Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Structural basis for catalysis and ubiquitin recognition by the *Severe acute respiratory syndrome coronavirus* papain-like protease

Chi-Yuan Chou,* Hsing-Yi Lai,‡ Hung-Yi Chen,‡ Shu-Chun Cheng,‡ Kai-Wen Cheng and Ya-Wen Chou

Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Tapei 112, Taiwan

‡ These authors made equal contributions.

Correspondence e-mail: cychou@ym.edu.tw

Received 23 August 2013
Accepted 12 November 2013

PDB reference: PL^{pro} C112S mutant, complex with ubiquitin, 4m0w

Papain-like protease (PL^{pro}) is one of two cysteine proteases involved in the proteolytic processing of the polyproteins of *Severe acute respiratory syndrome coronavirus* (SARS-CoV). PL^{pro} also shows significant *in vitro* deubiquitinating and de-ISGylating activities, although the detailed mechanism is still unclear. Here, the crystal structure of SARS-CoV PL^{pro} C112S mutant in complex with ubiquitin (Ub) is reported at 1.4 Å resolution. The Ub core makes mostly hydrophilic interactions with PL^{pro}, while the Leu-Arg-Gly-Gly C-terminus of Ub is located in the catalytic cleft of PL^{pro}, mimicking the P4–P1 residues and providing the first atomic insights into its catalysis. One of the O atoms of the C-terminal Gly residue of Ub is located in the oxyanion hole consisting of the main-chain amides of residues 112 and 113. Mutations of residues in the PL^{pro}–Ub interface lead to reduced catalytic activity, confirming their importance for Ub binding and/or catalysis. The structure also revealed an *N*-cyclohexyl-2-aminethanesulfonic acid molecule near the catalytic triad, and kinetic studies suggest that this binding site is also used by other PL^{pro} inhibitors. Overall, the structure provides a foundation for understanding the molecular basis of coronaviral PL^{pro} catalysis.

1. Introduction

Coronaviruses (CoVs) belong to the order *Nidovirales*, which are enveloped positive-stranded RNA viruses with a large genome of 26–31 kb (Gorbalenya *et al.*, 2006). They include important pathogens of humans and other animals (Weiss & Navas-Martin, 2005). In 2002–2003, a life-threatening atypical pneumonia, severe acute respiratory syndrome (SARS), caused by a novel CoV infected thousands of people, with a 10% fatality rate (World Health Organization, 2003). Since 2003, several less lethal human CoVs, such as HCoV-HKU1 and HCoV-NL63, have been reported (Fouchier *et al.*, 2004; Pyrc *et al.*, 2004; Woo *et al.*, 2005). However, in September 2012, *Middle East respiratory syndrome coronavirus* (MERS-CoV), which causes acute pneumonia and subsequent renal failure, with a 43% fatality rate, was identified (World Health Organization, 2012). Whole-genome analyses showed that both SARS-CoV and MERS-CoV are similar to bat CoVs (Li *et al.*, 2005; Lau *et al.*, 2005; Zaki *et al.*, 2012). These findings accentuate the possibility of the future re-emergence of SARS or SARS-like HCoVs that could lead to even more deadly outbreaks. Therefore, studies to understand these viruses and to develop novel antiviral inhibitors are necessary and urgent.

The coronavirus nonstructural polyproteins (pp1a and pp1ab) are cleaved by two types of viral cysteine proteases: a main protease and a papain-like protease (PL^{pro}; EC 3.4.22.46; Tan *et al.*, 2005). This processing is considered to be a suitable

antiviral target because it is required for viral maturation. In addition to its proteolytic activity, PL^{pro} also shows significant deubiquitinating and de-ISGylating (ISG, interferon stimulating gene) activities, although their physiological functions have not been completely clarified (Barretto *et al.*, 2005; Lindner *et al.*, 2005). Recent studies using cellular assays have suggested that SARS-CoV and *Murine hepatitis virus* (MHV) PL^{pro} are able to deubiquitinate interferon regulatory transcription factor 3 (IRF3), which can prevent its nuclear translocation and thus antagonize the induction of type I interferons (Zheng *et al.*, 2008; Clementz *et al.*, 2010). Other studies concluded that SARS-CoV PL^{pro} can trigger TGF-β1 production *via* ubiquitin (Ub) proteasome, p38 MAPK and ERK1/2-mediated signalling (Li *et al.*, 2012). These studies support the multifunctional nature of PL^{pro}, although the detailed mechanisms are still unclear.

