**Supporting information for**

# Inhibitor recognition specificity of MERS-CoV Papain-Like protease may differ from that of SARS-CoV

Hyun Lee[1,2], Hao Lei[1,2], Bernard D. Santarsiero[2], Joseph L. Gatuz[2], Amy J. Rice, Shuyi Cao[2], Kavankumar Patel[2], Michael Z. Szypulinski[2], Isabel Ojeda, Arun K. Ghosh[3], and Michael E. Johnson[2*]

[1]Equal contributors

[2]Center for Pharmaceutical Biotechnology and Department of Medicinal Chemistry and Pharmacognosy, University of Illinois at Chicago, 900 S. Ashland, IL 60607, USA;

[3]Departments of Chemistry and Medicinal Chemistry, Purdue University, 560 Oval Drive, West Lafayette, IN 47907 (USA)

**Correspondence should be addressed to M.E.J.* (mjohnson@uic.edu)

**Contents**

## 1. Supplementary Materials and Methods

### Cloning, expression, and purification of MERS-PLpro and SARS-PLpro

The MERS-PLpro gene (polyprotein residues 1482-1803) with a human rhinovirus 3C protease cleavage site at the N-terminus was prepared by codon-optimized gene synthesis (BioBasic Inc.) and cloned into a pET15b vector between NdeI and XhoI. Two liters of Rosetta2(DE3) cells containing the recombinant MERS-PLpro gene were grown to an $OD_{600}$ of 0.65 at 37 °C in LB medium. The cells were then induced with 0.5 mM IPTG for 15 hours at 25 °C. Cells were harvested and lysed by sonication in lysis buffer (1 mg $mL^{-1}$ lysozyme and protease inhibitor cocktail in buffer A: 50 mM Tris–HCl (pH 8.0), 500 mM NaCl, 20 mM Imidazole, and 5 mM β-mercaptoethanol). The His-tag fused MERS-PLpro was purified by a HisTrap HP column (GE Healthcare) with a gradient of buffer B (50 mM Tris–HCl (pH 8.0), 500 mM NaCl, 500 mM Imidazole, and 5 mM β-mercaptoethanol). The pooled MERS-PLpro was then dialyzed with buffer A without imidazole and the His-tag was cleaved by 2 units $mg^{-1}$ of HRV 3C protease (Novagen) overnight at 4 °C, producing a MERS-PLpro with only two extra residues (GP) at the N-terminus. Finally, MERS-PLpro was again loaded onto the HisTrap column to remove HRV protease, cleaved His-tags, and uncleaved His-tagged MERS-PLpro. MERS-PLpro was then further purified by Superdex 75 size exclusion column chromatography with a buffer containing 50 mM Tris-HCl (pH 7.5), 200 mM NaCl, and 1 mM TCEP. Purified protein with approximately 95% purity was flash frozen in liquid nitrogen and generated protein beads for storage at –80 °C. Two MERS-PLpro mutants (N109A and N109D) were generated by mutagenesis using Q5® Site-Directed Mutagenesis Kit (New England Biolabs). Both mutants were over-expressed and purified by the same method as the wild-type MERS-PLpro.

The SARS-PLpro gene (polyprotein residues 1541-1855) was cloned, over-expressed, and purified as previously described [1].

**Crystallization**

Protein was prepared for crystallization in 50 mM Tris, 50 mM NaCl and 1mM DTT at pH 8.0. Crystallization conditions were explored by screening nine commercial crystallization screens. Screening was performed by sitting-drop vapor-diffusion in 96-well Greiner plates and three gradient protein concentrations (5.0, 10.0, and 13.0 mg $mL^{-1}$) were tested. MERS-PLpro was crystallized in 1.0 M Sodium Citrate and 0.1 M Sodium Cacodylate at pH 6.5 in a twinned morphology. Optimization was conducted by grid expansion in hanging-drop vapor-diffusion in 24-well plates. Reservoir solutions were prepared in various combinations of Sodium Citrate (0.5-1.5 M, pH 6.5) and Sodium Cacodylate (0.05-0.2 M). Hanging drops were made by mixing 13.0 mg $mL^{-1}$ protein with reservoir solutions at 1:1/1:2/2:1 ratios. Single crystals were picked up and soaked in a reservoir solution with 20% Sucrose (v/v) before being flash-frozen in liquid nitrogen.

