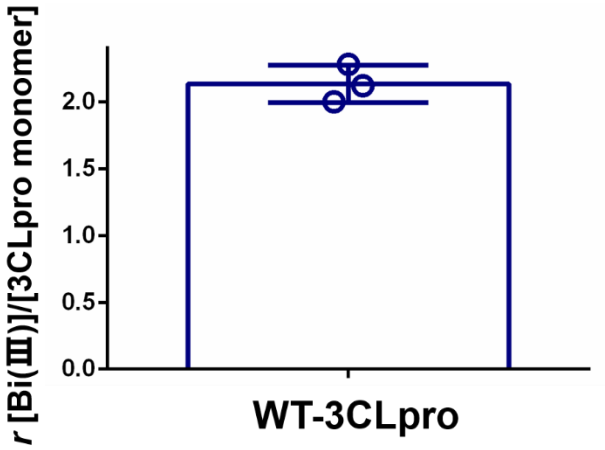


Figure S8. Bar chart shows the Bi(III)-binding capability of 3CLpro. Each experiment were performed in triplicates. The data are shown in mean ± sd.

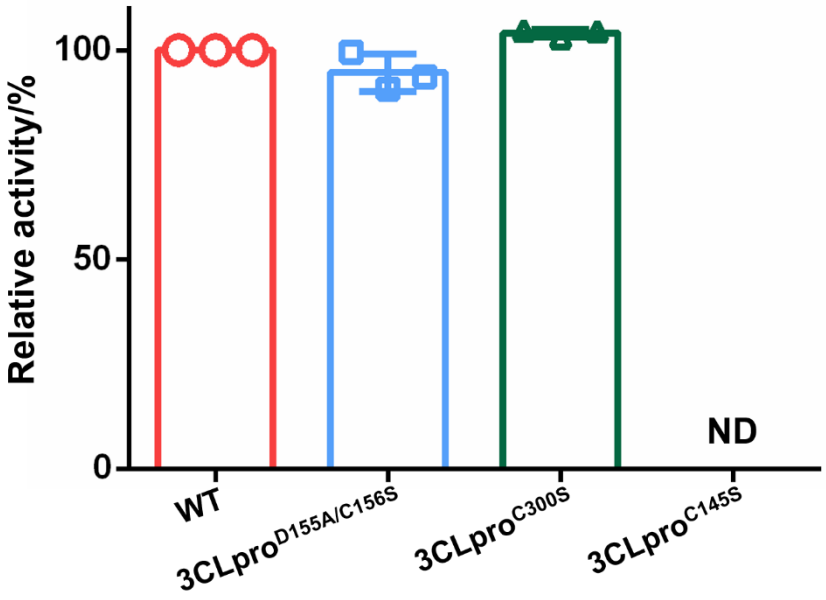


Figure S9. The relative protease activity of 3CLpro mutants. Each experiment were performed in triplicates. The data are shown in mean ± sd.

