THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 290, NO. 32, pp. 19403–19422, August 7, 2015

Ligand-induced Dimerization of Middle East Respiratory Syndrome (MERS) Coronavirus nsp5 Protease (3CL^{pro})

*IMPLICATIONS FOR nsp5 REGULATION AND THE DEVELOPMENT OF ANTIVIRALS**

Received for publication, March 11, 2015, and in revised form, June 3, 2015 Published, JBC Papers in Press, June 8, 2015, DOI 10.1074/jbc.M115.651463

Sakshi Tomar[‡1], Melanie L. Johnston[§], Sarah E. St. John[‡], Heather L. Osswald[§], Prasanth R. Nyalapatla[§], Lake N. Paul[¶], Arun K. Ghosh[§], Mark R. Denison[‖], and Andrew D. Mesecar[‡§2]

From the Departments of [‡]Biological Sciences and [§]Chemistry, Purdue University, West Lafayette, Indiana 47907, the [¶]Bindley Bioscience Center, Purdue University, West Lafayette, Indiana 47907, and the [‖]Departments of Pediatrics and Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee 37232

Background: 3CL^{pro} protease is required for coronaviral polyprotein processing and is only active as a dimer.
Results: MERS-CoV 3CL^{pro} is a weakly associated dimer requiring ligand binding for dimer formation.
Conclusion: Ligand-induced dimerization is a key mechanism for regulating the enzymatic activity of MERS-CoV 3CL^{pro} during polyprotein processing.
Significance: Activation via ligand-induced dimerization may add complexity for the development of MERS-CoV 3CL^{pro} inhibitors as antivirals.

All coronaviruses, including the recently emerged Middle East respiratory syndrome coronavirus (MERS-CoV) from the β-CoV subgroup, require the proteolytic activity of the nsp5 protease (also known as 3C-like protease, 3CL^{pro}) during virus replication, making it a high value target for the development of anti-coronavirus therapeutics. Kinetic studies indicate that in contrast to 3CL^{pro} from other β-CoV 2c members, including HKU4 and HKU5, MERS-CoV 3CL^{pro} is less efficient at processing a peptide substrate due to MERS-CoV 3CL^{pro} being a weakly associated dimer. Conversely, HKU4, HKU5, and SARS-CoV 3CL^{pro} enzymes are tightly associated dimers. Analytical ultracentrifugation studies support that MERS-CoV 3CL^{pro} is a weakly associated dimer (K_d ~52 μM) with a slow off-rate. Peptidomimetic inhibitors of MERS-CoV 3CL^{pro} were synthesized and utilized in analytical ultracentrifugation experiments and demonstrate that MERS-CoV 3CL^{pro} undergoes significant ligand-induced dimerization. Kinetic studies also revealed that designed reversible inhibitors act as activators at a low compound concentration as a result of induced dimerization. Primary sequence comparisons and x-ray structural analyses of two MERS-CoV 3CLpro and inhibitor complexes, determined to 1.6 Å, reveal remarkable structural similarity of the dimer interface with 3CL^{pro} from HKU4-CoV and HKU5-CoV. Despite this structural similarity, substantial differences in the dimerization ability suggest that long range interactions by the nonconserved amino acids distant from the dimer interface may control MERS-CoV 3CL^{pro} dimerization. Activation of MERS-CoV 3CL^{pro} through ligand-induced dimerization appears to be unique within the genogroup 2c and may potentially increase the complexity in the development of MERS-CoV 3CL^{pro} inhibitors as antiviral agents.

Coronaviruses (CoVs)[3] are enveloped, positive-strand RNA viruses that infect a variety of vertebrates, including bats, livestock, pets, poultry, and humans (1–3). Although human CoVs cause respiratory illnesses of mild to moderate severity (4–9), two recently emerged CoVs, severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), have demonstrated their potential to become a serious threat to public health. MERS-CoV emerged late in 2012, and unlike its predecessor SARS-CoV, MERS-CoV continues to exhibit up to a 35% fatality rate (10–12).

Based on the sequence analysis of seven genes of the replicase domain, MERS-CoV has been classified as a β-CoV genogroup 2c member, along with closely related bat coronaviruses HKU5 (*Pipistrellus* bat) and HKU4 (*Tylonycteris* bat) (13, 14). Increasing evidence suggests that bats may serve as zoonotic reservoirs for MERS-CoV (15, 16). Evidence presented by recent studies also supports the local zoonotic transmission of MERS-CoV from dromedary camels to humans (17, 18). Alarmingly, human-to-human transmission during close contact, especially in elderly or patients with underlying health conditions, has also been reported for MERS-CoV (19–22). In the wake of the recent upsurge in the laboratory-confirmed cases of MERS-CoV, including two recently identified cases in the United

* This work was supported, in whole or in part, by National Institutes of Health Grants AI08508 (to A. D. M.) and AI026603 (to A. D. M. and M. R. D.). This work was also supported by the Walther Cancer Foundation (to A. D. M.). All authors reviewed the results and approved the final version of the manuscript. The authors declare that they have no conflicts of interest with the contents of this article.

The atomic coordinates and structure factors (codes 4RSP and 4YLU) have been deposited in the Protein Data Bank (http://wwpdb.org/).

[1] Supported by a grant from the Purdue Research Foundation.

[2] To whom correspondence should be addressed: Dept. of Biological Sciences, Purdue University, 915 W. State St., West Lafayette, IN 47907. Tel.: 765-494-1924; E-mail: amesecar@purdue.edu.

[3] The abbreviations used are: CoV, coronavirus; MERS, Middle East respiratory syndrome; SARS, severe acute respiratory syndrome; nsp, nonstructural protein; 3CL^{pro}, 3-chymotrypsin-like protease; AUC, analytical ultracentrifugation; SV, sedimentation velocity; BME, β-mercaptoethanol; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; PDB, Protein Data Bank.

States (23), there is an urgent need to study and characterize the properties of important drug targets of MERS-CoV for the development of effective therapeutics.

Coronaviruses express a >800-kDa replicase polyprotein, which is processed by viral 3CL^{pro} protease (or nsp5) at 11 distinct cleavage sites to yield intermediate and mature nonstructural proteins (nsp) responsible for many aspects of virus replication (3, 24–26). Because of its indispensable role in the virus life cycle, 3CL^{pro} is an important target for therapeutic intervention against coronavirus infections (27–33).

A number of kinetic, biophysical, and x-ray structural studies have demonstrated that SARS-CoV 3CL^{pro} is only active *in vitro* as a tightly associated dimer with a dimer dissociation constant (K_d) in the low nanomolar range (34–38). The addition or deletion of amino acids, *e.g.* His_6 affinity tags, at either the N or C terminus drastically reduces the enzymatic rate and decreases the ability of SARS-CoV 3CL^{pro} to dimerize (37). Although cellular evidence for the auto-cleavage mechanism (*cis versus trans*) of 3CL^{pro} is lacking, models for how 3CL^{pro} cleaves itself from the polyprotein to form the mature dimer have been proposed based on *in vitro* studies using purified 3CL^{pro} (34, 39, 40). A current model posits that two inactive 3CL^{pro} molecules within two separate polyproteins recognize each other and form an immature dimer capable of cleaving the nsp4 ↓ nsp5 and nsp5 ↓ nsp6 sites in *trans*, followed by formation of an active and mature dimer that can then rapidly process other cleavage sites and multiple polyproteins. It has also been proposed that substrate-induced dimerization regulates the enzymatic activity of SARS-CoV 3CL^{pro} during virus replication; however, no experimental evidence of this has ever been demonstrated in infected cells (40). Although our knowledge of SARS-CoV 3CL^{pro} is extensive, the dimerization properties of 3CL^{pro} from MERS-CoV and other coronaviruses, as well as the factors regulating their enzymatic activity, remain largely unknown.

To understand the properties of MERS-CoV 3CL^{pro}, we conducted a series of kinetic, biophysical and x-ray structural studies. Here, we report a detailed kinetic and biophysical analysis of MERS-CoV 3CL^{pro} activity and dimerization. These kinetic and biophysical studies provide evidence for a weakly associated MERS-CoV 3CL^{pro} dimer. In addition, we utilized our previous knowledge on the design of potent SARS-CoV 3CL^{pro} peptidic inhibitors to design a series of inhibitors of MERS-CoV 3CL^{pro} that exhibit low micromolar potency. We demonstrate that MERS-CoV 3CL^{pro} requires the binding of a ligand for dimer formation, indicating that ligand-induced dimerization is likely a key mechanism in the regulation of MERS-CoV 3CL^{pro} activity during virus infection.

Experimental Procedures

Construct Design and Expression of MERS-CoV 3CL^{pro}—The gene encoding 3CL^{pro} protease of MERS-CoV (amino acid residues 3248–3553 in the replicase polyprotein, GenBank™ accession number AHC74086.1) was codon-optimized for optimal expression in *E. coli* (BioBasic Inc). The gene was subcloned into pET-11a expression vector with an N-terminal His_6 tag followed by the nsp4 ↓ nsp5 auto-cleavage site using the forward primer 5′-ATATACATATGCACCACCACCACCACCACAGCGGTGTTCTGCAGTCTGGTC-3′ and the reverse primer 5′-GACGGATCCTTACTGCATCACAACACCCATGATCTGC-3′. The construct was verified by DNA sequencing at the Purdue University Genomics Core Facility. This construct results in the expression of MERS-CoV 3CL^{pro} without any N- or C-terminal extensions. MERS-CoV 3CL^{pro} was expressed through auto-induction in *Escherichia coli* BL21-DE3 cells in the presence of 100 μg/ml carbenicillin as described previously (41). Cells were harvested by centrifugation at 5000 × *g* for 20 min at 4 °C, and the pellets were stored at −80 °C until further use.

MERS-CoV 3CL^{pro} Purification—Frozen pellets from 4 liters of bacterial cell culture were thawed on ice and resuspended in 250 ml of Buffer A (20 mM Tris, pH 7.5, 0.05 mM EDTA, 10% glycerol, and 5 mM β-mercaptoethanol (BME)), containing 500 μg of lysozyme and a small amount of DNase. Cells were then lysed using a single pass through a French press at 1200 p.s.i., and cell debris was removed from the cleared lysate by centrifuging at 29,000 × *g* for 30 min. Solid ammonium sulfate was added to the cleared lysate to a final concentration of 1 M through gradual mixing on ice.

Hydrophobic Interaction Chromatography—The cleared lysate, mixed with ammonium sulfate, was loaded at a flow rate of 3 ml/min onto a 60-ml phenyl-Sepharose 6 fast-flow high-sub column (Amersham Biosciences) equilibrated with Buffer B (50 mM Tris, pH 7.5, 1 M ammonium sulfate, 0.05 mM EDTA, 10% glycerol, and 5 mM BME). The column was then washed with 5× column volume (300 ml) of Buffer B at a flow rate of 4 ml/min. Protein was eluted using a 5× column volume (300 ml) linear gradient to 100% Buffer A. Fractions (12 ml) were collected, and those containing MERS-CoV 3CL^{pro}, as judged through SDS-PAGE analysis and specific activity measurements, were pooled (120 ml) and exchanged into 2 liters of Buffer A via overnight dialysis in a 10,000 molecular weight cutoff SnakeSkin® dialysis tubing (Thermo Scientific).

DEAE Anion-exchange Chromatography—Dialyzed sample from the previous step was loaded at a flow rate of 3 ml/min onto a 120- ml DEAE anion-exchange column (Amersham Biosciences) equilibrated with Buffer A. The column was then washed with 2× column volume (240 ml) of Buffer A at a flow rate of 4 ml/min. A linear gradient (total volume 480 ml) to 40% Buffer C (50 mM Tris, pH 7.5, 1 M NaCl, 0.05 mM EDTA, 10% glycerol, and 5 mM BME) was used to elute the protein. Fractions (6 ml) were collected, and those containing MERS-CoV 3CL^{pro} were pooled (66 ml) and dialyzed for 4 h in 4 liters of Buffer D (20 mM MES, pH 5.5, 0.05 mM EDTA, 10% glycerol, and 5 mM BME).

Mono S Cation-exchange Chromatography—Following dialysis, the pH of the sample was manually adjusted to 5.5 using 1 M solution of MES, pH 5.5, and any precipitated protein was removed by filtering through a 0.22-μm pore size Millex-GP filter (Millipore). The filtered sample was then loaded at a flow rate of 2 ml/min onto an 8-ml Mono S 10/100 column (Amersham Biosciences) equilibrated in Buffer D. The column was then washed with 5× column volume (40 ml) of Buffer D at a flow rate of 2 ml/min. Protein was eluted using a 25× column volume (200 ml) and a linear gradient to 50% Buffer E (50 mM MES, pH 5.5, 1 M NaCl, 0.05 mM EDTA, 10% glycerol, and 5 mM

BME). Fractions (2 ml) were collected, and those containing MERS-CoV 3CL^{pro} were pooled (22 ml) and concentrated to ~5 mg/ml.

Gel Filtration Chromatography—As the final purification step, the concentrated protein sample was loaded onto the preparation grade Superdex 75 26/60 gel filtration column (Amersham Biosciences) equilibrated with Buffer F (25 mM HEPES, pH 7.5, 10% glycerol, 2.5 mM dithiothreitol (DTT)). Protein was eluted isocratically at a flow rate of 1 ml/min with Buffer F. Fractions (2 ml) containing MERS-CoV 3CL^{pro} were pooled (total volume of 34 ml) and concentrated to ~5 mg/ml. For final storage of the purified MERS-CoV 3CL^{pro} enzyme, 300-μl protein aliquots were placed into 1-ml screw-cap vials, flash-frozen under liquid nitrogen, and then stored at −80 °C until further use.

Purification of SARS-CoV, HKU4-CoV, and HKU5-CoV 3CL^{pro}—SARS-CoV 3CL^{pro} and HKU5-CoV 3CL^{pro} with authentic N and C termini were expressed and purified as described previously (37, 42). HKU4-CoV 3CL^{pro} was purified utilizing a modified protocol from Ref. 42. Final protein yield was calculated based on the measurement of total activity units (μM product/min), specific activity (units/mg), and milligrams of protein obtained (Bio-Rad protein assay) after each chromatographic step.

*Synthesis of Compounds **1–11***—The peptidomimetic compounds with Michael acceptor groups (compounds **1**-**9**, Table 3) were synthesized via very similar methods to those published previously (30, 43). Synthesis of noncovalent peptidomimetic compounds **10** and **11** (Table 3) has been described previously (33).

Fluorescence-based Kinetic Assays—The enzymatic activity of 3CL^{pro} was measured using the following custom-synthesized peptide: HilyteFluor^{TM}-488-ESATLQSGLRKAK-(QXL^{TM}-520)-NH_2 (AnaSpec, Inc.). The HilyteFluor^{TM}-488 fluorescence group was internally quenched by QXL^{TM}-520 dye. This substrate works as a generic peptide substrate for 3CL^{pro} enzymes and was designed based on the nsp4 ↓ nsp5 cleavage sequence for many coronavirus 3CL^{pro} enzymes. The rate of enzymatic activity was determined at 25 °C by following the increase in fluorescence ($\lambda_{excitation}$ = 485 nm, $\lambda_{emission}$ = 528 nm, bandwidths = 20 nm) of Hilyte Fluor^{TM}-488 upon peptide hydrolysis by the enzyme as a function of time. Assays were conducted in black, half-area, 96-well plates (Corning Glass) in assay buffer (50 mM HEPES, pH 7.5, 0.1 mg/ml BSA, 0.01% Triton X-100, and 2 mM DTT) using a final reaction volume of 100 μl. The resulting florescence was monitored using a BioTek Synergy H1 plate reader. The rate of the reaction in arbitrary fluorescence units/s (AFU/s) was determined by measuring the initial slope of the progress curves, which were then converted to units of micromolars of product produced per min (μM/min) using experimentally determined values of fluorescence "extinction coefficient" as described previously (37). All reactions were carried out in triplicate.