The SARS-CoV PL^{pro} domain in nsp3 of the pp1a protein (residues 1541–1855) has been structurally characterized (Ratia *et al.*, 2006). The first 62 residues form an independent ubiquitin-like (Ubl) domain, while the other three domains, the palm, thumb and fingers domains, constitute a right-hand-like architecture. This topology is similar to the ubiquitin-specific proteases (USPs), one of the five distinct deubiquitinating protease (DUB) families, although their sequence identities are only about 10% (Nijman *et al.*, 2005; Ratia *et al.*, 2006). The catalytic triad, Cys112, His273 and Asp287, is located within the palm and thumb domains. In the fingertips region of the fingers domain, there is a zinc ion coordinated by four cysteine residues. Mutational studies have confirmed the importance of the catalytic triad and the zinc-binding motif for catalysis (Barretto *et al.*, 2005). Further studies suggested that in the presence of the denaturant urea, the zinc-binding domain may start to unfold during the first transition and lead to an 80% loss of enzymatic activity (Chou *et al.*, 2012). It has been demonstrated that PL^{pro} may display a differential domain-structure stability and a molten globule state in its folding. Furthermore, pp1a cleavage by SARS-CoV PL^{pro} has been found to be quite stringent, with selectivity for the substrate P1 (Gly), P2 (Gly), P4 (Leu) and P6 (hydrophobic residues) positions and with cleavage occurring between the P1 and P1′ positions (Han *et al.*, 2005). Despite this large body of knowledge on coronaviral PL^{pro}, in the absence of the structure of a substrate complex the molecular basis of its catalysis is still poorly understood.

Here, we report the crystal structure of the SARS-CoV PL^{pro} C112S mutant in complex with a Ub monomer. The structure indicates a well defined Ub core domain and a Leu-Arg-Gly-Gly C-terminus within the catalytic cleft. It also confirms the location of the oxyanion hole and the hydrogen donors. Similarly, the structure and site-directed mutagenesis suggest that mutation of the residue Glu168, which plays an important role in Ub core recognition, can cause a severe loss of DUB activity. One of the mutations, E168R, which can be used to mimic MERS-CoV PL^{pro}, has a significant loss of DUB activity. Furthermore, the crystal structure also contains an *N*-cyclohexyl-2-aminethanesulfonic acid (CHES) molecule near the catalytic triad, which was observed to possess a weak inhibitory effect. The structure provides a foundation for understanding the molecular basis of the catalytic mechanism of PL^{pro} and its substrate binding.

2. Materials and methods

2.1. Protein expression and purification

The SARS-CoV PL^{pro} (polyprotein residues 1541–1858) inserted into pET-22b(+) vector (Chou *et al.*, 2012) was expressed in *Escherichia coli* BL21 (DE3) cells (Novagen). In the construct, the secretion tag (pelB leader) was removed and the 6×His tag was retained at the C-terminus. Cultures were grown in LB medium at 37°C for 4 h, induced with 0.4 m*M* isopropyl β-D-1-thiogalactopyranoside and incubated overnight at 20°C. After centrifugation at 6000*g* at 4°C for 15 min, the cell pellets were resuspended in lysis buffer (20 m*M* Tris pH 8.5, 250 m*M* NaCl, 5% glycerol, 0.2% Triton X-100, 2 m*M* β-mercaptoethanol) and then lysed by sonication. The crude extract was then centrifuged at 12 000*g* at 4°C for 25 min to remove the insoluble pellet. The supernatant was incubated with 1 ml Ni–NTA beads at 4°C for 1 h and then loaded into an empty column. After allowing the supernatant to flow through, the beads were washed with wash buffer (20 m*M* Tris pH 8.5, 250 m*M* NaCl, 8 m*M* imidazole, 2 m*M* β-mercaptoethanol) and the protein was eluted with elution buffer (20 m*M* Tris pH 8.5, 30 m*M* NaCl, 150 m*M* imidazole, 2 m*M* β-mercaptoethanol). The protein was then loaded onto an S-100 gel-filtration column (GE Healthcare) equilibrated with running buffer (20 m*M* Tris pH 8.5, 100 m*M* NaCl, 2 m*M* dithiothreitol). The purity of the fractions collected was analyzed by SDS–PAGE and the protein was concentrated to 40 mg ml^{-1} using an Amicon Ultra-4 10 kDa centrifugal filter (Millipore). The typical yield of protein was 50 mg per litre of cell culture.