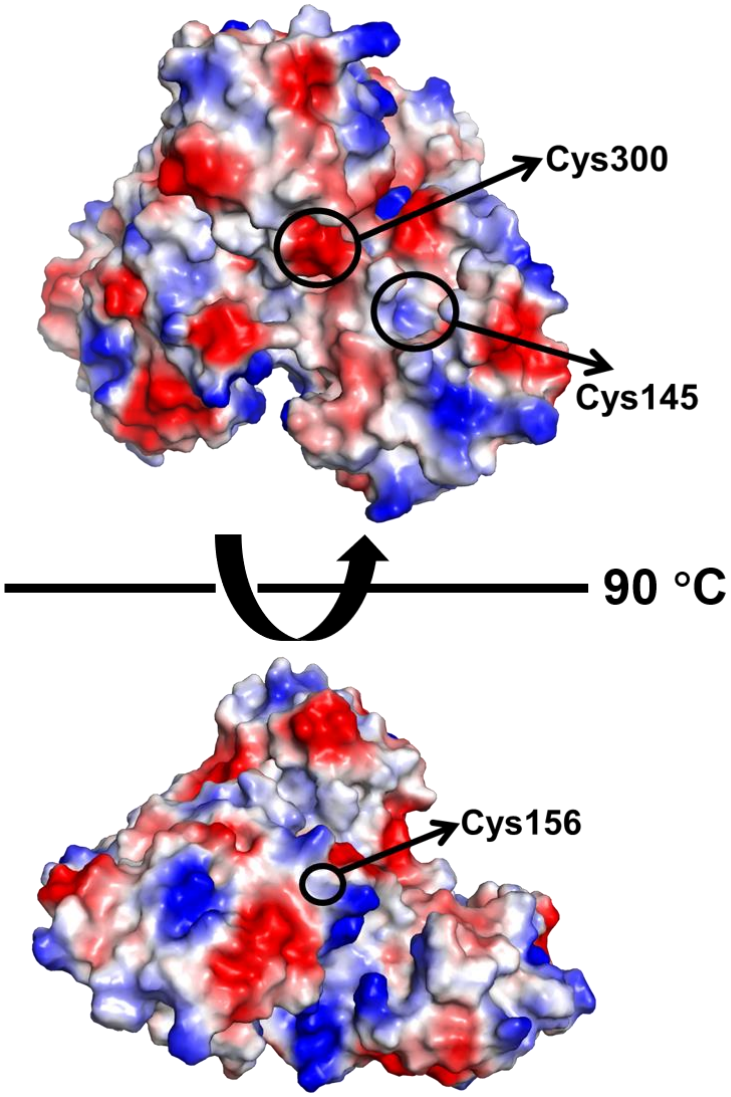


Figure S10. Protein electrostatic potential map of 3CLpro dimer with surface exposed three Cys residues highlighted.

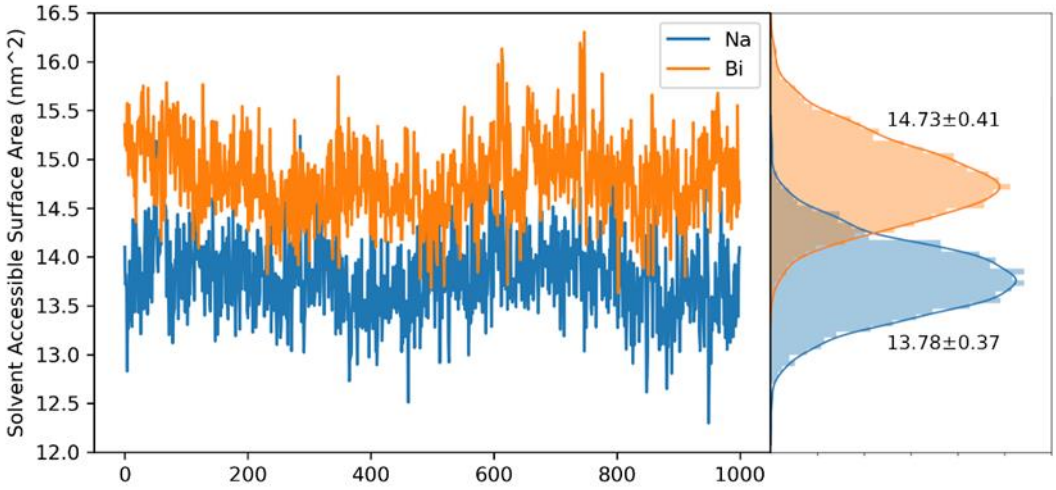


Figure S11. MM/GBSA binding free energy analysis. The binding free energy between the two protomers of 3CLpro is about -75.36 kcal/mol and -52.80 kcal/mol in the presence of Na (I) and Bi(III), respectively.