Determination of Enzymatic Efficiency—The apparent enzymatic efficiency for each of the 3CL^{pro} enzymes was determined by measuring the rate of enzymatic activity as a function of varying substrate concentration in 100-μl reactions. Reactions were initiated by the addition of enzyme to the wells of an assay plate containing varying concentrations of substrate. The final substrate concentrations varied over a range from 0 to 2 μM. The final enzyme concentrations for each 3CL^{pro} studied were as follows: MERS-CoV 3CL^{pro} at 1 μM, SARS-CoV 3CL^{pro} at 100 nM, HKU5-CoV 3CL^{pro} at 250 nM, and HKU4-CoV 3CL^{pro} at 200 nM. Because 3CL^{pro} enzymes cannot be saturated with this substrate at a substrate concentration that would still allow accurate fluorescent measurements without the inner filter effect, only the apparent k_{cat}/K_m values can be determined from the slope of the line that results from a plot of the enzymatic activity (y axis), normalized for the total enzyme concentration, against the substrate concentration (x axis).

Influence of Dimerization on the Activity of 3CL^{pro} Enzymes—The dependence of the enzymatic activity on the total enzyme concentration was determined using the FRET-based assay described above. The final enzyme concentrations were varied over a concentration range from 2 μM to 100 nM for MERS-CoV 3CL^{pro}, 500 to 10 nM for SARS-CoV 3CL^{pro}, 250 to 0.6 nM for HKU5-CoV 3CL^{pro}, and 200 to 10 nM for HKU4-CoV 3CL^{pro}. Reactions were initiated by the addition of substrate, at a final concentration of 2 μM, to the assay plates containing varying enzyme concentrations in the assay buffer. Initial rates were determined from the initial slopes of the progress curves at each enzyme concentration.

The rates of the 3CL^{pro}-catalyzed reactions measured over a range of enzyme concentrations can be fit to either Equation 1 or 2 to determine the values of the dissociation constant for the monomer-dimer equilibrium as well as the turnover numbers. Nonlinear regression and the program TableCurve 2D version 4.0 were used to fit the data to either Equation 1 or 2 below (44).

$$V_{max} = k_{cat,M}\frac{-K_d + \sqrt{K_d^2 + 8K_dC_T}}{4} + k_{cat,D}\frac{K_d + 4C_T - \sqrt{K_d^2 + 8K_dC_T}}{8} \quad \text{(Eq. 1)}$$

In Equation 1, V_{max} is the rate of the enzymatic activity calculated at each enzyme concentration (C_T); K_d is the monomer-dimer equilibrium dissociation constant, and $k_{cat,\,M}$ and $k_{cat,\,D}$ are the turnover numbers for the monomer and the dimer, respectively.

$$V_{max} = k_{cat}[D] = k_{cat}\frac{K_d + 4C_T - \sqrt{K_d^2 + 8K_dC_T}}{8} \quad \text{(Eq. 2)}$$

In Equation 2, V_{max}, C_T, and K_d have been described previously, and k_{cat} is the turnover number for the dimer only.

Inhibition Assays—To determine the percent inhibition for compounds **1–9**, the total concentration of the substrate was fixed at 1.0 μM, and the enzymes were fixed at 250 nM for SARS-CoV 3CL^{pro}, HKU5-CoV 3CL^{pro}, HKU4-CoV 3CL^{pro}, and at 500 nM for MERS-CoV 3CL^{pro}. DMSO stocks (100×) of the compounds were diluted a hundred-fold to a final concentration of 50 μM in 80 μl of the enzyme solution and incubated for 20 min. After 20 min, the enzymatic activity was measured as initial slope of the progress curve, obtained by initiating the reaction with 20 μl of 5 μM substrate. % inhibition was calculated using Equation 3.

$$\% \text{ inhibition} = \left(1 - \frac{(\text{rate}_{\text{sample}} - \text{rate}_{\text{neg}})}{(\text{rate}_{\text{pos}} - \text{rate}_{\text{neg}})}\right) \times 100 \quad \text{(Eq. 3)}$$

In Equation 3, $\text{rate}_{\text{sample}}$ is the initial slope of the progress curve in AFU/s measured in the presence of the compound; rate_{pos} is the initial slope measured in the absence of any compound, and rate_{neg} is the baseline substrate hydrolysis calculated in the absence of enzyme. All the reactions were carried out in triplicate and contained a final DMSO concentration of 1%. For compounds displaying more than 50% inhibition, a more extensive characterization of the inactivation kinetics was performed through progress curve analysis. To the reaction well, 20 μl of 5 μM substrate was added to a final concentration of 1 μM, and the total inhibitor concentration $[I]_{\text{total}}$ was varied from 0 to 50 μM. The reaction was initiated with the addition of 80 μl of MERS-CoV 3CL^{pro} to a final concentration of 500 nM. Fluorescence intensity was then measured over time as AFU_t for a period of 70 min. Equation 4 describes the resulting time course of reaction.

$$[P]_t = \frac{v_i}{k_{\text{obs}}}(1 - \exp(-k_{\text{obs}}t) + [P]_i \quad \text{(Eq. 4)}$$

In Equation 4, v_i is the initial velocity of the reaction; k_{obs} is the observed first-order rate constant for the reaction in the absence and presence of inhibitor; t is the time in minutes; $[P]_t$ is the concentration of product produced at time t, and $[P]_i$ is the initial product concentration, which is zero. Product concentrations were calculated from the values of AFU_t, using the experimentally determined fluorescence extinction coefficient. The resulting values of $[P]_t$ were then plotted against time t, and the data were fit to Equation 4 with $[P]_i = 0$ using the nonlinear regression program TableCurve 2D to derive the fitted parameters v_i and k_{obs} and their associated errors Δv_i and Δk_{obs}.

Values for each k_{obs} were then plotted against $[I]_{\text{total}}$ and the data were fit to Equation 5.

$$k_{\text{obs}} = \frac{k_{\text{inact}}[I]_{\text{total}}}{K_I + [I]_{\text{total}}} \quad \text{(Eq. 5)}$$

In Equation 5, k_{inact} defines the maximum rate of inactivation at infinite inhibitor concentration, and K_I defines the concentration of inhibitor that yields a rate of inactivation equal to $\frac{1}{2}k_{\text{inact}}$. The half-life of inactivation at infinite inhibitor concentration, which is a measure of inactivation efficiency, is defined as $t^{\infty}_{1/2} = 0.693/k_{\text{inact}}$.

AUC Analysis—To determine the oligomeric state of MERS-CoV 3CL^{pro}, sedimentation velocity experiments were performed at 20 °C on the Beckman-Coulter XLA ultracentrifuge using varying concentrations of MERS-CoV 3CL^{pro} (4–23 μM) in 25 mM HEPES, pH 7.5, 50 mM NaCl, and 1 mM tris(2-carboxyethyl)phosphine at 50,000 rpm. To characterize the effect of the ligand on the monomer-dimer equilibrium of MERS-CoV 3CL^{pro}, sedimentation velocity experiments were conducted on the Beckman-Coulter XLI instrument using different stoichiometric ratios of MERS-CoV 3CL^{pro} with compounds **6** and **10**. Samples were prepared by mixing 25 μM MERS-CoV 3CL^{pro} with 25, 50, and 100 μM compound **6** or **10** and incubating the mixture overnight at 4 °C before performing the experiments. Absorbance optics (280 nm) and interference optics were utilized for protein detection. Solvent density, viscosity, and partial specific volumes were calculated using SEDNTERP. SEDPHAT was used to fit the data to the monomer-dimer self-association model to estimate the sedimentation coefficients (s), apparent molecular weights, and K_d and k_{off} values from size distribution analysis. To obtain exact molecular weights, sedimentation equilibrium experiments were performed at concentrations of 3 and 17 μM MERS-CoV 3CL^{pro}. The experiments were done at 20 °C utilizing a two-channel centerpiece and run at multiple speeds (8100, 13,800 and 24,000 rpm) in a AN-60 Ti rotor.

MERS-CoV 3CL^{pro} Activation and Inhibition by a Noncovalent Inhibitor—The rates of the MERS-CoV 3CL^{pro}-catalyzed reactions were determined at final enzyme concentrations of 0.5, 1.0, and 2.0 μM and in the absence and presence of varying concentrations (0.1–60 μM) of compound **10**. The substrate concentration was fixed at 2.0 μM. DMSO stocks (100×) of compound **10** were diluted a hundred-fold in 80 μl of enzyme solution and incubated for 10 min. At the same time, a zero-inhibitor control reaction was set up by mixing DMSO to a final concentration of 1% into 80 μl of enzyme solution. After 10 min, the rate of the enzymatic activity was measured as the initial slope of the progress curve, obtained by initiating the reaction with 20 μl of 10 μM substrate. Equation 6 was utilized to calculate the percent activity.

$$\% \text{ activity} = \frac{(\text{rate}_{\text{sample}} - \text{rate}_{\text{neg}})}{(\text{rate}_{\text{pos}} - \text{rate}_{\text{neg}})} \times 100 \quad \text{(Eq. 6)}$$

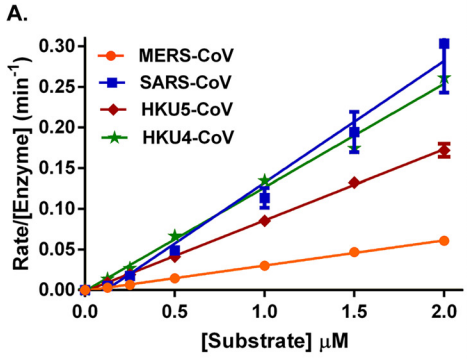
The $\text{rate}_{\text{sample}}$, rate_{pos}, and rate_{neg} are as described above for Equation 3.

MERS-CoV 3CL^{pro} Crystallization, X-ray Data Collection, and Structure Determination—Purified MERS-CoV 3CL^{pro} was concentrated to 1.6 mg/ml in 25 mM HEPES, pH 7.5, and 2.5 mM DTT. Inhibitor complexes of MERS-CoV 3CL^{pro} with compounds **6** and **11** were formed by incubating MERS-CoV 3CL^{pro} with the compounds in a 1:3 stoichiometric ratio at 4 °C overnight. After iterative rounds of optimization of the crystallization conditions based on the initial hits obtained from high throughput screening of Qiagen Nextel Screens, crystals of MERS-CoV 3CL^{pro} inhibitor complexes suitable for x-ray diffraction were grown by the hanging-drop, vapor diffusion method at 20 °C in 0.2 M sodium acetate, 0.1 M BisTris, pH 7.0, and 20% PEG-3350 for the MERS-CoV 3CL^{pro} and **6** complex, and 0.2 M ammonium acetate, 0.1 M BisTris, pH 5.5, 12% PEG-3350 for the MERS-CoV 3CL^{pro} and **11** complex. For x-ray data collection, crystals were flash-cooled in liquid nitrogen after dragging the crystals through a cryo-solution that contained the crystallization solution supplemented with 15% 2-methyl-2,4-pentanediol.

X-ray diffraction data were collected for MERS-CoV 3CL^{pro} and **6** and MERS-CoV 3CL^{pro} and **11** complexes at the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) Sector 31 and the Life Sciences Collaborative Access Team (LS-CAT) Sector 21 at the Advanced Photon Source, Argonne

TABLE 1
Purification summary of MERS-CoV 3CL^{pro} per liter of *E.coli* BL21-DE3 cells

Sample	Protein	Total activity units	Specific activity	Fold purification	% yield
	mg		*units/mg*		
Lysate	1102	1168	1	1	100
Phenyl-Sepharose	219	185	1	1	16
DEAE	22	189	8	8	16
Mono S	15	142	9	9	12
Superdex 75	12	114	10	10	10



B.

3CL^{pro} Enzyme	*k_{cat}/K_M (× 10^2) ($\mu M^{-1}.min^{-1}$)
MERS-CoV	3.1 ± 0.03
SARS-CoV	15.5 ± 0.9
HKU5-CoV	8.8 ± 0.1
HKU4-CoV	11.3 ± 0.4

FIGURE 1. **Comparison of enzymatic efficiencies (k_{cat}/K_m) of 3CL^{pro} enzymes from different CoVs.** *A,* rates for the enzymatic activity, normalized to the total enzyme concentration, are plotted as a function of varying substrate concentrations. Total concentration of each enzyme in the final reaction is as follows: MERS-CoV 3CL^{pro} at 1 μM; SARS-CoV 3CL^{pro} at 100 nM; HKU5-CoV 3CL^{pro} at 250 nM; and HKU4-CoV 3CL^{pro} at 200 nM. Slope of the *line* represents the apparent value of k_{cat}/K_m. *Error bars* represent the standard deviation for triplicate data. *B,* *, apparent value of k_{cat}/K_m for the nonsaturable substrate, calculated as the slope of the linear plot from panel *A*.

National Laboratory, respectively. Data were processed and scaled using Mosflm version 7.0.5 (45) and HKL2000 version 706 (46). The method of molecular replacement was used to obtain initial phases using the program PHASER-MR in Phenix suite version 1.8.4 (47). For MERS-CoV 3CL^{pro} and **6** complex, the x-ray structure of SARS-CoV 3CL^{pro} (PDB code 3V3M) was used as a phasing model (32). The final MERS-CoV 3CL^{pro} and **6** complex structure was then used to calculate the initial phases for the MERS-CoV 3CL^{pro} and **11** complex model. Automated model building using Autobuild in Phenix was initially used to build a preliminary model of the MERS-CoV 3CL^{pro} and **6** inhibitor complex. Each structure was then refined using iterative cycles of refinement using Phenix Refine coupled to manual model building using COOT (48) based on $F_o - F_c$ and $2F_o - F_c$ maps. Coordinates and molecular library files for inhibitor molecules were built using the program eLBOW in the Phenix suite. Water molecules were added to peaks in residual ($F_o - F_c$) density maps that were greater than 3σ using the "Find Water" function in COOT. MolProbity was used to assess structural quality of the final model (49). The measured structure factor amplitudes and the atomic coordinates for the final structures were deposited in the Protein Data Bank with accession codes 4RSP (MERS-CoV 3CL^{pro} and **6** complex) and 4YLU (MERS-CoV 3CL^{pro} and **11** complex), respectively. Structural superposition was performed using the method of least squares fitting of C-α atoms in COOT. PyMOL was used to generate figures of all the structures (50).

Results

Production of MERS-CoV 3CL^{pro} with Authentic N and C Termini—Insertion of the nsp4↓nsp5 cleavage site between the N-terminal His_6 tag and the coding region for MERS-CoV 3CL^{pro} results in autoprocessing of the His tag and overexpression of MERS-CoV 3CL^{pro} without any N-terminal extension in *E. coli* BL21-DE3 cells. MERS-CoV 3CL^{pro} was purified to high purity and an overall yield of 10% using four sequential chromatographic steps. A summary of the percent enzyme yield, total activity units, and the fold-purification after each chromatographic step is summarized in Table 1. Approximately 12 mg of highly pure MERS-CoV 3CL^{pro} can be obtained per liter of bacterial cell culture.

To verify the production of the enzyme with correct N and C termini, the molecular mass of purified MERS-CoV 3CL^{pro} was determined by MALDI to be 33.4 kDa, which is close to the theoretical molecular mass of 33.3 kDa for the authentic/mature MERS-CoV 3CL^{pro} monomer. Western blot analysis of purified MERS-CoV 3CL^{pro} using an anti-His_6 antibody also confirmed the absence of the N terminus His_6 tag associated with the expression plasmid (data not shown). These results demonstrate that the N-terminal His_6 tag is auto-catalytically removed by MERS-CoV 3CL^{pro} during its expression in *E. coli*, indicating MERS-CoV 3CL^{pro} is enzymatically active when expressed in *E. coli*.