2.2. Analytical ultracentrifugation (AUC) analysis

AUC experiments were performed on an XL-A analytical ultracentrifuge (Beckman, Fullerton, California, USA) using an An-50 Ti rotor (Cheng *et al.*, 2010; Hsieh & Chou, 2011; Chou *et al.*, 2012). Sedimentation-velocity experiments were performed using a double-sector epon charcoal-filled centrepiece (0.3 or 1.2 cm) at 20°C with a rotor speed of 42 000 rev min^{-1}. SARS-CoV PL^{pro} C112S mutant (0.2 mg ml^{-1}) or Ub (1 mg ml^{-1}) protein solutions (330 µl) and reference (370 µl) solutions were loaded into the centrepiece. For PL^{pro} C112S mutant (8 mg ml^{-1}) with Ub (2 mg ml^{-1}), the overnight-incubated sample (100 µl) and reference (120 µl) solutions were loaded into a thinner centrepiece (0.3 cm). The absorbance at 250 or 280 nm was monitored in continuous mode with a time interval of 300 s and a step size of 0.003 cm. Multiple scans at different time intervals were then fitted to a continuous $c(s)$ distribution model using *SEDFIT* (Schuck, 2000). All size and shape distributions were analyzed at a confidence level of $p = 0.95$ by maximal entropy regularization and a resolution N of 200 with sedimentation coefficients between 0 and 20 S. Additionally,

Table 1
Summary of crystallographic information for the SARS-CoV PL^{pro} C112S–Ub complex.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 47.5, b = 68.3, c = 68.4, α = 90, β = 95.7, γ = 90
Resolution (Å)	30–1.4 (1.48–1.40)
R_{merge}† (%)	3.5 (39.4)
$\langle I/\sigma(I)\rangle$	31.4 (3.3)
Completeness (%)	99.6 (99.1)
Multiplicity	3.8 (3.7)
Refinement	
No. of reflections	80733 (11418)
R factor‡ (%)	15.7
Free R factor§ (%)	18.0
No. of atoms	
Total	3535
Protein	3176
Ligand/ion	44/2
Water	313
B factors (Å^2)	
Protein	17.9
Ligand/ion	26.1/22.0
Water	25.4
R.m.s.d.	
Bond lengths (Å)	0.007
Bond angles (°)	1.4
Ramachandran analysis¶ (%)	
Favoured	98.0
Allowed	2.0
Disallowed	0
Rotamer outliers¶ (residues)	3

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the integrated intensity of a given reflection and $\langle I(hkl)\rangle$ is the mean intensity of multiple corresponding symmetry-related reflections. ‡ $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. § The free R factor is the R factor calculated using a random 5% of data that were excluded from the refinement. ¶ Ramachandran analysis and the rotamer outlier check were carried out using *MolProbity* (Chen *et al.*, 2010).

the AUC results were fitted to a hetero-association model ($A + B = AB$) using the *SEDPHAT* program to calculate the dissociation constant (Brown & Schuck, 2006).

2.3. Protein crystallization

Crystals of the SARS-CoV PL^{pro} C112S mutant in complex with bovine ubiquitin (Sigma) were obtained at 22°C by the sitting-drop vapour-diffusion method. The PL^{pro}–Ub complex was formed by mixing PL^{pro} and Ub in a 1:1 molar ratio. The final concentrations were 8 and 2 mg ml^{-1}, respectively. Initial crystal screens were set up at 22°C after incubating the protein mixture at 4°C overnight. After screening and optimization, crystals could be grown using a reservoir solution consisting of 18% PEG 3000, 0.1 *M* CHES pH 9.5. Single block-shaped crystals of 0.5 mm in size grew after 3 d. Analysis of the crystals by SDS–PAGE confirmed the presence of both PL^{pro} and Ub (Supplementary Fig. S1[1]). All crystals were cryoprotected in reservoir solution supplemented with 12%(*v*/*v*) glycerol and were flash-cooled in liquid nitrogen.

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: RR5055).

2.4. X-ray data collection, processing and structure determination

X-ray diffraction data were collected at 100 K on the SPXF beamline 13C1 at the National Synchrotron Radiation Research Center, Taiwan, ROC using a ADSC Quantum-315r