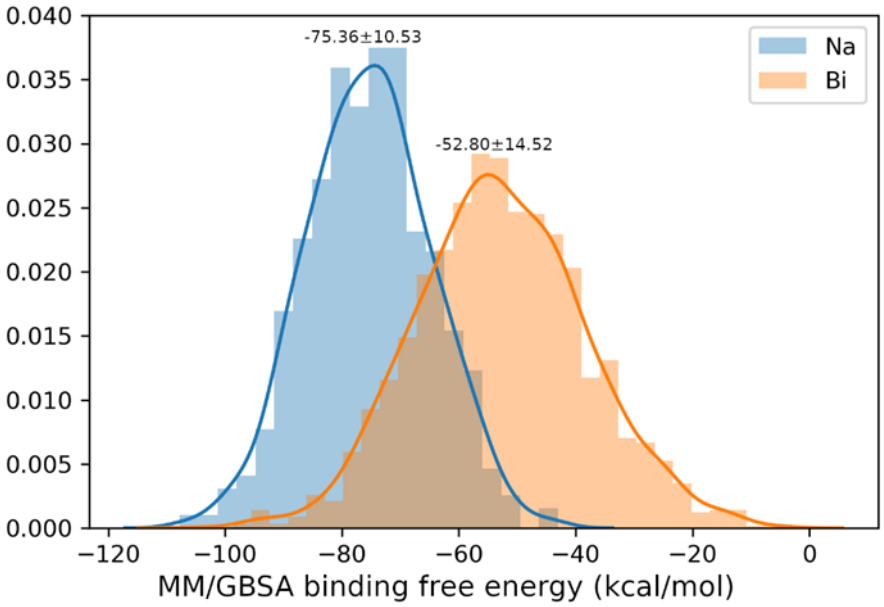


Figure S12. The solvent accessible surface area (SASA) of 3CLpro C-terminal helical segment (residue 292 to 303) in the presence of Bi(III) and Na(I), respectively. The SASA analysis indicated that Bi(III) binding would lead to the C-terminal helical segment more solvent-exposed about 100 $Å^2$ compared to that in the presence of Na(I).

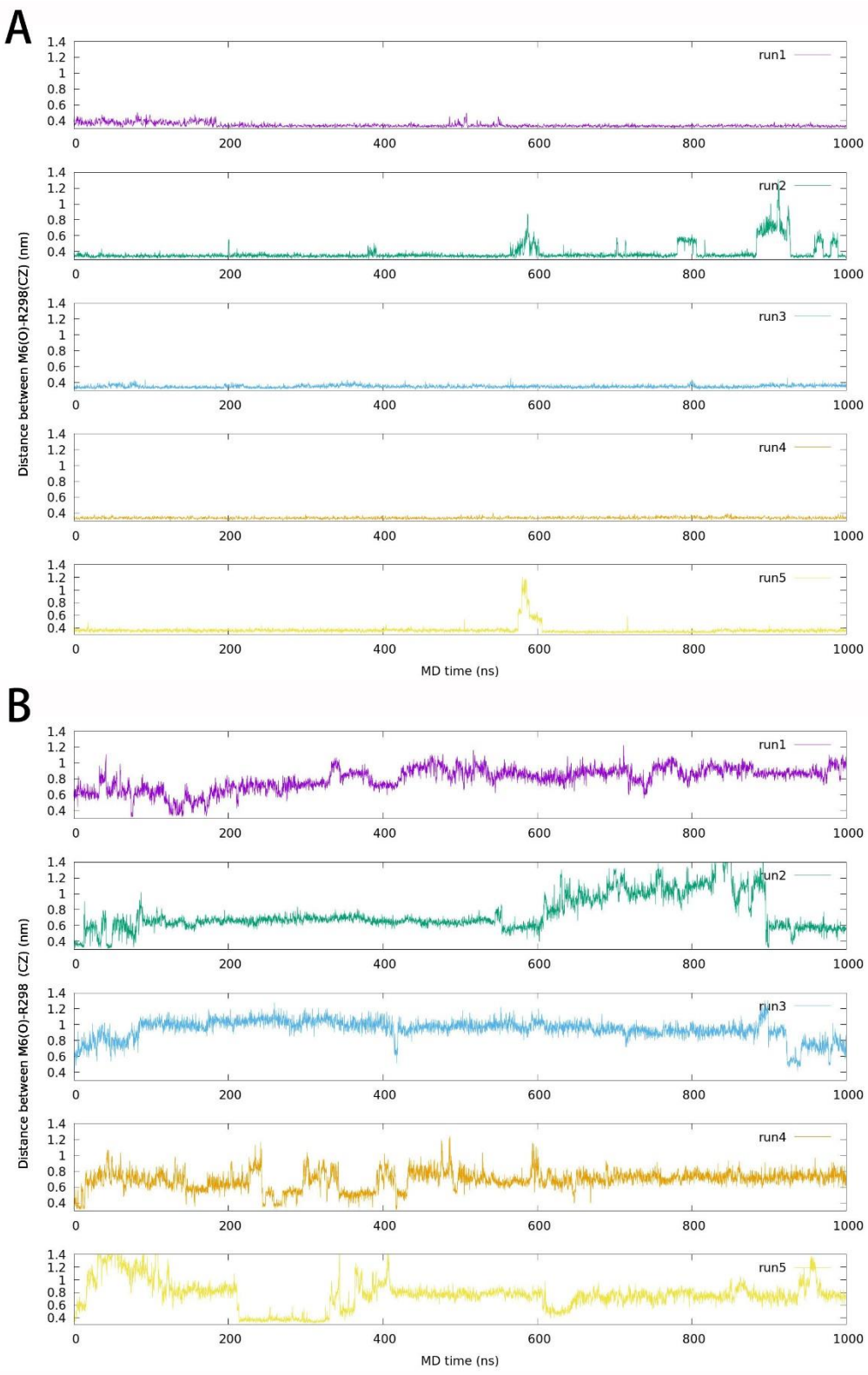


Figure S13. Distances between M6 backbone carbonyl oxygen and CZ atom of R298 of five 1000ns MD trajectories in the presence of (A) Na(I) and (B) Bi(III) ions.