**Tested compounds**

Compounds **1 – 6** were repurchased from Life Chemicals. Compound purity was determined by NMR, HPLC, and/or LC/MS to be ≥ 95%. Each confirmed hit compound was repurchased at least twice in separate batches and retested for its inhibitory activity.

**Preparation of antimicrobial focused Life Chemical library**

A custom 25,000 compound drug-like chemical library was constructed as previously described [2].

**Confirmation assay and $IC_{50}$ value determination by dose response curve**

All HTS hit compounds from the primary screens were cherry-picked and reanalyzed by continuous kinetic assay by hand for confirmation. For compounds that showed over 50% inhibition at 50 µM concentration in the confirmation assay, $IC_{50}$ values were determined using the same assay conditions as the primary screen. A series of increasing compound concentrations (0 to 200 µM final concentration at 2-fold serial dilution) in 100% DMSO were prepared. Then 3X (three times final concentration) compound solutions were prepared in assay buffer prior to assays. 8 µL of 3X enzyme solution was distributed into wells, and 8 µL of varying concentration of compounds (3X) were added and incubated for 5 minutes. The enzyme reaction was initiated by adding 8 µL of the substrate (3X), and its activity was continuously monitored for 6 minutes. The $IC_{50}$ values were calculated by fitting with the Hill equation (1), using Sigmaplot 12.0 where y is percent inhibition, x is inhibitor concentration, n is the slope of the concentration–response curve (Hill slope), and $V_{max}$ is maximal inhibition from three to four independent assays.

$$y = V_{max}\left(\frac{x^n}{IC_{50}{}^n + x^n}\right)$$

The enzyme omission assay was done following the same method as $IC_{50}$ determination, but without the PLpro enzyme in order to test for fluorescence signal interference from tested compounds.

**Mechanism of inhibition continued**

Competitive inhibition $$v = \frac{V_{max}}{\left(1+\frac{K_m}{[S]}\right)\left(1+\frac{[I]}{K_i}\right)}$$

Non-competitive inhibition $$v = \frac{V_{max}}{\left(1+\frac{[I]}{K_i}\right)\left(1+\frac{K_m}{[S]}\right)}$$

Uncompetitive inhibition $$v = \frac{V_{max}}{1+\frac{[I]}{K_i}+\frac{K_m}{[S]}}$$

Mixed-type inhibition $$v = \frac{V_{max}}{\left(\frac{K_m}{[S]}\right)\left(1+\frac{[I]}{K_i}\right)+\left(1+\frac{[I]}{\alpha K_i}\right)}$$

where $v$ is the reaction rate, $V_{max}$ is the maximum rate of the reaction, $K_m$ is the Michaelis-Menten constant for the substrate, [S] is the substrate concentration, [I] is the inhibitor concentration, $K_i$ is the dissociation constant of the inhibitor I to the free enzyme and $\alpha K_i$ is the dissociation constant for the inhibitor I to the ES complex. AICc values of four models were ranked, and the best fit equation for each PLpro enzyme was selected [3].

**Reversibility of Inhibition**

Each of the two PLpro enzymes was incubated with screened compounds at 20X the concentration of the $IC_{50}$ for each compound for 1 hour at room temperature in assay buffer containing 50 mM HEPES (pH 7.5), 5 mM DTT, 0.1 mg/ml BSA, 0.01% Triton X-100 (v/v) and 1% DMSO (v/v) in a final volume of 500 µL. Control PLpro enzymes without any compound were also prepared in the same way. Then, each sample including positive control without any compound was buffer exchanged with assay buffer using a desalting column (Pierce) to remove inhibitors. Each PLpro enzyme activity was measured in the same way as $IC_{50}$ measurements before and after buffer exchange.