MERS-CoV 3CL^{pro} Hydrolyzes a Fluorescent Peptide Substrate with Lower Efficiency than Other 3CL^{pro} Enzymes—A FRET-based peptide substrate was used to measure the enzymatic activity of MERS-CoV 3CL^{pro} as a function of substrate concentration over a substrate concentration range from 0 to 2.0 μM (Fig. 1*A*). We observed that MERS-CoV 3CL^{pro} cannot be saturated by the substrate over this concentration range, which is typical for other coronavirus 3CL^{pro} enzymes because the K_m values for peptide substrates approach 1 mM (51–54). Therefore, the slope of the kinetic response of MERS-CoV 3CL^{pro} to increasing substrate concentration was determined to derive an apparent (k_{cat}/K_m) value, which is a measure of

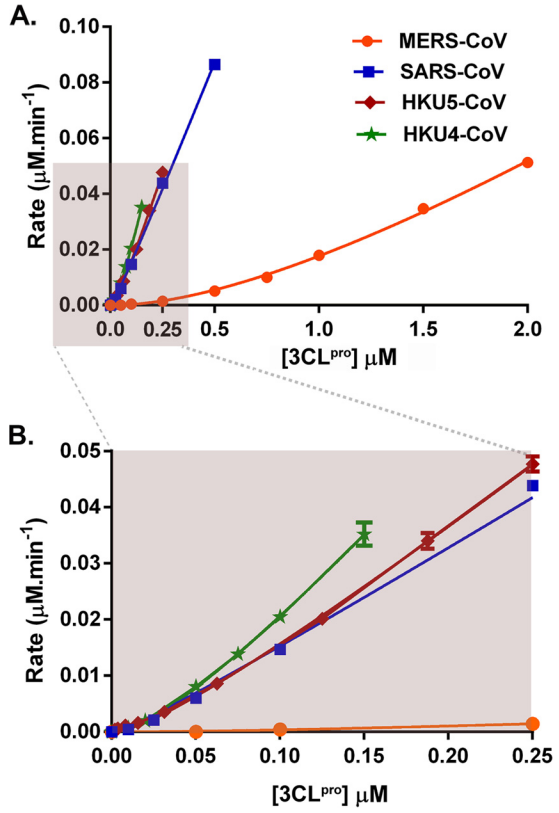


FIGURE 2. **Dependence of the enzymatic activity of MERS-CoV, HKU4-CoV, HKU5-CoV, and SARS-CoV 3CL^{pro} on the total enzyme concentration.** *A,* kinetic response of each CoV 3CL^{pro} to increasing enzyme concentration is plotted along with the resulting fit of the data to Equation 2. Resulting values for the apparent turnover number, k_{cat}, and the monomer-dimer equilibrium constant, K_d, are shown in Table 2. Final enzyme concentrations varied over the concentration ranges of 2 μM to 100 nM for MERS-CoV 3CL^{pro}, 500 to 10 nM for SARS-CoV 3CL^{pro}, 250 to 0.6 nM for HKU5-CoV 3CL^{pro}, and 200 to 10 nM for HKU4-CoV 3CL^{pro}. Final substrate concentration was fixed at 2 μM. Experiments were done in triplicate. *Error bars* represent the standard deviation for triplicate data. *Shaded box* represents the data that are plotted in *B. B,* enlarged view of the fitted data at low total enzyme concentrations, marked in *shaded box* in *A*, illustrating the nonlinear dependence of enzymatic activity on the total concentrations of 3CL^{pro} from SARS-CoV, HKU5-CoV, and HKU4-CoV.

enzymatic efficiency. We also determined and compared the apparent (k_{cat}/K_m) values for 3CL^{pro} enzymes from SARS-CoV, HKU5-CoV, and HKU4-CoV under similar experimental conditions (Fig. 1*B*). MERS-CoV 3CL^{pro} is able to hydrolyze the peptide substrate; however, the enzymatic efficiency of MERS-CoV 3CL^{pro} (k_{cat}/K_m = 3.1 ± 0.03 × 10^{-2} μM^{-1} min^{-1}) is noticeably lower than other 3CL^{pro} enzymes tested. Specifically, MERS-CoV 3CL^{pro} was 5-fold less efficient at processing the peptide substrate when compared with SARS-CoV 3CL^{pro}. Even among the β-CoVs from the same 2c genogroup (MERS, HKU5, and HKU4), MERS-CoV 3CL^{pro} was the least efficient enzyme.

MERS-CoV 3CL^{pro} Is a Weakly Associated Dimer—Because a dimer has consistently been shown to be the catalytically active form of all 3CL^{pro} enzymes studied to date, we tested the hypothesis that the lower enzymatic efficiency of MERS-CoV 3CL^{pro} is a result of the reduction in its ability to dimerize. Therefore, we determined the dependence of the enzymatic activity of MERS-CoV 3CL^{pro} on the total enzyme concentration and compared it with other 3CL^{pro} enzymes from HKU4, HKU5, and SARS coronaviruses (Fig. 2).

TABLE 2
Comparison of the apparent turnover number, k_{cat}, and the monomer-dimer dissociation constant, K_d, for 3CL^{pro} from different CoVs

3CL^{pro}	Nonlinear fitting of kinetic data[a]	
	k_{cat}[b]	K_d
	min^{-1}	μM
MERS-CoV	0.2 ± 0.02	7.8 ± 1.3
SARS-CoV	0.47 ± 0.03	0.06 ± 0.01
HKU5-CoV	0.53 ± 0.02	0.06 ± 0.01
HKU4-CoV	0.84 ± 0.07	0.1 ± 0.03

[a] Values were determined through nonlinear fitting of the kinetic data to Equation 2.
[b] k_{cat} represents the apparent turnover number.

It is immediately apparent from the data plotted in Fig. 2 that the response of MERS-CoV 3CL^{pro} enzymatic activity to an increasing enzyme concentration is nonlinear. The strong curvature suggests that a dimer is either the most active form or the only active form of MERS-CoV 3CL^{pro}. To determine the mechanism of dimerization, the data in Fig. 2 were first fit to Equation 1 (see "Experimental Procedures"), which describes a model where both the monomer and the dimer are active. A fit of the data to Equation 1 yielded a negative turnover value for the monomer ($k_{cat,\,M}$), suggesting the monomer is inactive and that the dimer is the only active form of the enzyme. Therefore, the data were fit to Equation 2 (see "Experimental Procedures"), which considers only the dimer as the active form of the enzyme. The kinetic data for all four 3CL^{pro} enzymes, MERS-CoV, HKU4-CoV, HKU5-CoV, and SARS-CoV, fit well to this model, and the resulting values for the monomer-dimer equilibrium dissociation constant, K_d, and apparent turnover number, k_{cat}, for each enzyme are provided in Table 2.

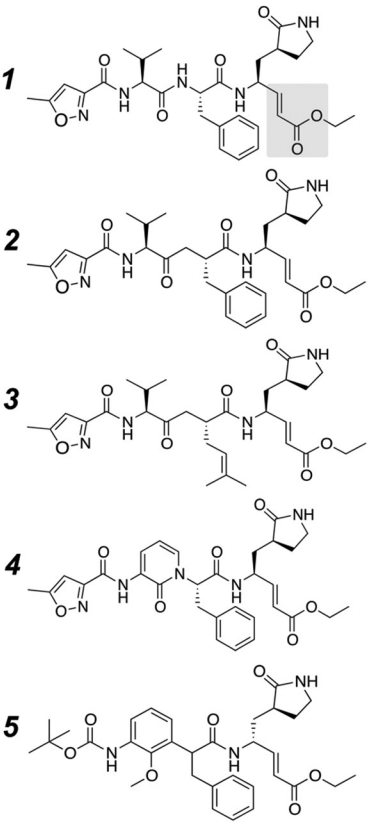
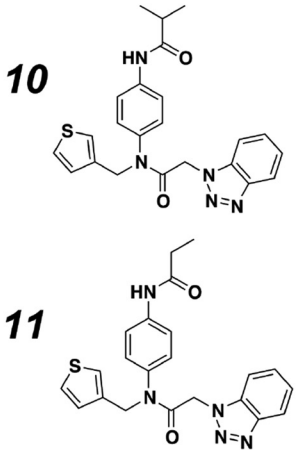
The lower k_{cat} value for MERS-CoV 3CL^{pro}, when compared with other coronavirus 3CL^{pro} enzymes, indicates a moderate reduction (2–4-fold) in its ability to turn over the substrate, which is consistent with the observed lower apparent (k_{cat}/K_m) value. In contrast, there is a substantial reduction in the ability of MERS-CoV 3CL^{pro} to dimerize compared with the other 3CL^{pro} enzymes. Based on the K_d values, the capacity of MERS-CoV 3CL^{pro} to dimerize is ~78–130-fold weaker than the other enzymes (Table 2). These results indicate that the MERS-CoV 3CL^{pro} dimer is much more weakly associated than the other coronavirus 3CL^{pro} enzymes studied, and these results raise questions as to the structural and mechanistic differences among the 3CL^{pro} enzymes that ultimately regulate protease activity during coronavirus replication.

MERS-CoV 3CL^{pro} Inhibition by Designed Peptidomimetic Compounds—In an effort to develop potent inhibitors of MERS-CoV 3CL^{pro}, we designed and synthesized nine peptidomimetic compounds containing a Michael acceptor group, *i.e.* an α,β-unsaturated carbonyl, capable of irreversibly reacting with the active site cysteine of MERS-CoV 3CL^{pro} (Table 3). These compounds were designed and synthesized based on our understanding and knowledge of the interactions of similar inhibitor molecules with SARS-CoV 3CL^{pro} (30, 31). At a concentration of 50 μM, compounds **6–9** displayed more than 50% inhibition of MERS-CoV 3CL^{pro} and were further evaluated for their ability to inactivate the enzyme in a time- and concentration-dependent manner (Fig. 3). Data from the kinetic progress curve for compound **6** (Fig. 3), as well as for

TABLE 3

Chemical structures and inhibitory activity of compounds 1 to 11 against MERS-CoV 3CL^{pro}

The Michael acceptor group for compound 1 is shaded to highlight this group for all the compounds. The stereochemistry at the benzyl stereocenter of compound 5 is a 1:1 mixture of enantiomers (racemic); therefore, the compound was tested as a mixture of diastereomers.

Peptidomimetic compounds with Michael-acceptor groups					Non-covalent peptidomimetics	
						
Cmpd	% Inhi *	k_{inact} [b]	$t_{1/2}^{\infty}$ [c]	K_I [d]	Cmpd	IC_{50} [e]
1 [a]	46	nd	nd	nd	***10***	>100
2 [a]	11	nd	nd	nd	***11***	>100
3 [a]	21	nd	nd	nd		
4 [a]	0	nd	nd	nd		
5	0	nd	nd	nd		
6	99	0.81 ± 0.08	0.86 ± 0.08	3.6 ± 0.8		
7	100	0.84 ± 0.05	0.83 ± 0.05	4.7 ± 0.6		
8	100	1.12 ± 0.20	0.62 ± 0.11	9.0 ± 2.3		
9	100	1.13 ± 0.20	0.61 ± 0.11	9.9 ± 2.6		

* % inhibition was measured as the % loss in enzymatic activity after 20 min of incubation of 500 nM MERS-CoV 3CL^{pro} with 50 μM of the compound.

[a] As compounds 1–5 showed <50% inhibition of MERS-CoV 3CL^{pro}, values of k_{inact}, $t_{1/2}^{\infty}$ and K_I were not determined (nd) for these compounds.

[b] k_{inact} is $\times 10^{-3}$ s^{-1}.

[c] $t_{1/2}^{\infty}$ is $\times 10^{3}$ s.

[d] K_I is in μM.

[e] IC_{50} values for compounds 10 and 11 were calculated from a dose- response curve determined after 10 min of incubation of 1 μM MERS-CoV 3CL^{pro} with varying concentrations of compounds. IC_{50} is in μM.

compounds **7**-**9** (data not shown), were fit to the appropriate equations (see under "Experimental Procedures") to obtain the kinetic parameters, k_{inact}, $t^{\infty}_{1/2}$, and K_I, and the resulting values are provided in Table 3.

We identified four compounds, **6**–**9**, as micromolar inhibitors of MERS-CoV 3CL^{pro} with K_I values less than 10 μM (Table 3). Analysis of structure-activity relationships of these compounds suggests that the S_2 subsite pocket of MERS-CoV 3CL^{pro} is small and can only accommodate a smaller P_2-isobutyl substituent (compounds **6**-**9**) but not bigger substituents such as P_2-benzyl or P_2-isobutylenyl (compounds **1**-**5**). It was also observed that replacing the P_4-ethoxy (compound **6**) with P_4-isopropoxy (compounds **7** and **8**) had no effect on the inhibitory activity of the compounds. Finally, these compounds provide an excellent chemical scaffold to study the molecular details of interactions of substrate-like compounds with the enzyme and to develop more potent inhibitors of MERS-CoV 3CL^{pro} for therapeutic intervention.

To evaluate broad spectrum specificity of these compounds, we also calculated % inhibition of SARS-CoV 3CL^{pro}, HKU5-CoV 3CL^{pro}, and HKU4-CoV 3CL^{pro} after 20 min of incubation in the presence of 50 μM compounds **6**–**9**. Except for compound **9**, which inhibited SARS-CoV 3CL^{pro} by 76%, we observed 100% inhibition of all other enzymes in the presence of compounds **6**-**9**. Furthermore, we performed progress curve analysis of HKU5-CoV 3CL^{pro} and HKU4-CoV 3CL^{pro} in the presence of varying concentrations of compounds **6**–**9**. The K_I values of compounds **6**–**9** for HKU5-CoV 3CL^{pro} are 0.49 ± 0.16, 0.60 ± 0.21, 1.30 ± 0.53, and 0.47 ± 0.06 μM, respectively. The K_I values of compounds **6**–**9** for HKU4-CoV 3CL^{pro} are 0.39 ± 0.14, 0.50 ± 0.17, 0.85 ± 0.33, and 0.64 ± 0.25 μM, respectively. These data suggest that peptidomimetic compounds **6**–**9** have the potential to be developed as coronavirus 3CL^{pro} inhibitors with broad spectrum specificity.

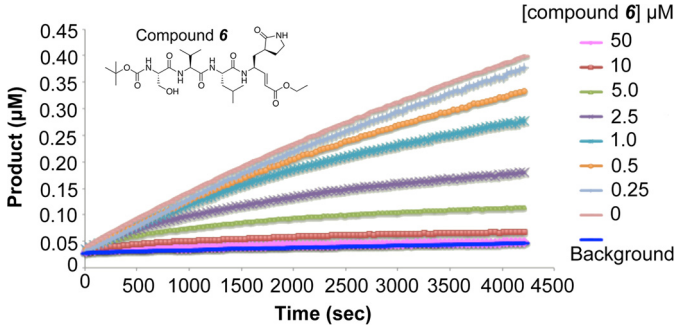


FIGURE 3. **Progress curves for the MERS-CoV 3CL^{pro}-catalyzed reaction in the presence of compound 6.** Time-dependent hydrolysis of 1 μM substrate catalyzed by 500 nM MERS-CoV 3CL^{pro} was measured over a time period of 70 min and at fixed variable concentrations of compound **6** ranging from 0 to 50 μM. Values for the inactivation kinetic parameters k_{inact}, $t^{\infty}_{1/2}$, and K_I were calculated by fitting the progress curve data to Equations 4 and 5. Chemical structure of compound **6** is shown in the *inset*.

Weak Association of the MERS-CoV 3CL^{pro} Dimer Is Supported by AUC Studies—To further explore the mechanism of MERS-CoV 3CL^{pro} dimerization, we performed analytical ultracentrifugation sedimentation velocity (AUC-SV) studies at varying concentrations of MERS-CoV 3CL^{pro} (Fig. 4*A*). Unlike enzyme kinetics, AUC allows determination of the monomer-dimer equilibrium constant (K_d) in the absence of substrate. MERS-CoV 3CL^{pro} displayed a continuous size distribution at different protein concentrations. Two distinct peaks corresponding to monomer (2.9 S) and dimer (3.9 S) species are observed, with the dimer peak becoming more pronounced at higher enzyme concentrations (Fig. 4*A*). We fit the AUC data to a monomer-dimer equilibrium model to determine the values for K_d and k_{off}, where K_d is the equilibrium dissociation constant for a monomer from the dimer, and k_{off} is the rate constant for dissociation of the monomer from the dimer. The resulting best fit value for K_d is 52 ± 5 μM and that for k_{off} is 10^{-4} s^{-1}. The K_d value of 52 μM for MERS 3CL^{pro} is dramatically different from SARS-CoV 3CL^{pro}, which has reported K_d values ranging from low nanomolar up to 10 μM depending on the enzyme construct used and the experimental conditions and methods utilized to determine the dissociation constant (37). The dimer affinity of MERS-CoV 3CL^{pro} is substantially weaker than that for SARS-CoV 3CL^{pro}, when comparing the same enzyme construct, *i.e.* the enzyme without any N- or C-terminal modifications. The AUC-SV calculated K_d value for MERS-CoV 3CL^{pro} is ~150,000 times higher than the value of 0.35 nM determined for SARS-CoV 3CL^{pro} (34).