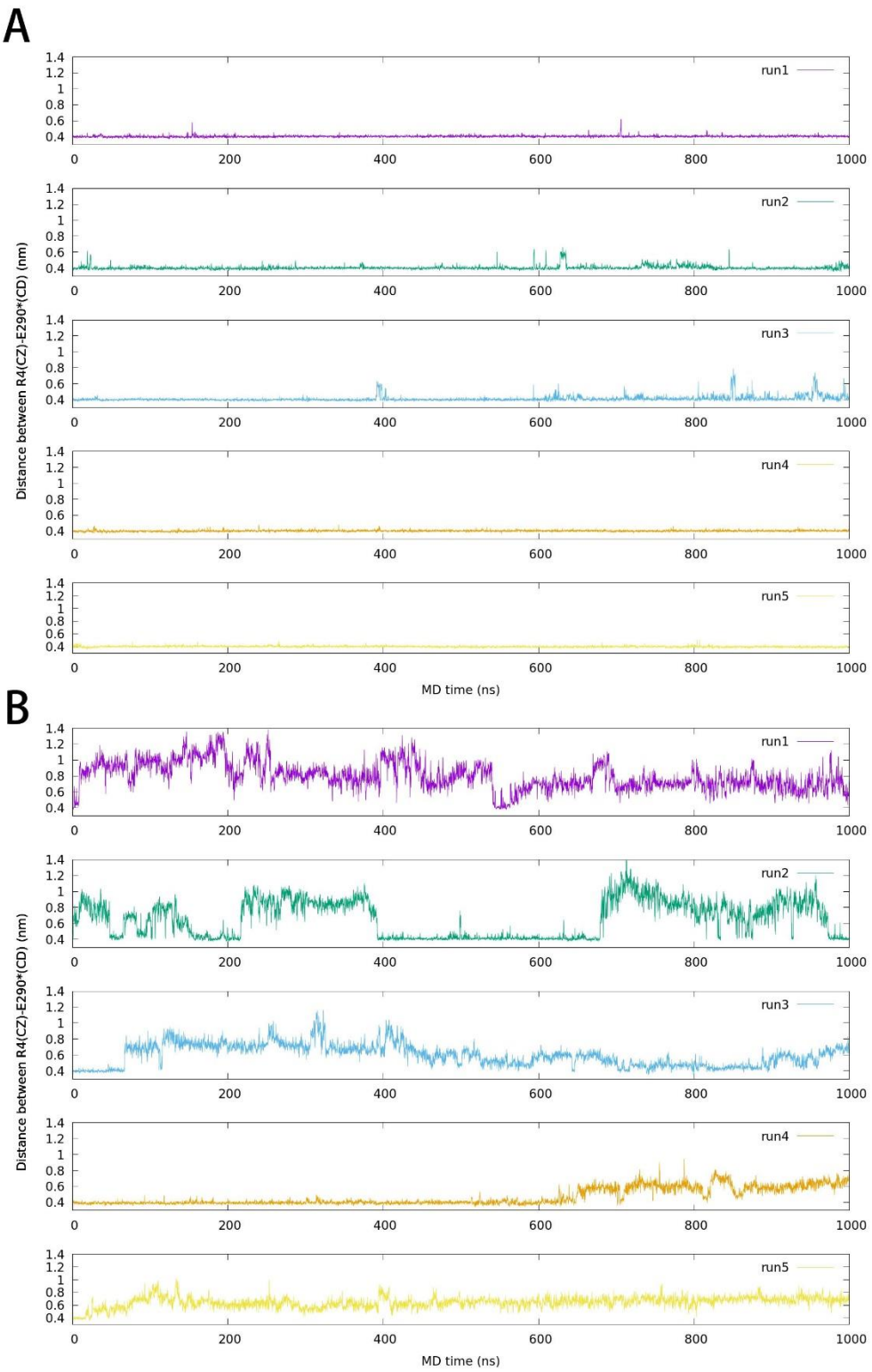


Figure S14. Distances between CZ atom of R4 and CD atom of E290* of five 1000ns MD trajectories in the presence of (A) Na(I) and (B) Bi(III) ions.