## 2. Supplementary Tables

**Table S1. Data-collection and refinement statistics for MERS PLpro**

| MERS-CoV PLpro crystal | 4RNA | 4PT5 |
|---|---|---|
| **Data collection** | | |
| Space group | C2 (No. 5) | C2 (No. 5) |
| Cell constants a, b, c, $\beta$ | 98.90Å 48.78Å 87.46Å 121.52° | 100.31Å 47.00Å 87.99Å 121.967° |
| Resolution (Å) | 1.79 (1.89-1.79) | 2.56 (2.72-2.56) |
| Total No. of reflections | 164507 | 44081 |
| No. of averaged reflections (unique reflections) | 33175 | 11133 |
| $R_{merge}$ (%) | 7.9 (134.8) | 8.8 (93.4) |
| CC(1/2) (%) | 99.9 (71) | |
| $\langle I/\sigma(I)\rangle$ | 6.7 (1.2) | 11.9 (1.5) |
| Completeness (%) | 98.9 (94.1) | 97.2 (90.8) |
| **Refinement** | | |
| Resolution range (Å) | 19.73-1.79 | 19.88-2.59 |
| No. of reflections | 31482 | 10484 |
| No. of reflections in test set | 1681 | 521 |
| Completeness (%) | 98.9 | 97.6 |
| $R_{cryst}$ (%) | 0.192 | 0.231 |
| $R_{free}$ (%) | 0.245 | 0.298 |
| Wilson *B* factor (Å$^2$) | 32.4 | 58.6 |
| Average *B* factor (protein) (Å$^2$) | 41.0 | 76.0 |
| No. of protein molecules in asymmetric unit | 1 | 1 |
| **R.m.s.d.s from ideal geometry** | | |
| Bond lengths (Å) | 0.019 | 0.005 |
| Bond angles (°) | 2.209 | 0.872 |
| **Ramachandran plot** | | |
| Favoured (%) | 98 | 92.16 |
| Allowed (%) | 2 | 7.21 |
| Disallowed (%) | 0 | 0.63 |
| No. of solvent molecules | 204 | 22 |

**Table S2. Comparison of kinetic parameters with z-RLRGG-AMC substrates**.

| | $K_M$ (μM) | *kcat* ($s^{-1}$) | *kcat*/ $K_M$ ($M^{-1}s^{-1}$) |
|---|---|---|---|
| SARS-PLpro | 75.9 ± 5.9 | 62.4 | $8.2 \times 10^5$ |
| MERS-PLpro | 142 ± 14 | 2.5 | $1.8 \times 10^4$ |

Both PLpro enzymes were incubated with the substrates in 50 mM HEPES, pH 7.5, 2 mM GSH, 0.1 mg/mL BSA, and 0.01% Triton X-100.

## 3. Supplementary Figures

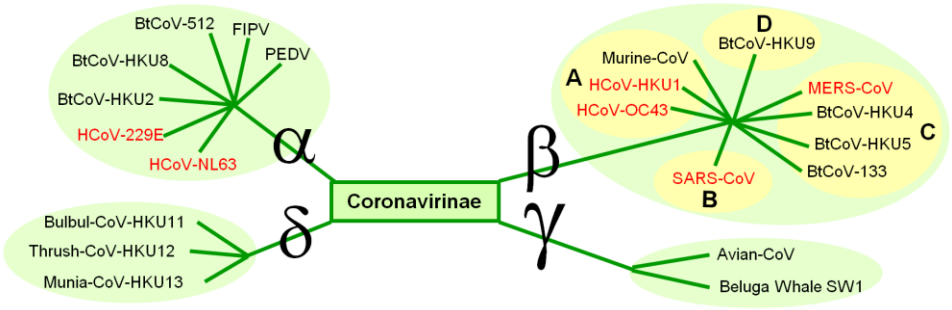



**Figure S1. Classification of coronaviruses and MERS-PLpro.** Classification of a few major coronaviruses in the coronavirinae subfamily divided into four genuses - alphacoronavirus, betacoronavirus, gammacoronavirus, and deltacoronavirus. Further division of the betacoronavirus lineage into subgroups: HCoV (human coronavirus), BtCoV (bat coronavirus), PEDV (porcine epidemic diarrhea virus), FIPV (feline infectious peritonitis virus), SARS (severe acute respiratory syndrome), MERS (Middle East respiratory syndrome).