The AUC results (Fig. 4*A*) show that the monomer peak at ~2.9S does not gradually shift peak position toward the dimer peak at ~3.9S with increasing concentrations of MERS-CoV 3CL^{pro}; rather, the two peaks change in area, which is indicative

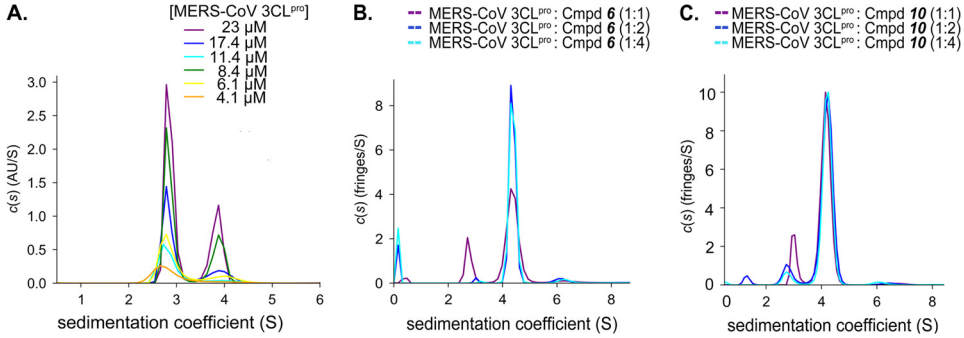


FIGURE 4. **AUC-SV analyses of ligand-induced dimerization of MERS-CoV 3CL^{pro}.** *A*, sedimentation coefficient distribution for varying concentrations of MERS-CoV 3CL^{pro} (4.1 to 23 μM) with sedimentation coefficient values of 2.9S and 3.9S for the monomer and the dimer, respectively. The best fit value for AUC-SV-calculated K_d is 52 ± 5 μM. *B*, sedimentation coefficient distribution of MERS-CoV 3CL^{pro} (25 μM) in the presence of different stoichiometric ratios of compound **6** (25, 50, and 100 μM). *C*, sedimentation coefficient distribution of MERS-CoV 3CL^{pro} (25 μM) in the presence of different stoichiometric ratios of compound **10** (25, 50, and 100 μM). A significant shift in the 2.9S peak (monomer) to a 4.1S peak (dimer) is detected upon addition of increasing concentrations of compounds **6** and **10.**

of very slow monomer-dimer exchange rate (k_{off} ~10^{-4} s^{-1}) and the formation of hydrodynamically stable monomer and dimer species (55). This k_{off} value is 1000 times slower than the k_{off} value (10^{-1} s^{-1}) reported for SARS-CoV 3CL^{pro} indicating that the SARS-CoV enzyme has a significantly more rapid monomer-dimer exchange rate (56). These observations support a model whereby the MERS-CoV 3CL^{pro} dimer is weakly associated, suggesting the enzyme exists mainly as a monomer in solution.

MERS-CoV 3CL^{pro} Undergoes Extensive Ligand-induced Dimerization—The weak association of MERS-CoV 3CL^{pro} monomers engenders the following questions. "Are higher levels of expression of 3CL^{pro} in MERS-CoV-infected cells necessary to allow formation of active dimer?" "Are other mechanisms such as substrate- or ligand-induced dimerizations involved in activating 3CL^{pro}?" To explore the latter question of ligand-induced dimerization of MERS-CoV 3CL^{pro}, we performed AUC experiments in the presence of compound **6**, which acts as a substrate mimetic and mechanism-based inhibitor, also known as a suicide substrate. Peptidomimetic compounds such as compound **6**, which contains a Michael acceptor group, interact and react with the active site cysteine of cysteine proteases to covalently modify them. We utilized compound **6** to form a covalent MERS-CoV 3CL^{pro} and inhibitor **6** complex that is stable over long periods of time, making it amenable to analysis by AUC-SV experiments. In contrast, incubation of a normal peptide substrate with the enzyme would lead to immediate hydrolysis of the substrate and dissociation of the products from the enzyme, confounding AUC experiments and subsequent data analysis.

MERS-CoV 3CL^{pro} was incubated with varying concentrations of compound **6** in stoichiometric ratios of 1:1, 1:2, and 1:4. The modified enzyme was then subjected to AUC studies to determine the influence of compound **6** on the monomer-dimer equilibrium (Fig. 4*B*). A significant shift in the area under 2.9S peak (monomer) to 4.1S peak (dimer) is detected upon addition of increasing concentrations of compound **6**. We obtained similar results when AUC studies were performed utilizing a complex of MERS-CoV 3CL^{pro} with a noncovalent peptidomimetic inhibitor (compound **10**, Figs. 4*C*). The transition of MERS-CoV 3CL^{pro} from monomer to dimer in the presence of compounds **6** and **10** suggests that the enzyme undergoes extensive dimerization upon substrate binding.

MERS-CoV 3CL^{pro} Is Activated by Ligand-induced Dimerization—The observed ligand-induced dimerization of MERS-CoV 3CL^{pro}, as demonstrated through AUC studies, prompted us to investigate whether or not the enzymatic activity of MERS-CoV 3CL^{pro} could be increased at low concentrations of a compound via ligand-induced dimerization. To do so, we chose to use a noncovalent peptidomimetic compound (compound **10**, Fig. 5*A*) that we previously identified as an inhibitor of SARS-CoV 3CL^{pro}. Because of the time-dependent, irreversible nature of the reaction between compound **6** and MERS-CoV 3CL^{pro}, use of compound **6** was not ideal for these kinetic studies as it would further complicate kinetic data analysis.

The kinetic response of MERS-CoV 3CL^{pro} to increasing concentrations of compound **10** was first measured at a single enzyme concentration of 1.0 μM (Fig. 5*A*). Interestingly, an increase in the activity of MERS-CoV 3CL^{pro}, as high as 195%, was observed in the presence of low inhibitor concentrations (0.1 to 20 μM). Inhibition of enzymatic activity was observed only at higher inhibitor concentrations (40 μM or greater). These results suggest that at low concentrations, compound **10** binds to a monomer and induces the formation of a dimer. The resulting dimer then has one free active site that is capable of processing the substrate. At higher concentrations of inhibitor, the substrate and inhibitor directly compete for the free active site.

The model of activation and inhibition suggested by the data at 1 μM enzyme would predict that at higher enzyme concentrations less activation by a compound would be observed at lower inhibitor concentrations, and the inhibition of activity would be detected at lower inhibitor concentrations because the equilibrium would be pushed toward dimer formation. In contrast, lower enzyme concentrations would result in higher activation by compounds, and inhibition by the compound would occur at significantly higher compound concentrations. Therefore, we further measured the activity of MERS-CoV 3CL^{pro} at two additional enzyme concentrations (0.5 and 2.0 μM) in the presence of varying concentrations of compound **10**. Remarkably, we observed that the activation effect was most pronounced at the lowest MERS-CoV 3CL^{pro} concentration tested (0.5 μM), and the effect decreased as the enzyme concentration was increased (1.0 and 2.0 μM) (Fig. 5*A*). Moreover, inhibition by compound **10** occurred at lower compound concentrations when higher concentrations of enzyme were used. These observations further support a model whereby enzyme activation can occur through ligand-induced dimerization.

The activation and inhibition of MERS-CoV 3CL^{pro} by compound **10** can be explained by a simple kinetic model depicted in Fig. 5*B*. The MERS-CoV 3CL^{pro} monomer exists in equilibrium with the dimer, and their relative concentrations depend on the total enzyme concentration. In the absence of substrate or compound, the K_d value is 52 μM, and the equilibrium is represented by the *gray spheres* (*blue box*) in Fig. 5*B*. The monomer is unable to hydrolyze the substrate and is therefore inactive. Binding of inhibitor (Fig. 5*B, green triangle*) to the monomer results in monomer to dimer switch leading to the formation of a dimer that contains inhibitor bound in one of the active sites. Once the dimer is formed, the substrate binds in the second active site and catalysis takes place. Under high inhibitor concentrations, however, the inhibitor molecule directly competes with substrate for the free dimer active site, and inhibition of the enzymatic activity is observed as a result.

We would also expect to observe induced dimerization and activation in the presence of the substrate. Indeed, the monomer-dimer kinetic studies performed in Fig. 2 were performed at a fixed concentration of substrate at 2 μM. In this experiment, the K_d value for the MERS-CoV 3CL^{pro} dimer was determined to be 7.8 μM, which is lower than the K_d value determined in the absence of substrate using AUC, thereby supporting substrate-induced dimerization. Given the high K_m value of 3CL^{pro} for the peptide substrate (51–54), even higher substrate concentrations would be required to observe substrate activation in a plot

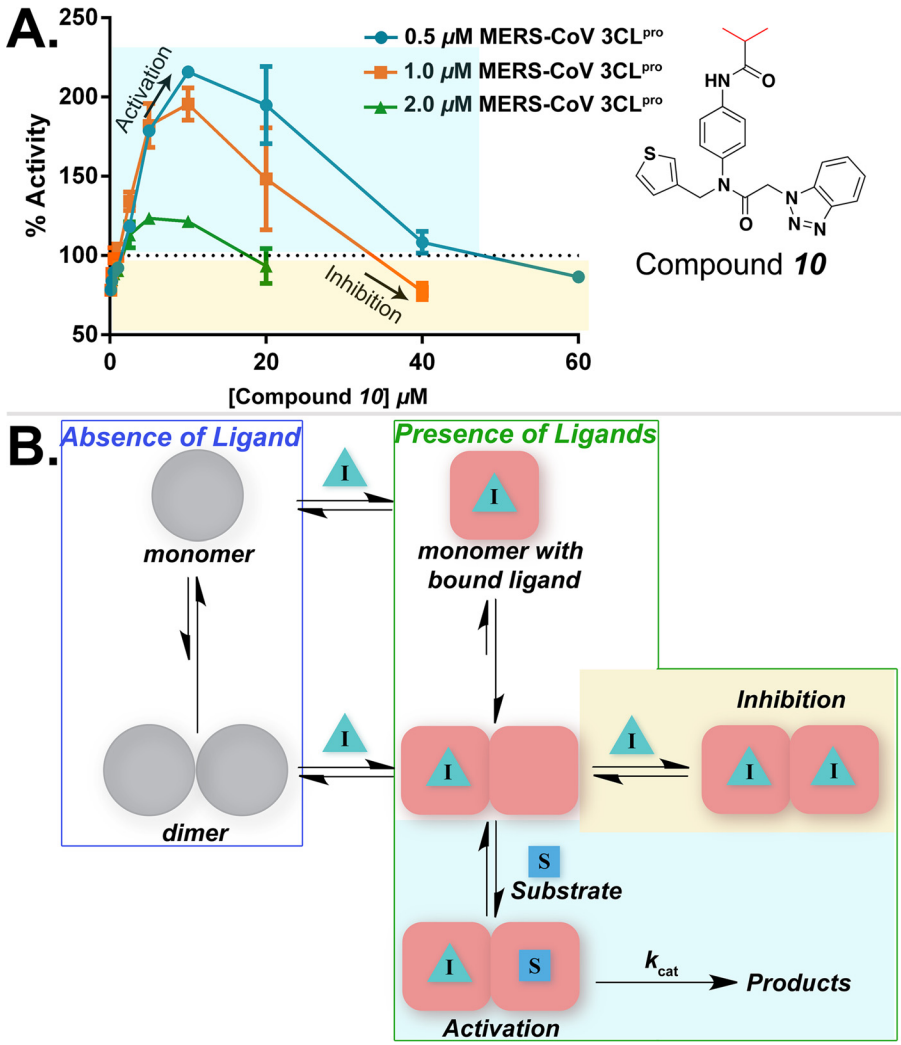


FIGURE 5. **Activation of MERS-CoV 3CL^{pro} via ligand-induced dimerization.** *A,* enzymatic activity of 0.5, 1.0, and 2.0 μM MERS-CoV 3CL^{pro} was measured in the absence and presence of varying concentrations of compound **10.** Substrate concentration was fixed at 2.0 μM. % activity, normalized to zero inhibitor enzymatic activity, was plotted as a function of increasing inhibitor concentrations. *Error bars* represent the standard deviation for triplicate data. Increase in enzymatic activity (highlighted in *cyan-shaded box*) is observed in the presence of low concentrations of compound **10**. Inhibition of enzymatic activity is observed at higher inhibitor concentrations (highlighted in *yellow-shaded box*). *B,* kinetic model describing the equilibrium between different species of MERS-CoV 3CL^{pro} that are formed in the absence (*blue box*) and presence (*green box*) of a ligand is shown. Based on the AUC-calculated K_d value of ~ 52 μM, MERS-CoV 3CL^{pro} primarily exists as a monomer in solution in the absence of a ligand. Upon ligand binding (inhibitor *I* in our case) to the monomer, the monomer-dimer equilibrium shifts toward dimer formation. Next, under lower inhibitor concentrations (*cyan-shaded box*), substrate (*S*) binds in the second active site and catalysis takes place. However, under higher inhibitor concentrations (*yellow-shaded box*), inhibitor directly competes with the substrate for the second active site, and inhibition of the enzymatic activity is observed.

of catalytic activity *versus* substrate concentration. However, we are limited to use our FRET-based substrate only at low concentrations due to a significant inner filter effect at higher concentrations of substrate. Therefore, a compound that both mimics substrate and has higher binding affinity can act as a useful surrogate for the substrate, allowing the observation of ligand-induced dimerization and activation even at low substrate concentrations.

X-ray Structure of MERS-CoV 3CL^{pro} in Complex with Compound **6**—To gain atomic level detail and molecular insight into the mechanism for substrate-induced dimerization of MERS-CoV 3CL^{pro}, we attempted to crystallize and determine the x-ray structures of the unliganded MERS-CoV 3CL^{pro} monomer and the MERS-CoV 3CL^{pro} covalently modified with compound **6**. Unfortunately, we were unable to crystallize the unliganded MERS-CoV 3CL^{pro} monomer after multiple attempts, but we were able to crystallize and determine the x-ray structure of MERS-CoV 3CL^{pro} in complex with compound **6** to a resolution of 1.6 Å. The statistics for x-ray data collection, processing, and refinement are summarized in Table 4. The MERS-CoV 3CL^{pro} and **6** complex crystallized as a biologically relevant, symmetrical dimer in space group *C*2 with one monomer in the asymmetric unit. Electron density for the entire protein was clearly visible and strong electron density ($F_o - F_c > 4\sigma$) was present for compound **6** within the active site (Fig. 6*A*).