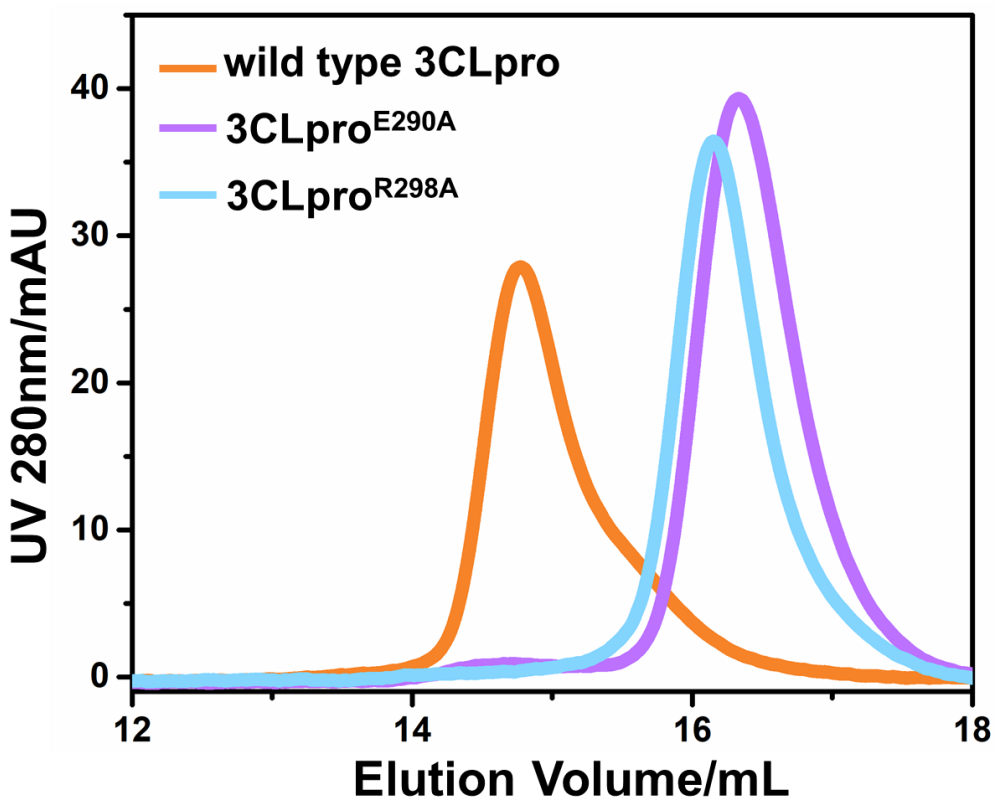


Figure S15. Size-exclusive chromatography analysis of wild type 3CLpro, 3CLpro^{E290A} and 3CLpro^{R298A}.