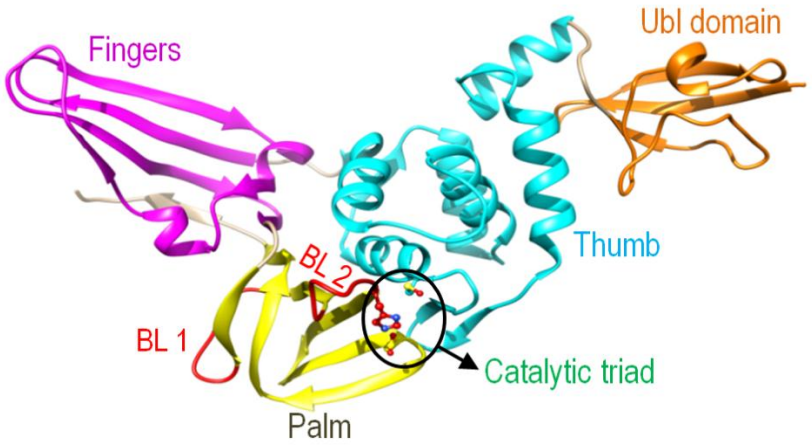



**Figure S2. Crystal structure of MERS-PLpro.** MERS-PLpro crystal structure overview (PDB: 4RNA). The overall structure is composed of four domains indicated by different colors, namely a Ubl-domain (orange), a thumb (cyan), a palm (yellow) and fingers (magenta). The protease catalytic site is located in the interface between the thumb and palm domains, indicated by a black circle.

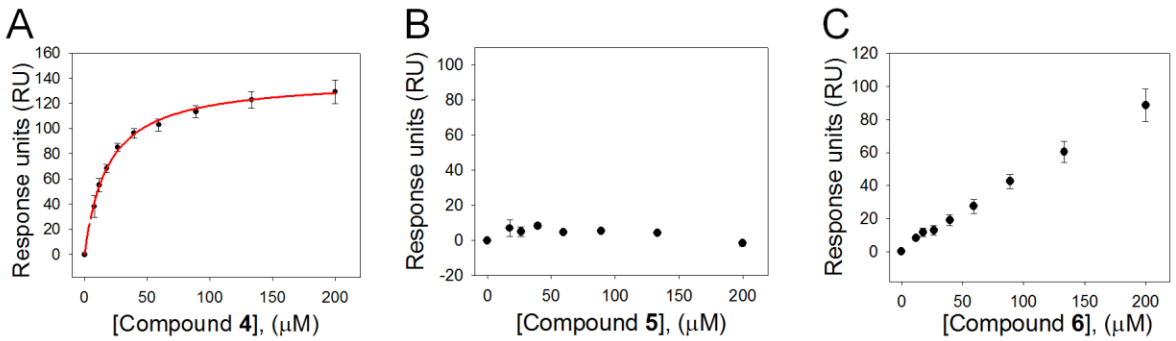
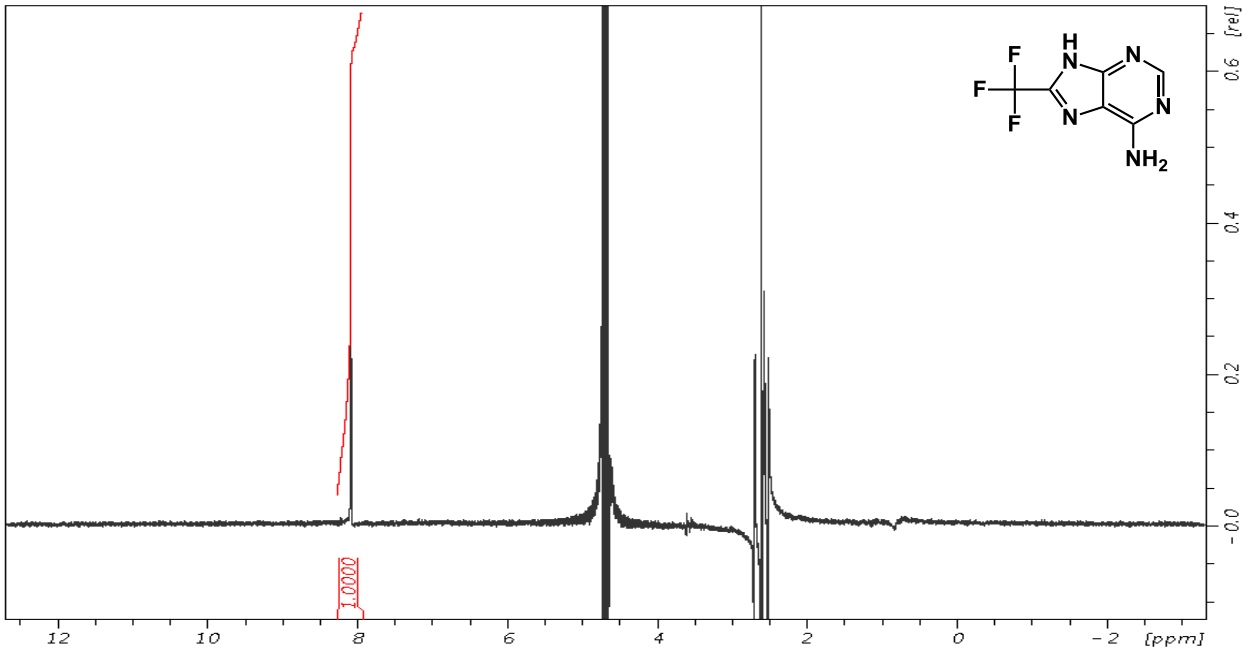