MERS-CoV 3CL^{pro} Has a Smaller S_2 Pocket than SARS-CoV 3CL^{pro}—The active site of MERS-CoV 3CL^{pro} bound with compound **6** is shown in Fig. 6, *A* and *B*. Compound **6** is covalently bound to the active site cysteine (Cys-148) via a 1.8 Å bond between the γ-sulfur and the electrophilic β-carbon of the Michael acceptor. The P'_1-ethyl ester carbonyl, which mimics the carbonyl of the scissile bond in a substrate, forms a hydro-

TABLE 4
X-ray data collection and refinement statistics

	MERS-CoV 3CL^{pro}·6	MERS-CoV 3CL^{pro}·11
Beamline	LRL-CAT sector 31 ID-D	LS-CAT sector 21 ID-G
Data collection		
Wavelength (Å)	0.9793	0.9786
Resolution range (Å)	19.35–1.62 (1.68–1.62)[a]	50.00–2.10 (2.14–2.10)[a]
Protein monomers in asymmetric unit	1	4
Space group	*C*2	$P2_1$
Unit cell dimensions		
a, *b*, *c* (Å)	106.49, 57.31, 48.88	63.44, 114.93, 92.34
α, β, γ (°)	90, 112.78, 90	90, 90.89, 90
Total no. of reflections	63,855	816,216
No. of unique reflections	32,851	76,865
Multiplicity	1.9 (1.9)[a]	2.2 (2.2)[a]
Completeness (%)	95.0 (93.8)[a]	96.8 (93.8)[a]
Mean $I/\sigma I$	5.2 (1.3)[a]	11.17 (1.83)[a]
R_{merge} (%))[b]	8.3 (67.2)[a]	8.8 (58.6)[a]
Refinement		
Resolution range (Å)	19.35–1.62	42.59–2.10
No. of reflections in working set	30824	76623
No. of reflections in test set	2026	2019
R_{work} (%)[c]	17.8	15.91
R_{free} (%)[c]	21.7	21.51
No. of non-hydrogen atoms		
Protein/water	2380/208	9383/995
r.m.s.d.,[d] bond lengths (Å)	0.007	0.013
r.m.s.d., bond angles (°)	1.09	1.35
Ramachandran favored (%)	99	98
Ramachandran outliers (%)	0	0
Molprobity clash score	3.3	1.94
Average *B*-factor (Å^2)	20.4	33.1
Protein	19.8	32.5
Ligands	16.6	41.1
Solvent	27.7	37.9

[a] Values in parentheses are for highest resolution shell.
[b] $R_{merge} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|/\Sigma_h\Sigma_i I_i(h)$, where $I_i(h)$ is the *i*th measurement and $\langle I(h)\rangle$ is the weighted mean of all measurements of $I(h)$.
[c] R_{work} and $R_{free} = h(|F(h)_o| - |F(h)_c|)/h\,|F(h)_o|$ for reflections in the working and test sets, respectively.
[d] r.m.s.d. is root mean square deviation.

gen bond with the backbone NH of Gly-146 that forms part of the oxyanion hole (Fig. 6*B*). Within the S_1 subsite, the P_1-lactam carbonyl, which is a surrogate for the amide of P_1-glutamine of substrates, participates in a hydrogen bonding interaction with the imidazole ring of His-166, and the P_1-lactam NH forms a hydrogen bond with the carboxylate oxygen of Glu-169. The P_2-backbone amide NH forms a hydrogen bond with the side chain carbonyl of Gln-192 (Fig. 6*B*). The P_2-leucine side chain atoms of the inhibitor make hydrophobic contacts with the side chains of Met-168 and Leu-49 that line the S_2 subsite pocket. Moreover, compared with the equivalent residue Thr-25 in SARS-CoV 3CL^{pro}, Met-25 in the S_2 pocket of MERS-CoV 3CL^{pro} is expected to reduce the size of the hydrophobic pocket, which is supported by our observed SAR described above.

The smaller size of the S_2 pocket in MERS-CoV 3CL^{pro} is also consistent with the preference for a smaller leucine residue at the P_2 position of cleavage sites instead of a bulkier phenylalanine or methionine residue. Indeed, analysis of the preference for leucine or phenylalanine at the P_2 position for the 11 3CL^{pro} cleavage sites within the polyprotein of MERS-CoV shows that none of the 11 cleavage sites contain a phenylalanine residue at this position (Fig. 6*C*). Leucine is the predominantly favored residue at this position followed by methionine. Analysis of the cleavage sites from SARS-CoV, HKU4-CoV, and HKU5-CoV shows that none of the 11 cleavage sites from group 2c members (MERS-CoV, HKU4-CoV, and HKU5-CoV) contain a phenylalanine residue at the P_2 position; however, the SARS-CoV nsp5 ↓ nsp6 cleavage site contains a phenylalanine residue at this position.

Other interactions are also observed to play a significant role in stabilizing the MERS-CoV 3CL^{pro}-compound **6** complex. The P_3-carbonyl and P_3-NH participate in hydrogen bonding interactions with the backbone NH and carbonyl of Glu-169. The P_4-serine side chain is within hydrogen bonding distance of the side chain carboxamide of Gln-195 and the backbone carbonyl of Lys-191.

X-ray Structure of MERS-CoV 3CL^{pro} in Complex with a Noncovalent Inhibitor—We were also able to obtain diffraction quality crystals of MERS-CoV 3CL^{pro} in complex with compound **11**, which has an almost identical chemical structure as that of compound **10** (Fig. 6*D*). We previously showed that compounds similar to **10** and **11** act as potent noncovalent inhibitors of 3CL^{pro} from SARS-CoV (33). The x-ray structure of compound **11** bound to MERS-CoV 3CL^{pro} was determined to a resolution of 2.1 Å and the x-ray data collection, processing, and refinement statistics are summarized in Table 4. The MERS-CoV 3CL^{pro} and **11** complex crystallized in space group $P2_1$ with two biologically relevant dimers in the asymmetric unit. The overall root mean square deviation between the C-α atoms of the four chains was less than 1 Å, with the highest C-α root mean square deviation of 0.719 Å between chains C and D. Strong electron density ($F_o - F_c > 4\sigma$) was present for compound **11** within all the four active sites of the two dimers (Fig. 6*D*).

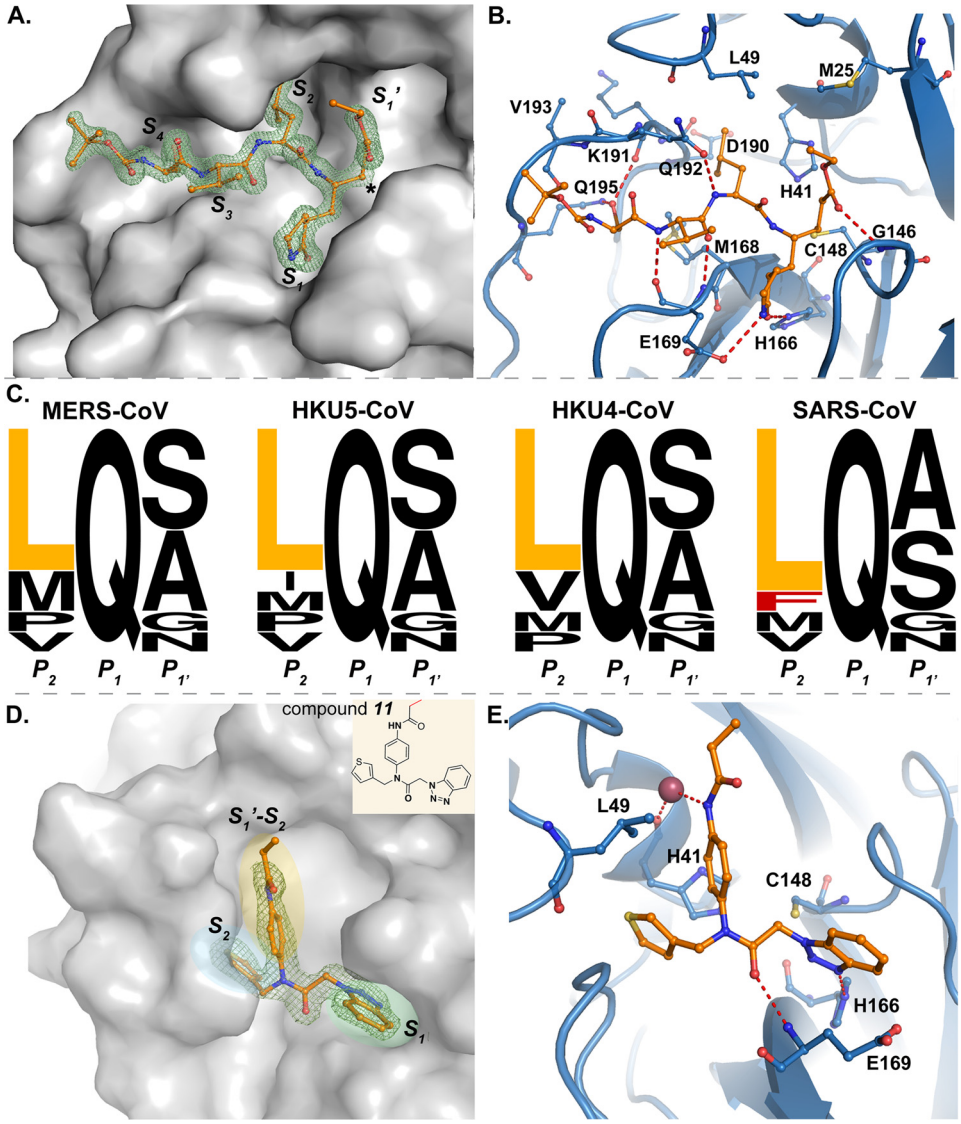


FIGURE 6. **X-ray crystal structure of MERS-CoV 3CL^{pro} in complex with inhibitors.** *A,* solvent-accessible surface (*gray-shaded surface*) of MERS-CoV 3CL^{pro} and compound **6** complex. Compound **6** is displayed in *ball and stick* model with atoms colored as follows: carbons (*orange*), nitrogens (*blue*), and oxygens (*red*). Electron density associated with compound **6** is shown as an $F_o - F_c$ electron density difference map contoured to 3σ (*green mesh*). Substrate binding pockets S_4-S'_1 are labeled, where *asterisk* indicates the electrophilic carbon of compound **6** that forms a C–S covalent bond with the active site cysteine Cys-148. *B,* MERS-CoV 3CL^{pro} and compound **6** complex with the MERS-CoV 3CL^{pro} backbone represented as a *ribbon* model and relevant amino acids that interact with compound **6** represented as *ball and sticks*. MERS-CoV 3CL^{pro} carbon atoms are colored *blue*, and compound **6** carbon atoms are colored *orange*. Nitrogen atoms are colored *blue*, and oxygen atoms are colored *red*. Catalytic residues Cys-148 and His-41 are also shown. Hydrogen bonds are depicted as *red dashed lines*. *C,* sequence logos showing amino acid conservation for the 11 polyprotein cleavage sites of different 3CL^{pro} enzymes (MERS-CoV, HKU5-CoV, HKU4-CoV, and SARS-CoV), generated using the WebLogo server (63). Residues P_2-P'_1 are shown. *Height* of each letter corresponds to the amino acid conservation at that position. *D,* solvent-accessible surface (*gray-shaded surface*) of MERS-CoV 3CL^{pro} and compound **11** complex. Compound **11** is displayed in *ball and stick* model. Electron density associated with compound **11** is shown as a $2F_o - F_c$ electron density difference map contoured to 1.5σ (*green mesh*). Functional groups of compound **11** with their corresponding binding pockets are highlighted in *yellow*, *green*, and *blue ellipses*. Chemical structure of compound **11** is shown in the *inset*. *E,* interactions between MERS-CoV 3CL^{pro} and compound **11** are illustrated. Catalytic residues Cys-148 and His-41 are also shown. Hydrogen bonds are depicted as *red dashed lines*.

The binding orientation for compound **11** in the active site of MERS-CoV 3CL^{pro} is similar to the binding orientation of related compounds in the active site of SARS-CoV 3CL^{pro} (PDB code 4MDS). The benzotriazole group binds in the S_1 subsite; phenyl propionamidyl occupies the S'_1-S_2 subsite, and the thiophene group binds in the S_2 subsite. Compound **11** also forms two direct and one water-mediated hydrogen bond interactions with amino acids in the MERS-CoV 3CL^{pro} active site (Fig. 6*E*). The *N*3 of the benzotriazole ring forms a hydrogen bond with the side chain ϵ-nitrogen of conserved His-166, and the central acetamide oxygen forms a hydrogen bond with the backbone NH of conserved Glu-169. The NH of the phenyl propionamidyl group interacts with backbone carbonyl oxygen of the catalytic His-41 residue through a water-mediated hydrogen bond, and the imidazole ring of His-41 engages with the phenyl ring of phenyl propionamidyl group through T-shaped π stacking. The phenyl ring also form hydrophobic contacts with Leu-49.

Interactions at the 3CL^{pro} Dimer Interface—Analysis of the MERS-CoV 3CL^{pro} and **6** and MERS-CoV 3CL^{pro} and **11** crystal structures reveals key differences between the dimer interface of MERS-CoV and SARS-CoV 3CL^{pro} (PDB code 2ALV)

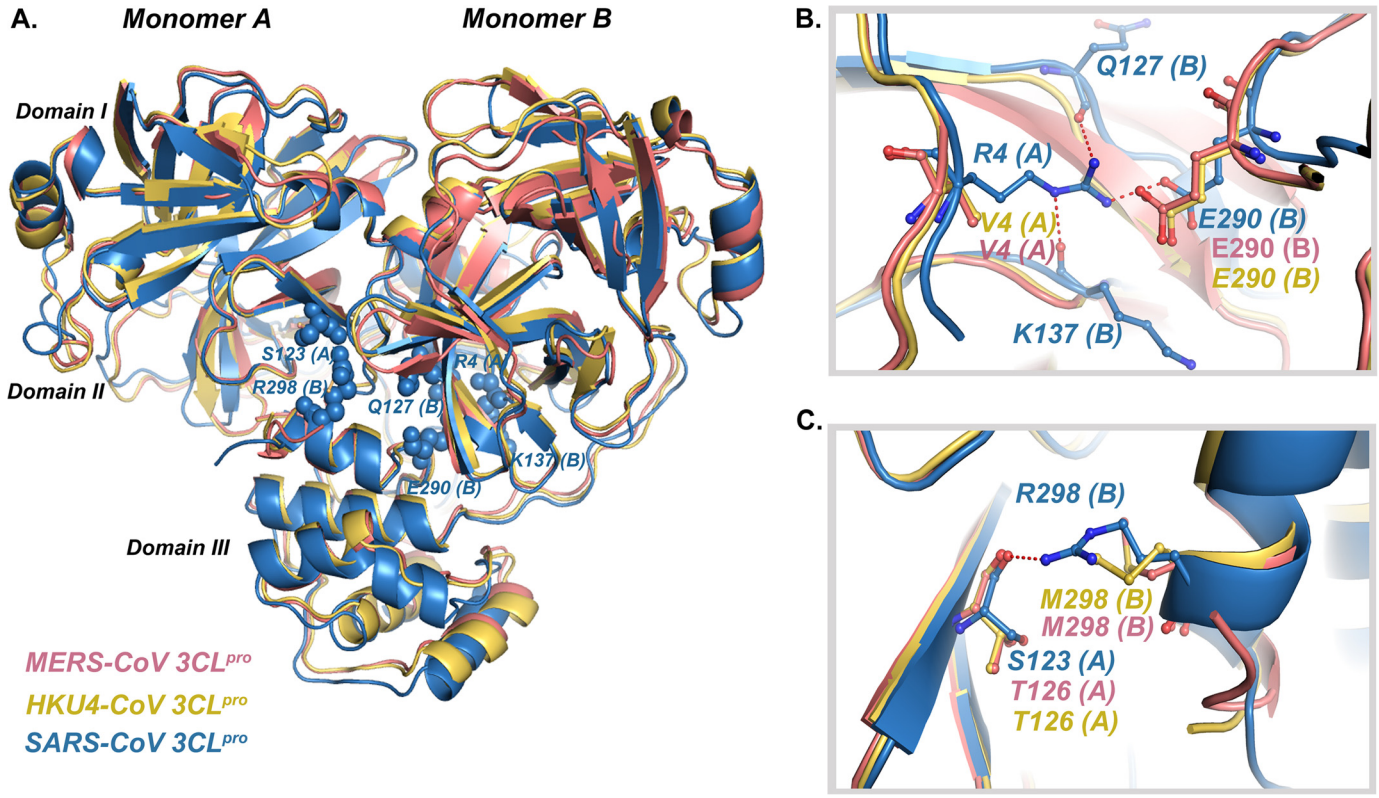


FIGURE 7. **Comparison of x-ray crystal structures of 3CL^{pro} dimers from MERS-CoV, HKU4-CoV, and SARS-CoV.** *A,* superposition of dimers of MERS-CoV 3CL^{pro} (*pink color*), HKU4-CoV 3CL^{pro} (*yellow color*, PDB code 2YNB), and SARS-CoV 3CL^{pro} (*blue color*, PDB code 2ALV). For SARS-CoV 3CL^{pro}, residues Arg-4 and Ser-123 from monomer A, and residues Gln-127, Lys-137, Glu-290, and Met-298 from monomer B are represented as *spheres*. *B,* for SARS-CoV 3CL^{pro}, interactions between the side chain of Arg-4 from monomer A and Gln-127, Glu-290, and Lys-137 residues from monomer B are shown. The corresponding residues in MERS-CoV 3CL^{pro} and HKU4-CoV 3CL^{pro} are Val-4 in monomer A and Glu-290 in monomer B, which do not interact at the dimer interface. *C,* for SARS-CoV 3CL^{pro}, Ser-123 from monomer A engages in hydrogen bonding with Arg-298 from monomer B across the dimer interface. The corresponding residue in monomer B of MERS-CoV 3CL^{pro} and HKU4-CoV 3CL^{pro} is Met-298, which does not participate in any interaction with Thr-126 from monomer A across the dimer interface.