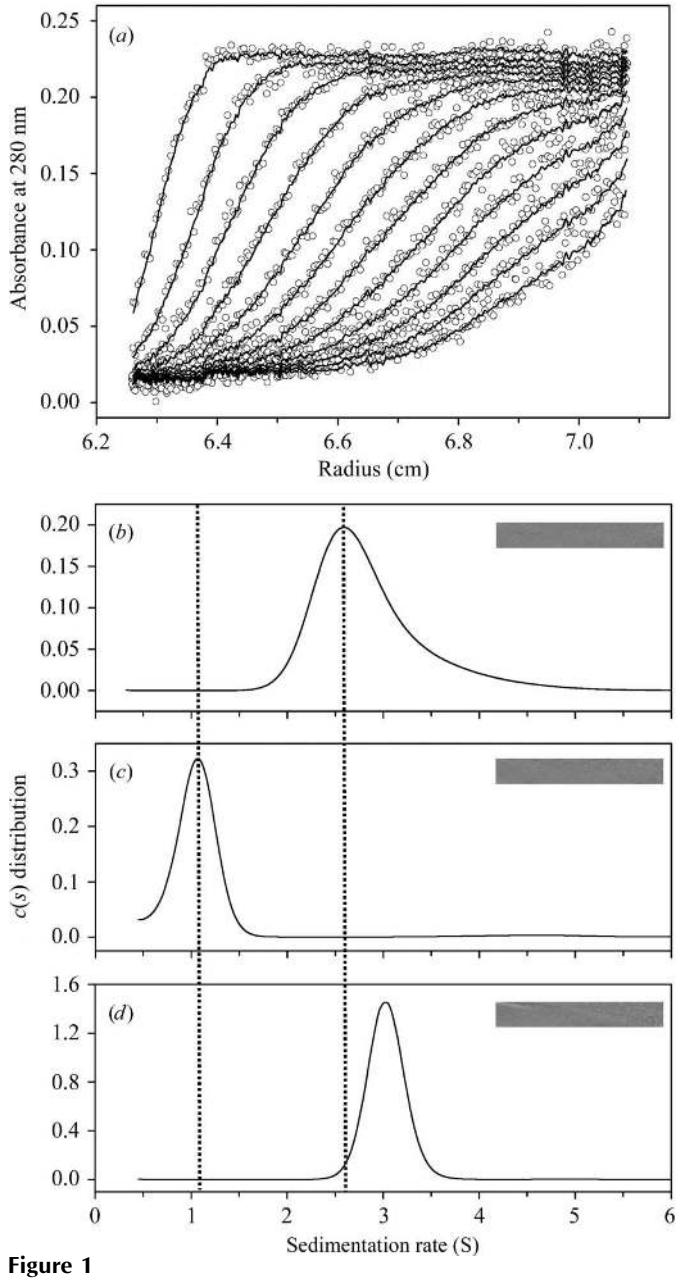

Figure 1
The SARS-CoV PL^{pro} C112S mutant forms a stable complex with Ub. (*a*) Traces of absorbance at 280 nm of the PL^{pro} C112S mutant in 50 m*M* phosphate buffer pH 7.4 during the sedimentation-velocity experiment. The protein concentration was 0.2 mg ml^{-1}. For clarity, only every third scan is shown. The circles represent experimental data and the lines are the results obtained after fitting to the Lamm equation using *SEDFIT* (Chou *et al.*, 2011; Schuck, 2000). (*b*), (*c*) and (*d*) show the continuous $c(s)$ distributions of PL^{pro} C112S mutant (0.2 mg ml^{-1}), Ub (1 mg ml^{-1}) and C112S mutant with Ub (8 and 2 mg ml^{-1}), respectively. The two vertical dotted lines indicate the positions of PL^{pro} (s = 2.72) and Ub (s = 1.04) alone. The residual bitmaps of the raw data and the best-fit results are shown in the insets.

CCD detector (X-ray wavelength of 0.976 Å). The diffraction images were processed and scaled with the *HKL*-2000 package (Otwinowski & Minor, 1997). The structure was solved by the molecular-replacement method with *Phaser* (McCoy *et al.*, 2007) using the structure of wild-type PL^{pro} (PDB entry 2fe8; Ratia *et al.*, 2006) as the search model. Manual rebuilding of the structure model was performed with *Coot* (Emsley & Cowtan, 2004). Structure refinement was carried out with *REFMAC* (Murshudov *et al.*, 2011). The data-processing and refinement statistics are summarized in Table 1. The crystal structure has been deposited in the Protein Data Bank (PDB entry 4m0w).