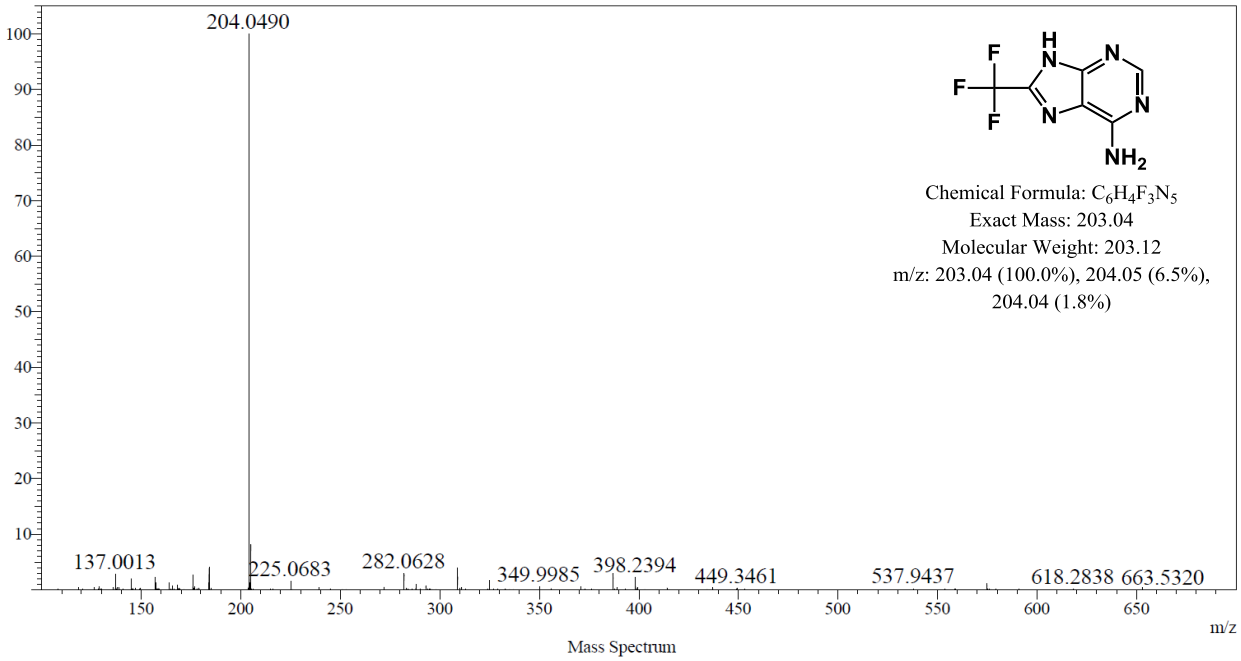

**Figure S3. SPR results.** (*A*) The binding curve of compound **4** fitted to a single rectangular hyperbolic equation (See Methods). The $K_D$ was determined to be 18.4 μM. (*B*) Response units of compound **5** at a series of increasing concentrations (0 – 200 μM), showing no binding to MERS-PLpro. (*C*) Response units of compound **6** at a series of increasing concentrations (0 – 200 μM), showing non-specific binding.

**4. NMR Spectra and MS analysis of our newly identified compound.**

Compounds 1 – 6 were repurchased from Life Chemicals. Compound purity was determined by NMR, HPLC, and/or LC/MS to be ≥ 95%. Each confirmed hit compound was repurchased at least twice in separate batches and retested for its inhibitory activity.

Compound **4** (F2124-0890) was purchased from Life Chemicals. The company provided us quality and purity control data that showed > 95% purity, which was confirmed through NMR and Mass spectra taken at UIC (shown below).[4]








Chemical Formula: $C_6H_4F_3N_5$
Exact Mass: 203.04
Molecular Weight: 203.12
m/z: 203.04 (100.0%), 204.05 (6.5%), 204.04 (1.8%)

| Elmt | Val. | Min | Max | Elmt | Val. | Min | Max | Elmt | Val. | Min | Max | Elmt | Val. | Min | Max | Use Adduct |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H | 1 | 0 | 15 | O | 2 | 0 | 3 | Cl | 1 | 0 | 3 | I | 3 | 0 | 0 | H |