(Fig. 7) (30). Two arginine residues, Arg-4 and Arg-298 (Fig. 7, *A–C*), form some of the key interactions at the dimer interface of SARS-CoV 3CL^{pro}, and mutation of either of these amino acids results in a drastic loss of dimerization in SARS-CoV 3CL^{pro} (36, 38). Interestingly, these two arginine residues (Arg-4 and Arg-298) are substituted in MERS-CoV 3CL^{pro} by two hydrophobic residues (Val-4 and Met-298) that are unable to participate in the formation of hydrogen bonds or salt bridges. Therefore, we initially thought that the loss of these key interactions might simply explain the >100,000-fold weaker dimerization observed for MERS-CoV 3CL^{pro} compared with SARS-CoV 3CL^{pro}. Surprisingly, however, structural analysis of the dimer interface from the available x-ray structure of HKU4-CoV 3CL^{pro} (PDB code 2YNB; Fig. 7, *B* and *C*), and primary sequence alignment of 3CL^{pro} from MERS-CoV, HKU5-CoV, HKU4-CoV and SARS-CoV (Fig. 8) revealed that Val-4 and Met-298 are conserved between all the β-CoV 2c members studied here. Substantial differences between the ability of MERS-CoV 3CL^{pro} and HKU4/HKU5-CoV 3CL^{pro} to dimerize, despite their high sequence identity, led us to the hypothesis that nonconserved residues between MERS-CoV and other β-CoV 2c members that are remote from the dimer interface may play a significant role in dimer formation.

Analysis of Nonconserved Residues of MERS-CoV 3CL^{pro}—Analysis of our current crystal structures does not reveal a clear mechanism for the monomer to dimer switch of MERS-CoV 3CL^{pro} upon ligand binding. Therefore, we attempted to identify the nonconserved residues in MERS-CoV 3CL^{pro} that might affect enzymatic activity due to their proximity to key residues involved in substrate binding and/or dimer formation.

Based on a sequence alignment, MERS-CoV 3CL^{pro} contains ~24 nonconserved amino acids (*pink arrows* in Fig. 8). Upon analyzing the position of these amino acids in the crystal structure, we observed that a remarkable number of these amino acids are present in the loop regions. Fig. 9*A* illustrates the nonconserved residues present in the loop regions as *gray* (monomer A) and *pink* (monomer B) spheres. Interestingly, we also observed that there are hot spots in the protein structure where most of these amino acids are clustered. These hot spots include the N-terminal region, the active site region, the interdomain loop (loop between the catalytic fold and domain III), and the domain III. In MERS-CoV 3CL^{pro}, nonconserved amino acid His-8, which forms van der Waals contacts with Lys-155 of the same monomer and Thr-128 of the other monomer, is present at the end of the N-terminal finger (Fig. 9, *B* and *C*), whereas amino acids Asp-12 and Ala-15 are part of the N-terminal helix (Fig. 9*B*). Additionally, amino acids Thr-128, Lys-155, and Ser-158 are present within 6 Å of the N-terminal region (Fig. 9*B*). Substitution to these amino acids in MERS-CoV 3CL^{pro} might have changed the protein dynamics in a way

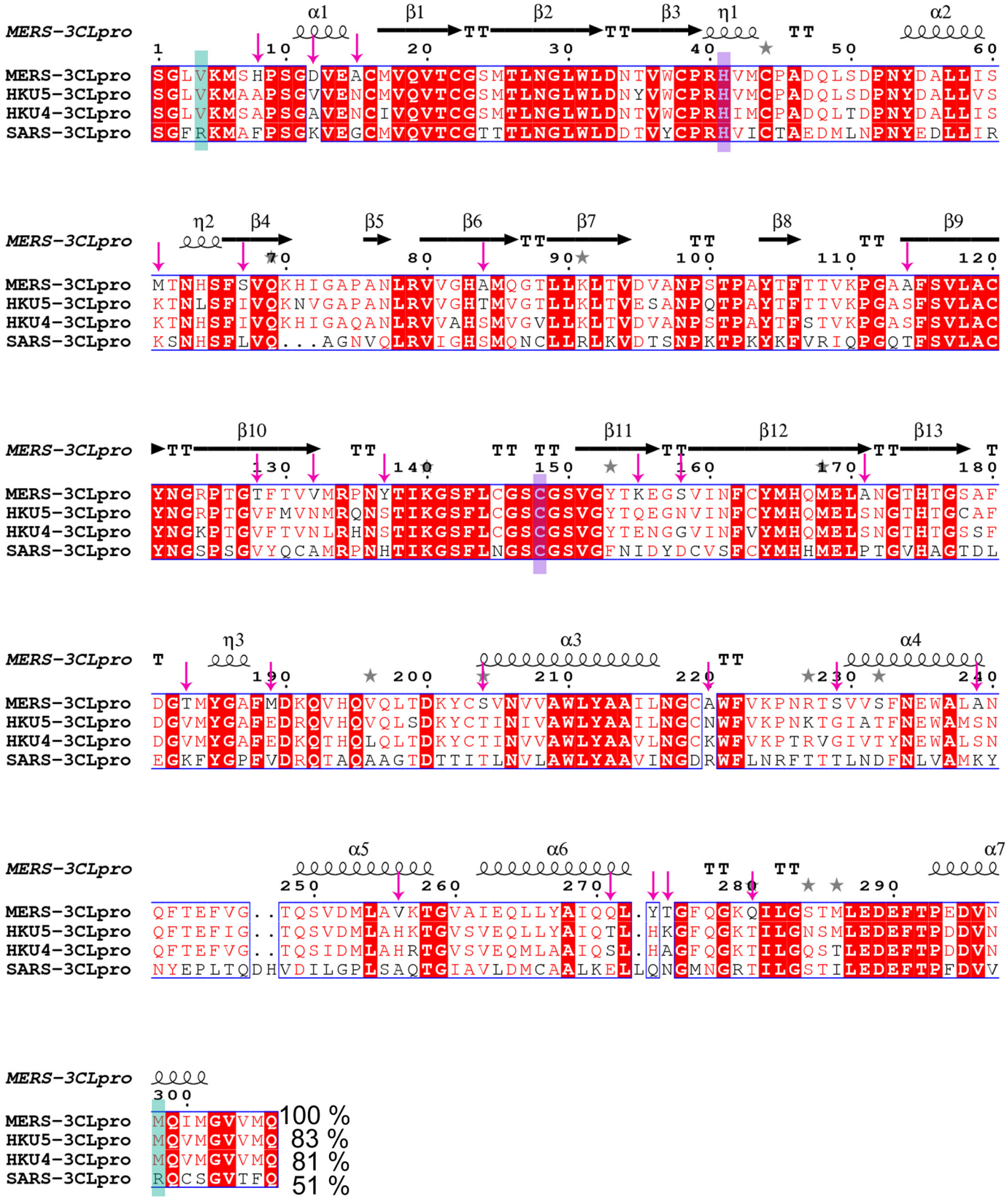


FIGURE 8. **Sequence alignment of 3CL^{pro} enzymes from MERS-CoV, HKU5-CoV, HKU4-CoV, and SARS-CoV.** Programs MultAlin (64) and ESPript (65) were used for the sequence alignment and visualization. Secondary structural elements of MERS-CoV 3CL^{pro} are represented as *spirals* for α-helix, *arrows* for β-strands, η for 3_{10} helix, and *T* for β-turns. Residues Val-4 and Met-298 in MERS-CoV, HKU5-CoV, HKU4-CoV 3CL^{pro}, and Arg-4 and Arg-298 in SARS-CoV are shown in a *green box*; catalytic residues His-41 and Cys-148 are highlighted in a *purple box*. The nonconserved residues of MERS-CoV 3CL^{pro} are marked with *pink arrows*. % identity with MERS-CoV 3CL^{pro} is shown.

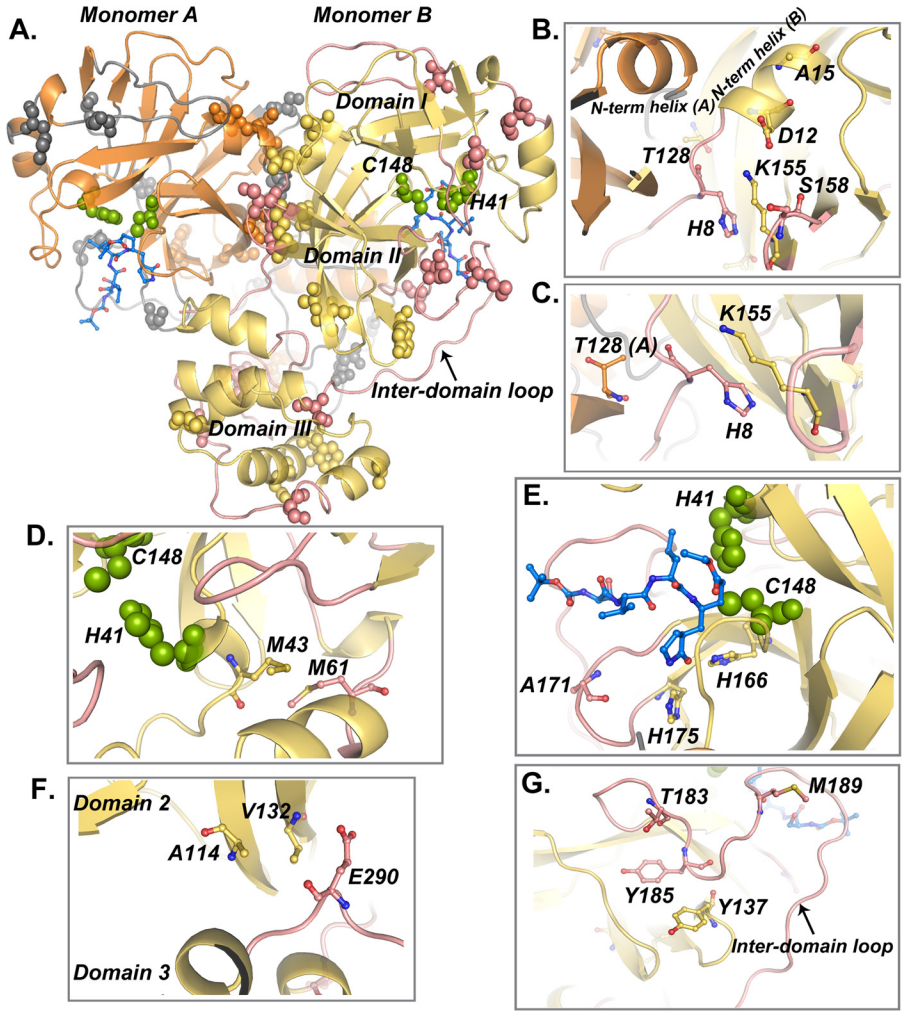


FIGURE 9. **Analysis of the nonconserved amino acids of MERS-CoV 3CL^{pro}.** *A,* representation of MERS-CoV 3CL^{pro} dimer with monomers A and B colored in *orange* and *yellow*, respectively. Nonconserved residues that are present in the loop regions are shown as *spheres* in *gray* and *pink* for monomers A and B, respectively. Other nonconserved residues are represented as *spheres* with the corresponding chain color. Domains I–III and the inter-domain loop are labeled. Catalytic residues His-41 and Cys-148 are shown as *green spheres*. Inhibitor molecule is shown in both active sites in *blue sticks*. *B–G,* residues of monomer B are shown (*yellow* and *pink*), unless otherwise labeled. *B,* clustering of some of the nonconserved amino acids, His-8, Asp-12, Ala-15, Thr-128, Lys-155, and Ser-158, near the N-terminal region is shown. N-terminal helices for both monomers are labeled. *C,* His-8 from the N-terminal region forms van der Waals contacts with Lys-155 of the same monomer and Thr-128 of the other monomer in the dimer. *D,* nonconserved residue Met-61 forms hydrophobic contacts with the Met-43 residue, which is in close proximity to the catalytic residue His-41. *E,* loop containing the nonconserved residue Ala-171 forms the S_1 pocket along with residues His-166 and His-175. *F,* Val-132 forms hydrophobic contacts with a residue within the same domain (Ala-114), as well as Glu-290 from domain III. *G,* nonconserved residue Tyr-137 makes hydrophobic contacts with Tyr-185; Tyr-185 along with two other nonconserved residues Thr-183 and Met-189 are present on the inter-domain loop.

that only ligand binding populates the monomer conformation, which is more amenable to dimer formation.

We also observe that some of the nonconserved residues in MERS-CoV 3CL^{pro} are located in proximity to the substrate-binding site and might contribute toward ligand-induced dynamic changes favorable for dimer formation. For example, nonconserved amino acid Met-61 forms hydrophobic interactions with Met-43, which in turn is in close proximity to the catalytic residue His-41 (Fig. 9*D*). Residue Ala-171 is present on a loop, and this loop, along with conserved residues His-166 and His-175, forms the S_1 subsite for binding the P_1 amino acid of the substrate (Fig. 9*E*). In addition to its influence on substrate binding, Ala-171 may also contribute toward dimer formation upon substrate binding due to its close proximity with Glu-169. This glutamate residue in SARS-CoV 3CL^{pro} (Glu-166) has been established as a key residue linking the substrate-binding site to the dimer interface (56). Val-132 forms hydrophobic interaction with other nonconserved residue Ala-114 within domain II (Fig. 9*F*). Additionally, Val-132 is present within van der Waals contact distance of Glul-290 from extra-helical domain III (Fig. 9*F*). It is noteworthy that Glu-290 forms a salt bridge with Arg-4 across the dimer interface in SARS-CoV 3CL^{pro}. However, this interaction is not formed in MERS-CoV 3CL^{pro} due to the substitution of Arg-4 with Val-4. Tyr-137 forms hydrophobic contacts with the conserved residue Tyr-185 (Fig. 9*G*).

Besides amino acid Val-132 that connects domains II and III, residue Tyr-185, along with two other nonconserved residues, Thr-183 and Met-189, is present on the inter-domain loop that connects the catalytic fold (domains I and II) with the extra-helical domain III (Fig. 9*G*). Flexibility within these residues might affect the orientation of domain III required for dimer formation.