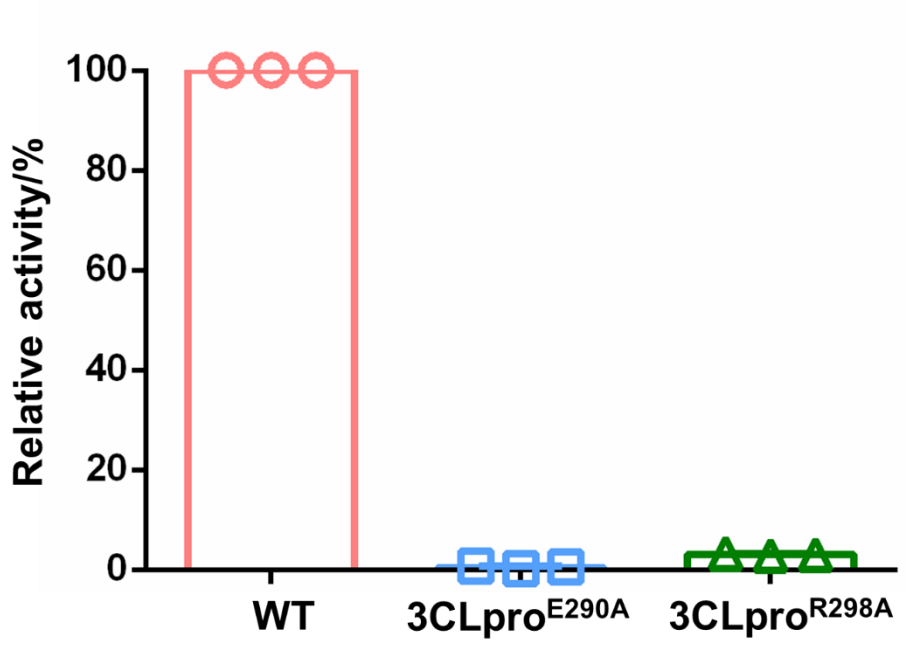


Figure S16. The relative protease activity of 3CLpro^{E290A} and 3CLpro^{R298A} mutants. Each experiment were performed in triplicates. The data are shown in mean ± sd.

Table S1. The primers used for PCR amplification and site-directed mutagenesis

Primer name	**Sequence**
3CLpro_pGEX_BamHI_Forward	CTTGTAGGATCCGCGGTACTGCAGAGC GGTTTCCGTAAGATGGCG
3CLpro_pGEX_XhoI_Reverse	CTTGTACTCGAGTTAGTGGTGGTGGTG GTGGTGGGGTCCTTGAAAGGTCACACC GCT
3CLpro_C145S_Forward	CTGAACGGTAGCAGCGGCAGCGTGG
3CLpro_C145S_Reverse	CCACGCTGCCGCTGCTACCGTTCAG
3CLpro_C156S_Forward	CATTGACTACGATAGCGTTAGCTTTTG
3CLpro_C156S_Reverse	CAAAAGCTAACGCTATCGTAGTCAATG
3CLpro_D155AC156S_Forward	ACATTGACTACGCTAGCGTTAGCTT
3CLpro_D155AC156S_Reverse	AAGCTAACGCTAGCGTAGTCAATGT
3CLpro_C300S_Forward	GTGGTTCGTCAGAGCAGCGGTGTGA
3CLpro_C300S_Reverse	TCACACCGCTGCTCTGACGAACCAC
3CLpro_E290A_Forward	GAGGACGcATTCACCCCGTTTGATGTGG TTCGTCA
3CLpro_E290A_Reverse	GGTGAATgCGTCCTCCAGCAGCGCGCT ACCCAGAA
3CLpro_R298A_Forward	GATGTGGTTgcTCAGTGCAGCGGTGTGA CCTTTCA
3CLpro_R298A_Reverse	CTGAgcAACCACATCAAACGGGGTGAA TTCGTCCT

Table S2. Accessible surface areas scores of 3CLpro cysteines calculated by AREAIMOL software

Site	**Score**
Cys300	9.7
Cys156	8.3
Cys145	7.6
Cys85	5.3
Cys44	1.5
Cys22	1.4
Cys16	0.0
Cys38	0.0
Cys117	0.0
Cys128	0.0
Cys160	0.0
Cys265	0.0

[1] X. Xue, H. Yang, W. Shen, Q. Zhao, J. Li, K. Yang, C. Chen, Y. Jin, M. Bartlam, Z. Rao, *J. Mol. Biol.* **2007**, *366*, 965-975.

[2] H. Yang, W. Xie, X. Xue, K. Yang, J. Ma, W. Liang, Q. Zhao, Z. Zhou, D. Pei, J. Ziebuhr, R. Hilgenfeld, K. Y. Yuen, L. Wong, G. Gao, S. Chen, Z. Chen, D. Ma, M. Bartlam, Z. Rao, *PLoS Biol* **2005**, *3*, e324-e324.

[3] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J Chem Theory Comput* **2015**, *11*, 3696-3713.

[4] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, W. M. Skiff, *J. Am. Chem. Soc.* **1992**, *114*, 10024-10035.

[5] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577-8593.

[6] D. A. B.-S. Case, I. Y.; Brozell, S. R.; Cerutti, D. S.;, T. E. Cheatham, III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.;, D. G. Ghoreishi, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.;, R. H. Harris, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee,, S. L. T. S.; LeGrand, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.;, D. J. M. Mermelstein, K. M.; Miao, Y.; Monard, G.; Nguyen, C.;, H. O. Nguyen, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.;, A. S. Roitberg, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C., J. S.-F. L.; Smith, R.; Swails, J.; Walker, R. C.; Wang, J.;, H. W. Wei, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A., *AMBER 2018; University of California: San Francisco, 2018.*, **2018**.