| 2H | 1 | 0 | 0 | F | 1 | 0 | 4 | Fe | 2 | 0 | 0 | Hg | 1 | 0 | 0 | |
| B | 3 | 0 | 0 | Si | 4 | 0 | 0 | Co | 2 | 0 | 0 | | | | | |
| C | 4 | 0 | 18 | P | 3 | 0 | 0 | Cu | 2 | 0 | 0 | | | | | |
| N | 3 | 0 | 7 | S | 2 | 0 | 0 | Br | 1 | 0 | 0 | | | | | |

Error Margin (ppm): 5
HC Ratio: unlimited
Max Isotopes: all
MSn Iso RI (%): 75.00

DBE Range: -1.0 - 30.0
Apply N Rule: yes
Isotope RI (%): 1.00
MSn Logic Mode: AND

Electron Ions: both
Use MSn Info: no
Isotope Res: 10000
Max Results: 25

| Rank | Score | Formula (M) | Ion | Meas. m/z | Pred. m/z | Df. (mDa) | Df. (ppm) | Iso | DBE |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 70.96 | C6 H4 N5 F3 | [M+H]+ | 204.0490 | 204.0492 | -0.2 | -0.98 | 70.96 | 6.0 |

## 5. References


1. Lee, H., et al. Synergistic Inhibitor Binding to the Papain-Like Protease of Human SARS Coronavirus: Mechanistic and Inhibitor Design Implications. *ChemMedChem* **2013,** *8* (8), 1361-72.
2. Lee, H., et al. High-throughput screening (HTS) and hit validation to identify small molecule inhibitors with activity against NS3/4A proteases from multiple hepatitis C virus genotypes. *PLoS One* **2013,** *8* (10), e75144.
3. Burnham, K. P. a. A., D.R. *Model Selection and Inference*. Springer-Verlag: 1998.
4. Lee, H., et al. Identification of novel drug scaffolds for inhibition of SARS-CoV 3-Chymotrypsin-like protease using virtual and high-throughput screenings. *Bioorg Med Chem* **2014,** *22* (1), 167-77.