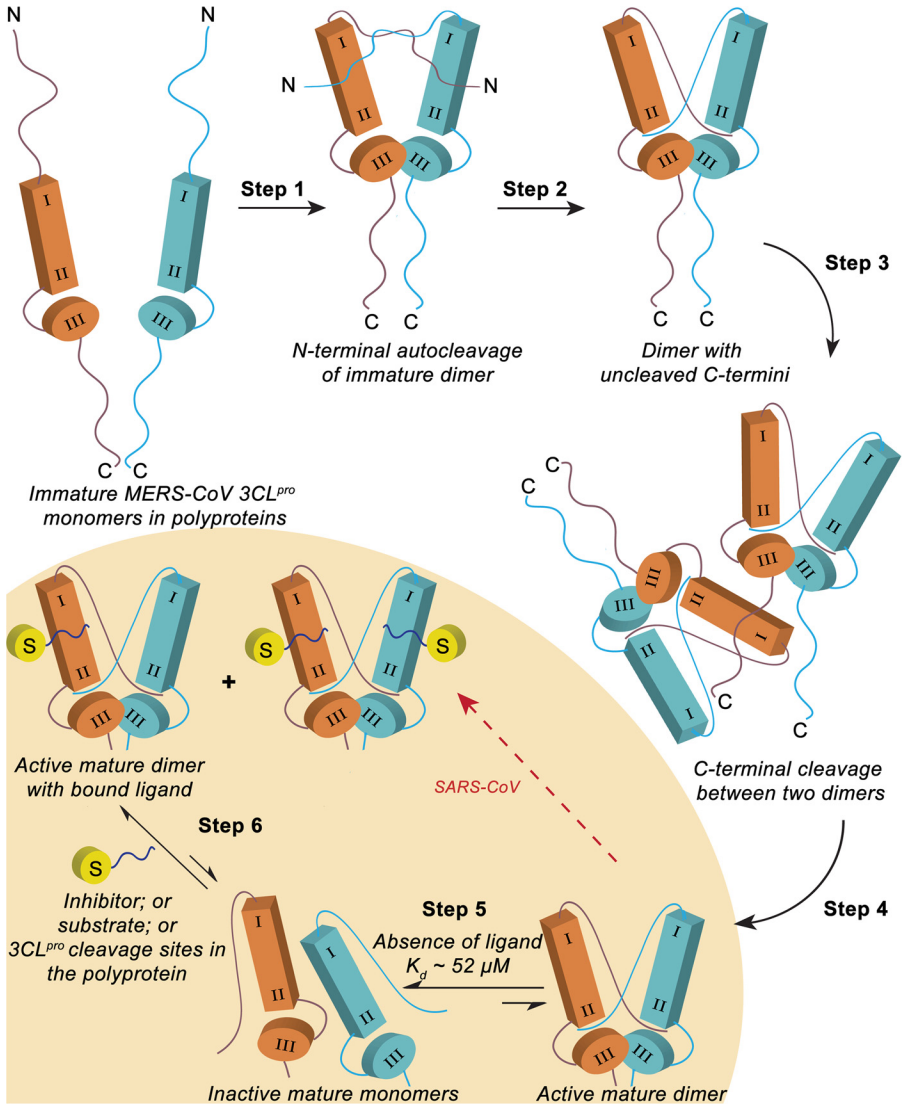


FIGURE 10. **Proposed model for polyprotein processing in MERS-CoV regulated by ligand-induced dimerization of MERS-CoV 3CL^{pro}.** MERS-CoV 3CL^{pro} domains I and II are together represented as the *rectangular box*, and domain III is represented as a *cylinder*. The N and C termini are labeled, and the *yellow cylinder* labeled *S* represents a ligand that can be a peptide inhibitor, peptide substrate, or 3CL^{pro} cleavage sites in the polyprotein. Various steps required for the auto-release of 3CL^{pro} from the polyprotein and subsequent processing of the polyprotein cleavage sites are described in the text. Suggested by our AUC and kinetic studies, the *shaded* region (*steps 5* and *6*) highlights the additional steps MERS-CoV 3CL^{pro} would undertake during polyprotein processing and have been described in the kinetic model depicted in Fig. *5B*.

Discussion

Model for Regulation of the Enzymatic Activity of MERS-CoV 3CL^{pro} during Polyprotein Processing—Enzymatic activity of coronavirus 3CL^{pro} is required for the processing of viral polyproteins at 11 distinct cleavage sites, allowing the release of nonstructural proteins that subsequently form a replication complex for virus genome replication. Because of its indispensable role in the virus life cycle, regulation of the enzymatic activity of 3CL^{pro} is instrumental for efficient replication of coronaviruses. Based on our experimental results, we propose a model to explain the mechanism for regulating the enzymatic activity of MERS-CoV 3CL^{pro} in the context of polyprotein processing during virus infection (Fig. 10).

A number of *in vitro* studies performed on SARS-CoV 3CL^{pro} have established the mechanism for 3CL^{pro} auto-release from the polyprotein (34, 39, 40). Based upon these studies and our data on MERS-CoV 3CL^{pro}, we propose the polyprotein processing model in Fig. 10. The steps proposed for auto-release of MERS-CoV 3CL^{pro} from the polyprotein (*steps 1–4*, Fig. 10) have been adapted from Chen *et al.* (39), where it is suggested that the N-terminal auto-processing does not require the formation of a mature 3CL^{pro} dimer for SARS-CoV. Based on the differences between the properties of SARS-CoV 3CL^{pro} and MERS-CoV 3CL^{pro}, as highlighted in our studies, we added two additional steps (*steps 5* and *6*, Fig. 10) that MERS-CoV 3CL^{pro} may need to utilize for efficient polyprotein processing. In Fig. 10, *step 1*, two immature MERS-CoV 3CL^{pro} monomers in the polyprotein approach each other and form an immature dimer via interactions between domain III, which allows each of the monomers to insert their N termini into the active site of the other monomer. In *step 2*, the N termini are cleaved, and the dimer with uncleaved C termini adopts a conformation similar

to the mature dimer. Our observation of auto-cleavage of the N-terminal His_6 tag from MERS-CoV 3CL^{pro} during expression in bacterial cells supports *steps 1* and *2*, where formation of an immature dimer capable of auto-processing the N terminus occurs. In *step 3*, two dimers with uncleaved C termini approach each other, followed by insertion of the C terminus from one dimer into one of the active sites of the other dimer. In *step 4*, the C termini are cleaved and mature dimer is released from the polyprotein.

For SARS-CoV, the 3CL^{pro} dimer formed in *step 4* continues to process cleavage sites in the polyprotein, effectively skipping *steps 5* and *6* (*red arrow* in Fig. 10) because the dimer is tightly associated. However, the high K_d value of MERS-CoV 3CL^{pro} dimer suggests that the active and mature dimer may dissociate into inactive, mature monomers in the absence of any ligand (*step 5*). In order for polyprotein processing to proceed, another step (*step 6*) must occur. In *step 6*, a substrate S, *e.g.* one of the 11 polyprotein cleavage sites, would induce dimer formation and hence activate catalysis and cleavage at the substrate recognition sites. Our AUC results and the kinetic activation studies performed in the absence and presence of inhibitors support *steps 5* and *6* where the inactive but mature monomers require binding of a ligand to undergo ligand-induced dimerization and formation of an active, mature dimer that can then process the polyprotein cleavage sites.

Nonconserved Amino Acids of MERS-CoV 3CL^{pro} May Regulate the Dimer Formation—Long range interactions have been reported to modulate dimerization and activity of 3CL^{pro} enzymes. Barrila *et al.* (57) demonstrated that mutation of a conserved amino acid Ser-147, which is distant from the dimer interface, results in a total loss of dimerization and enzymatic activity of SARS-CoV 3CL^{pro}. Although Ser-147 does not form direct interactions at the dimer interface, disruption of the dimer upon mutation stems from the fact that Ser-147 makes several interactions with other residues involved in forming a hydrogen bonding network within SARS-CoV 3CL^{pro}. Site-directed mutagenesis studies on domain III of SARS-CoV 3CL^{pro}, where N214A and S284A/T285A/I286A mutants were characterized, revealed that despite being present on an entirely different domain, these residues affect catalysis through a network of residues undergoing correlated motions across the entire protease (58, 59). Utilizing 3CL^{pro} temperature-sensitive mutants of MHV, Stobart *et al.* (60) have also demonstrated that second-site mutation physically distant from the temperature-sensitive mutation suppresses the temperature-sensitive phenotype through long range interactions, thereby regulating 3CL^{pro} enzymatic activity during polyprotein processing and virus replication.

Our studies also suggest that long range interactions among the nonconserved residues can significantly alter the properties of MERS-CoV 3CL^{pro}. A detailed analysis of nonconserved residues of MERS-CoV 3CL^{pro} among β-CoV 2c members identified hot spots, including the N-terminal finger and helix, the active site region, the inter-domain loop, and the domain III, where these residues are clustered. Several studies done on SARS-CoV 3CL^{pro} have demonstrated that amino acids from the N-terminal finger, the N-terminal helix, and domain III significantly contribute toward dimer formation.

In addition to the direct interactions at the dimer interface, correct orientation between the catalytic fold and domain III is also crucial for dimer formation. Wu *et al.* (61) showed that the most dramatic difference between the crystal structures of monomer and the ligand-bound dimer of the R298A mutant of SARS-CoV 3CL^{pro} was a 33° rotation of domain III (38). This rotation results in a steric clash between domain III from two monomers and would essentially block dimer formation. However, upon addition of a ligand, domain III of the R298A mutant adopts the correct orientation and results in the formation of a dimer structure. Similar to the SARS-CoV 3CL^{pro} R298A mutant, ligand binding into the active site of the MERS-CoV 3CL^{pro} monomer possibly stabilizes the inter-domain loop conformation that maintains domain III in the correct orientation for dimer formation. Most of the nonconserved residues within domain III are present on the surface and also are distant from the dimer interface. These residues may be involved in providing the flexibility required for conformational changes during the monomer to dimer switch.

We have identified several amino acids in MERS-CoV 3CL^{pro} that may contribute to the dimer formation upon ligand binding. However, single amino acid mutagenesis alone is unlikely to reveal significant differences in the dimerization properties. As demonstrated by Myers *et al.* (62) for ornithine decarboxylase, the response of single amino acid to ligand binding may be limited to only local conformational changes and may not have significant contribution toward dimer stability. However, local conformational changes in a network of residues may propagate larger effects that stabilize dimer formation upon ligand binding. Analysis of the nonconserved residues of MERS-CoV 3CL^{pro} discussed here sets forth a framework to perform systematic single or multiple mutagenesis studies to gain insights into the mechanism for ligand-induced dimerization of the enzyme.

Development of 3CL^{pro} Inhibitors with Broad Spectrum Specificity—Insights into the mechanistic and structural similarities as well as differences between 3CL^{pro} enzymes from different coronavirus subgroups are instrumental for the development of 3CL^{pro} inhibitors with broad spectrum specificity. To evaluate the broad spectrum specificity of our peptidomimetic compounds, we determined their inhibitory activity against 3CL^{pro} from MERS-CoV, SARS-CoV, HKU5-CoV, and HKU4-CoV. Our inhibitory data and K_I values clearly show that compounds **6**–**9** inhibit all the 3CL^{pro} enzymes tested here. The x-ray structure of MERS-CoV 3CL^{pro} in complex with compound **6** revealed that out of eight direct hydrogen bonds formed between compound **6** and MERS-CoV 3CL^{pro}, four of these hydrogen bonds involve interactions with conserved structural elements of the peptide backbone of the enzyme. Furthermore, the amino acids that form hydrogen bonds with compound **6** through side chain interactions are conserved in all the coronavirus 3CL^{pro} enzymes evaluated here, as well as 3CL^{pro} enzymes from other β-coronaviruses like MHV, OC43, and HKU1. These results suggest that canonical structural features exist among the 3CL^{pro} enzymes that can be exploited for structure-based design of broad spectrum inhibitors.

For the noncovalent inhibitor compound **11**, the x-ray structure reveals two direct hydrogen bonding interactions between

the compound and MERS-CoV 3CL^{pro}. One of the hydrogen bonds forms with the side chain ϵ-nitrogen of conserved His-166, and the second involves the backbone NH of conserved Glu-169. We speculate these interactions remain conserved in other 3CL^{pro} enzymes as well, because His-166 and Glu-169 amino acids are conserved in all 3CL^{pro} enzymes. In fact, the crystal structure of SARS-CoV 3CL^{pro} in complex with an inhibitor similar to compound **11** (PDB code 4MDS) reveals that the interactions of the inhibitor with the amino acids His-166 and Glu-169 are conserved.

The identification of 3CL^{pro}-inhibitor interactions utilizing conserved elements of the protein structure, including the peptide backbone and conserved side chains of active site residues, suggests that the development of broad-spectrum inhibitors of coronavirus 3CL^{pro} is feasible.

Our studies here demonstrate the unique properties of MERS-CoV 3CL^{pro} among β-CoV 2c members, evident from the requirement for a ligand to induce dimerization. Although the peptidomimetic compounds containing a Michael acceptor group (for example, compounds **6–9**) induce dimer formation of MERS-CoV 3CL^{pro}, the irreversible nature of their reaction with the active site cysteine ensures complete inhibition of the enzyme at stoichiometric ratios in a time-dependent manner. On the contrary, noncovalent peptidomimetic compounds (for example, compounds **10** and **11**) inhibit the enzymatic activity of MERS-CoV 3CL^{pro} only at high compound concentrations. Based on these observations, compounds that irreversibly modify the 3CL^{pro} active site may serve as better candidates for the development of inhibitors for MERS-CoV 3CL^{pro}.

Potential Complexity in the Development of MERS-CoV 3CL^{pro} Inhibitors as Antiviral Agents—Induced dimerization of MERS-CoV 3CL^{pro}, as seen in the presence of peptidomimetic inhibitors, has significant implications in the development of antiviral agents targeting MERS-CoV 3CL^{pro}. As a consequence of enzyme activation, the development of an effective antiviral agent may necessitate the development of a compound that can inhibit the MERS-CoV 3CL^{pro} monomer and stabilize it without inducing dimerization and/or inhibit the active sites of the dimer at low doses, ensuring inactivation of both the active sites within the dimer. On the contrary, it is also possible that the presence of an inhibitor could enhance the activity of MERS-CoV 3CL^{pro} to an extent that results in a complete loss of the temporal and spatial regulation of the enzymatic activity, thereby disrupting viral genome replication. Ramifications of ligand-induced dimerization and activation of MERS-CoV 3CL^{pro}, as seen in the presence of lower concentrations of inhibitor, will need to be further explored in virus-infected cells.

Author Contributions—S. T. and A. D. M. conceived and coordinated the study and wrote the paper. S. T., M. L. J. and S. E. S. J. designed, performed, and analyzed the experiments shown in Figs. 1 and 2. S. T. and L. N. P. designed, performed, and analyzed the experiments shown in Fig. 4. S. T. and M. L. J. designed, performed, and analyzed the experiments shown in Fig. 5. S. T. and A. D. M. determined the crystal structures. H. L. O., P. R. N., and A. K. G. synthesized compounds **1–9**. S. E. S. J., A. K. G., and M. R. D. provided substantial contributions to analysis and interpretation of data.

Acknowledgments— Crystallization and DNA sequencing were partially supported by the Purdue Center for Cancer Research Macromolecular Crystallography and DNA Sequencing Shared Resources, which are partially supported by National Institutes of Health Grant P30 CA023168. We acknowledge the LS-CAT and LRL-CAT beam line staff for their help in acquiring x-ray data. Use of the Advanced Photon Source, an Office of Science User Facility operated for the United States Department of Energy Office of Science by Argonne National Laboratory, was supported by the United States Department of Energy under Contract DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corp. and the Michigan Technology Tri-Corridor Grant 085P1000817. Use of the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline at Sector 31 of the Advanced Photon Source was provided by Eli Lilly Co., which operates the facility. We thank Yahira Baez-Santos, Dia Beachboard and Sergey Savinov for their helpful suggestions during manuscript preparation.

References

1. Alenius, S., Niskanen, R., Juntti, N., and Larsson, B. (1991) Bovine coronavirus as the causative agent of winter dysentery: serological evidence. *Acta Vet. Scand.* **32,** 163–170
2. Barber, D. M., Nettleton, P. F., and Herring, J. A. (1985) Disease in a dairy herd associated with the introduction and spread of bovine virus diarrhoea virus. *Vet. Rec.* **117,** 459–464
3. Perlman, S., and Netland, J. (2009) Coronaviruses post-SARS: update on replication and pathogenesis. *Nat. Rev. Microbiol.* **7,** 439–450
4. McIntosh, K., Becker, W. B., and Chanock, R. M. (1967) Growth in suckling-mouse brain of "IBV-like" viruses from patients with upper respiratory tract disease. *Proc. Natl. Acad. Sci. U.S.A.* **58,** 2268–2273
5. Cavallaro, J. J., and Monto, A. S. (1970) Community-wide outbreak of infection with a 229E-like coronavirus in Tecumseh, Michigan. *J. Infect. Dis.* **122,** 272–279
6. Hamre, D., and Procknow, J. J. (1966) A new virus isolated from the human respiratory tract. *Proc. Soc. Exp. Biol. Med.* **121,** 190–193
7. Woo, P. C., Lau, S. K., Chu, C. M., Chan, K. H., Tsoi, H. W., Huang, Y., Wong, B. H., Poon, R. W., Cai, J. J., Luk, W. K., Poon, L. L., Wong, S. S., Guan, Y., Peiris, J. S., and Yuen, K. Y. (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* **79,** 884–895
8. van der Hoek, L., Pyrc, K., Jebbink, M. F., Vermeulen-Oost, W., Berkhout, R. J., Wolthers, K. C., Wertheim-van Dillen, P. M., Kaandorp, J., Spaargaren, J., and Berkhout, B. (2004) Identification of a new human coronavirus. *Nat. Med.* **10,** 368–373
9. Cui, L. J., Zhang, C., Zhang, T., Lu, R. J., Xie, Z. D., Zhang, L. L., Liu, C. Y., Zhou, W. M., Ruan, L., Ma, X. J., and Tan, W. J. (2011) Human coronaviruses HCoV-NL63 and HCoV-HKU1 in hospitalized children with acute respiratory infections in Beijing, China. *Adv. Virol.* **2011,** 129134
10. Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D., and Fouchier, R. A. (2012) Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* **367,** 1814–1820
11. Butler, D. (2012) Clusters of coronavirus cases put scientists on alert. *Nature* **492,** 166–167
12. WHO-Global Alert and Response (2014) Middle East respiratory syndrome coronavirus (MERS-CoV)–Saudi Arabia. World Health Organization, Geneva, Switzerland
13. van Boheemen, S., de Graaf, M., Lauber, C., Bestebroer, T. M., Raj, V. S., Zaki, A. M., Osterhaus, A. D., Haagmans, B. L., Gorbalenya, A. E., Snijder, E. J., and Fouchier, R. A. (2012) Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *MBio* **3,** e00473
14. Chan, J. F., Li, K. S., To, K. K., Cheng, V. C., Chen, H., and Yuen, K. Y. (2012) Is the discovery of the novel human betacoronavirus 2c EMC/2012 (HCoV-EMC) the beginning of another SARS-like pandemic? *J. Infect.* **65,**

477–489
15. Annan, A., Baldwin, H. J., Corman, V. M., Klose, S. M., Owusu, M., Nkrumah, E. E., Badu, E. K., Anti, P., Agbenyega, O., Meyer, B., Oppong, S., Sarkodie, Y. A., Kalko, E. K., Lina, P. H., Godlevska, E. V., *et al.* (2013) Human betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. *Emerg. Infect. Dis.* **19,** 456–459
16. Memish, Z. A., Mishra, N., Olival, K. J., Fagbo, S. F., Kapoor, V., Epstein, J. H., Alhakeem, R., Durosinloun, A., Al Asmari, M., Islam, A., Kapoor, A., Briese, T., Daszak, P., Al Rabeeah, A. A., and Lipkin, W. I. (2013) Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. *Emerg. Infect. Dis.* **19,** 1819–1823
17. Reusken, C. B., Haagmans, B. L., Müller, M. A., Gutierrez, C., Godeke, G. J., Meyer, B., Muth, D., Raj, V. S., Smits-De Vries, L., Corman, V. M., Drexler, J. F., Smits, S. L., El Tahir, Y. E., De Sousa, R., van Beek, J., *et al.* (2013) Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. *Lancet Infect. Dis.* **13,** 859–866
18. Perera, R. A., Wang, P., Gomaa, M. R., El-Shesheny, R., Kandeil, A., Bagato, O., Siu, L. Y., Shehata, M. M., Kayed, A. S., Moatasim, Y., Li, M., Poon, L. L., Guan, Y., Webby, R. J., Ali, M. A., Peiris, J. S., and Kayali, G. (2013) Seroepidemiology for MERS coronavirus using microneutralisation and pseudoparticle virus neutralisation assays reveal a high prevalence of antibody in dromedary camels in Egypt, June 2013. *Euro Surveill.* **18,** pii=20574
19. Cauchemez, S., Van Kerkhove, M. D., Riley, S., Donnelly, C. A., Fraser, C., and Ferguson, N. M. (2013) Transmission scenarios for Middle East Respiratory Syndrome Coronavirus (MERS-CoV) and how to tell them apart. *Euro Surveill.* **18,** 20503
20. Assiri, A., McGeer, A., Perl, T. M., Price, C. S., Al Rabeeah, A. A., Cummings, D. A., Alabdullatif, Z. N., Assad, M., Almulhim, A., Makhdoom, H., Madani, H., Alhakeem, R., Al-Tawfiq, J. A., Cotten, M., Watson, S. J., *et al.* (2013) Hospital outbreak of Middle East respiratory syndrome coronavirus. *N. Engl. J. Med.* **369,** 407–416
21. Breban, R., Riou, J., and Fontanet, A. (2013) Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk. *Lancet* **382,** 694–699
22. Cotten, M., Watson, S. J., Kellam, P., Al-Rabeeah, A. A., Makhdoom, H. Q., Assiri, A., Al-Tawfiq, J. A., Alhakeem, R. F., Madani, H., AlRabiah, F. A., Al Hajjar, S., Al-nassir, W. N., Albarrak, A., Flemban, H., Balkhy, H. H., *et al.* (2013) Transmission and evolution of the Middle East respiratory syndrome coronavirus in Saudi Arabia: a descriptive genomic study. *Lancet* **382,** 1993–2002
23. Centers for Disease Control (2014) *MERS in the U.S.* CDC, Atlanta, GA
24. Lu, Y., Lu, X., and Denison, M. R. (1995) Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59. *J. Virol.* **69,** 3554–3559
25. Lai, M. M., and Cavanagh, D. (1997) The molecular biology of coronaviruses. *Adv. Virus Res.* **48,** 1–100
26. Ziebuhr, J., Snijder, E. J., and Gorbalenya, A. E. (2000) Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* **81,** 853–879
27. Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., and Hilgenfeld, R. (2003) Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **300,** 1763–1767
28. Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G. F., Anand, K., Bartlam, M., Hilgenfeld, R., and Rao, Z. (2003) The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 13190–13195
29. Yang, H., Bartlam, M., and Rao, Z. (2006) Drug design targeting the main protease, the Achilles' heel of coronaviruses. *Curr. Pharm. Des.* **12,** 4573–4590
30. Ghosh, A. K., Xi, K., Ratia, K., Santarsiero, B. D., Fu, W., Harcourt, B. H., Rota, P. A., Baker, S. C., Johnson, M. E., and Mesecar, A. D. (2005) Design and synthesis of peptidomimetic severe acute respiratory syndrome chymotrypsin-like protease inhibitors. *J. Med. Chem.* **48,** 6767–6771
31. Ghosh, A. K., Gong, G., Grum-Tokars, V., Mulhearn, D. C., Baker, S. C., Coughlin, M., Prabhakar, B. S., Sleeman, K., Johnson, M. E., and Mesecar, A. D. (2008) Design, synthesis and antiviral efficacy of a series of potent chloropyridyl ester-derived SARS-CoV 3CLpro inhibitors. *Bioorg. Med. Chem. Lett.* **18,** 5684–5688
32. Jacobs, J., Grum-Tokars, V., Zhou, Y., Turlington, M., Saldanha, S. A., Chase, P., Eggler, A., Dawson, E. S., Baez-Santos, Y. M., Tomar, S., Mielech, A. M., Baker, S. C., Lindsley, C. W., Hodder, P., Mesecar, A., and Stauffer, S. R. (2013) Discovery, synthesis, and structure-based optimization of a series of *N*-(*tert*-butyl)-2-(*N*-arylamido)-2-(pyridin-3-yl) acetamides (ML188) as potent noncovalent small molecule inhibitors of the severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL protease. *J. Med. Chem.* **56,** 534–546
33. Turlington, M., Chun, A., Tomar, S., Eggler, A., Grum-Tokars, V., Jacobs, J., Daniels, J. S., Dawson, E., Saldanha, A., Chase, P., Baez-Santos, Y. M., Lindsley, C. W., Hodder, P., Mesecar, A. D., and Stauffer, S. R. (2013) Discovery of *N*-(benzo[1,2,3]triazol-1-yl)-*N*-(benzyl)acetamido)phenyl) carboxamides as severe acute respiratory syndrome coronavirus (SARS-CoV) 3CLpro inhibitors: identification of ML300 and noncovalent nanomolar inhibitors with an induced-fit binding. *Bioorg. Med. Chem. Lett.* **23,** 6172–6177
34. Hsu, M. F., Kuo, C. J., Chang, K. T., Chang, H. C., Chou, C. C., Ko, T. P., Shr, H. L., Chang, G. G., Wang, A. H., and Liang, P. H. (2005) Mechanism of the maturation process of SARS-CoV 3CL protease. *J. Biol. Chem.* **280,** 31257–31266
35. Lai, L., Han, X., Chen, H., Wei, P., Huang, C., Liu, S., Fan, K., Zhou, L., Liu, Z., Pei, J., and Liu, Y. (2006) Quaternary structure, substrate selectivity and inhibitor design for SARS 3C-like proteinase. *Curr. Pharm. Des.* **12,** 4555–4564
36. Chen, S., Zhang, J., Hu, T., Chen, K., Jiang, H., and Shen, X. (2008) Residues on the dimer interface of SARS coronavirus 3C-like protease: dimer stability characterization and enzyme catalytic activity analysis. *J. Biochem.* **143,** 525–536
37. Grum-Tokars, V., Ratia, K., Begaye, A., Baker, S. C., and Mesecar, A. D. (2008) Evaluating the 3C-like protease activity of SARS-Coronavirus: recommendations for standardized assays for drug discovery. *Virus Res.* **133,** 63–73
38. Shi, J., Sivaraman, J., and Song, J. (2008) Mechanism for controlling the dimer-monomer switch and coupling dimerization to catalysis of the severe acute respiratory syndrome coronavirus 3C-like protease. *J. Virol.* **82,** 4620–4629
39. Chen, S., Jonas, F., Shen, C., Hilgenfeld, R., and Higenfeld, R. (2010) Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. *Protein Cell* **1,** 59–74
40. Li, C., Qi, Y., Teng, X., Yang, Z., Wei, P., Zhang, C., Tan, L., Zhou, L., Liu, Y., and Lai, L. (2010) Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like proteinase. *J. Biol. Chem.* **285,** 28134–28140
41. Báez-Santos, Y. M., Barraza, S. J., Wilson, M. W., Agius, M. P., Mielech, A. M., Davis, N. M., Baker, S. C., Larsen, S. D., and Mesecar, A. D. (2014) X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. *J. Med. Chem.* **57,** 2393–2412
42. Agnihothram, S., Yount, B. L., Jr., Donaldson, E. F., Huynh, J., Menachery, V. D., Gralinski, L. E., Graham, R. L., Becker, M. M., Tomar, S., Scobey, T. D., Osswald, H. L., Whitmore, A., Gopal, R., Ghosh, A. K., Mesecar, A., *et al.* (2014) A mouse model for Betacoronavirus subgroup 2c using a bat coronavirus strain HKU5 variant. *MBio* **5,** e00047–00014
43. Ghosh, A. K., Xi, K., Grum-Tokars, V., Xu, X., Ratia, K., Fu, W., Houser, K. V., Baker, S. C., Johnson, M. E., and Mesecar, A. D. (2007) Structure-based design, synthesis, and biological evaluation of peptidomimetic SARS-CoV 3CLpro inhibitors. *Bioorg. Med. Chem. Lett.* **17,** 5876–5880
44. Shen, Y., Chou, C. Y., Chang, G. G., and Tong, L. (2006) Is dimerization required for the catalytic activity of bacterial biotin carboxylase? *Mol. Cell* **22,** 807–818
45. Leslie, A. G., and Powell, H. R. (2007) Processing diffraction data with MOSFLM. *NATO Sci. Ser. ii Math.* **245,** 41–51
46. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* **276,** 307–326
47. Adams, P. D., Afonine, P. V., Bunkóczi, 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., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221
48. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
49. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21
50. DeLano, W. L. (2014) *The PyMOL Molecular Graphics System,* Version 1.7.2.1, Schrödinger, LLC, New York
51. Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma, L., Lai, B., Pei, J., Liu, Y., Chen, J., and Lai, L. (2004) Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. *J. Biol. Chem.* **279,** 1637–1642
52. Shi, J., Wei, Z., and Song, J. (2004) Dissection study on the severe acute respiratory syndrome 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: defining the extra domain as a new target for design of highly specific protease inhibitors. *J. Biol. Chem.* **279,** 24765–24773
53. Blanchard, J. E., Elowe, N. H., Huitema, C., Fortin, P. D., Cechetto, J. D., Eltis, L. D., and Brown, E. D. (2004) High-throughput screening identifies inhibitors of the SARS coronavirus main proteinase. *Chem. Biol.* **11,** 1445–1453
54. Kaeppler, U., Stiefl, N., Schiller, M., Vicik, R., Breuning, A., Schmitz, W., Rupprecht, D., Schmuck, C., Baumann, K., Ziebuhr, J., and Schirmeister, T. (2005) A new lead for nonpeptidic active-site-directed inhibitors of the severe acute respiratory syndrome coronavirus main protease discovered by a combination of screening and docking methods. *J. Med. Chem.* **48,** 6832–6842
55. Brown, P. H., Balbo, A., and Schuck, P. (2008) Characterizing protein-protein interactions by sedimentation velocity analytical ultracentrifugation. *Curr. Protoc. Immunol.* Chapter 18, Unit 18.15
56. Cheng, S. C., Chang, G. G., and Chou, C. Y. (2010) Mutation of Glu-166 blocks the substrate-induced dimerization of SARS coronavirus main protease. *Biophys. J.* **98,** 1327–1336
57. Barrila, J., Bacha, U., and Freire, E. (2006) Long-range cooperative interactions modulate dimerization in SARS 3CLpro. *Biochemistry* **45,** 14908–14916
58. Shi, J., Han, N., Lim, L., Lua, S., Sivaraman, J., Wang, L., Mu, Y., and Song, J. (2011) Dynamically-driven inactivation of the catalytic machinery of the SARS 3C-like protease by the N214A mutation on the extra domain. *PLoS Comput. Biol.* **7,** e1001084
59. Lim, L., Shi, J., Mu, Y., and Song, J. (2014) Dynamically-driven enhancement of the catalytic machinery of the SARS 3C-like protease by the S284-T285-I286/A mutations on the extra domain. *PLoS One* **9,** e101941
60. Stobart, C. C., Lee, A. S., Lu, X., and Denison, M. R. (2012) Temperature-sensitive mutants and revertants in the coronavirus nonstructural protein 5 protease (3CLpro) define residues involved in long-distance communication and regulation of protease activity. *J. Virol.* **86,** 4801–4810
61. Wu, C. G., Cheng, S. C., Chen, S. C., Li, J. Y., Fang, Y. H., Chen, Y. H., and Chou, C. Y. (2013) Mechanism for controlling the monomer-dimer conversion of SARS coronavirus main protease. *Acta Crystallogr. D Biol. Crystallogr.* **69,** 747–755
62. Myers, D. P., Jackson, L. K., Ipe, V. G., Murphy, G. E., and Phillips, M. A. (2001) Long-range interactions in the dimer interface of ornithine decarboxylase are important for enzyme function. *Biochemistry* **40,** 13230–13236
63. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14,** 1188–1190
64. Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16,** 10881–10890
65. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42,** W320–W324