2.5. Site-directed mutagenesis and enzyme-kinetic assay

Mutants of SARS-CoV PL^{pro} were produced using the QuikChange kit (Stratagene) and were verified by DNA sequencing. The enzyme activity of PL^{pro} was measured by a colorimetry-based peptide-cleavage assay involving the 6-mer peptide substrate FRLKGG-*para*-nitroanilide (FG6-pNA; GL Biochem Ltd, Shanghai, People's Republic of China). This substrate is cleaved at the Gly-pNA bond to release free pNA, which can be monitored by the increase in absorbance at 405 nm. The amount of pNA released by proteolysis can be calculated using a standard curve generated with analytical grade pNA. The protease-activity assay was performed in 50 m*M* phosphate pH 7.4 at 30°C. The substrate stock solution was made up at 6 m*M* and the working concentrations were from 0.5 to 5 m*M*. The concentration of the wild-type PL^{pro}, E168A, E168D, E168R and Y265F mutants was 1 µ*M*, while that of the L163Q, D165A and Y265A mutants was 9 µ*M*. The steady-state enzyme-kinetic parameters were obtained by fitting the initial velocity data to the Michaelis–Menten equation. The program *SigmaPlot* 12 (Systat Software Inc., Richmond, California, USA) was used for data analysis.

2.6. Deubiquitination assays

The fluorogenic substrate ubiquitin-7-amino-4-trifluoromethylcoumarin (AFC; Boston Biochem Co., USA) added at 0.5 µ*M* to 50 m*M* phosphate buffer pH 7.4 and varying concentrations of wild-type PL^{pro} or its mutants (0.17 µ*M* for wild type, L163Q, E168D and Y265F and 0.51 µ*M* for D165A, E168A, E168R and Y265A) were used for deubiquitination assays. The enzymatic activity at 30°C was determined by continuously monitoring the fluorescence emission increase upon substrate cleavage at excitation and emission wavelengths of 350 and 485 nm, respectively, in a Perkin Elmer LS 50B luminescence spectrometer.

2.7. Inhibition assays

For the inhibition studies, 0.5 m*M* FG6-pNA and various concentrations of CHES (0–600 m*M*) were mixed with PL^{pro} (0.4 µ*M*) at 30°C and the activity was continuously monitored. The initial velocities of the inhibition reactions were plotted against the concentrations of CHES to obtain the IC_{50} value.

To characterize the mutual effect of CHES and a known PL^{pro} inhibitor, 6-thioguanine (6TG; Chou *et al.*, 2008), the activity of PL^{pro} was measured in the presence of 6TG (0–15 µ*M*) at various CHES concentrations (0–200 m*M*). The FG6-pNA concentration was held constant at 0.4 m*M* and the enzyme was at 1 µ*M*. Data for multiple inhibitions were fitted to the equation

$$v = \frac{v_0}{\left(1 + \frac{I}{K_i} + \frac{J}{K_j} + \frac{IJ}{\alpha K_i K_j}\right)}, \quad (1)$$

where v is the initial velocity in the presence of both inhibitors, I and J are the concentrations of the two inhibitors, v_0 is the velocity in the absence of inhibitors, K_i and K_j are the apparent dissociation constants for the two inhibitors and α is a measure of the degree of interaction of the two inhibitors (Yonetani & Theorell, 1964).

3. Results and discussion

3.1. A stable complex of SARS-CoV PL^{pro} with Ub

To delineate the molecular basis of catalysis by SARS-CoV PL^{pro}, the goal of our experiment was to determine the binding mode of peptidyl substrates or Ub to SARS-CoV PL^{pro}. However, efforts to crystallize the wild-type PL^{pro} complex were unsuccessful. Therefore, we tried to replace the wild-type PL^{pro} with less active or inactive mutants that still maintained the substrate-binding affinity. After many trials, an inactive catalytic triad mutation, C112S, was co-crystallized with Ub in a 1:1 molar ratio.

To confirm that the formation of the complex of the PL^{pro} C112S mutant with Ub was genuine and not because of crystal packing, we performed AUC experiments to characterize the PL^{pro}–Ub complex. Fig. 1(*a*) shows a typical absorbance trace at 280 nm of the PL^{pro} C112S mutant during the experiment. After fitting the signals to a continuous size-distribution model, it was clear that the C112S mutant was monomeric, with a sedimentation coefficient of 2.72 S (Fig. 1*b*), consistent with that of wild-type PL^{pro} (Chou *et al.*, 2012). AUC analysis indicated that Ub is also a monomer, with a sedimentation coefficient of 1.04 S (Fig. 1*c*). After overnight incubation at the crystallization concentration (PL^{pro} at 8 mg ml^{-1} and Ub at 2 mg ml^{-1}), a PL^{pro} C112S–Ub complex with a sedimentation coefficient of 3.03 S could be observed (Fig. 1*d*), which is consistent with a complex with 1:1 stoichiometry (Dam *et al.*, 2005). In contrast, AUC experiments for wild-type PL^{pro} did not yield the PL^{pro}–Ub complex, only separate PL^{pro} and Ub signals (Supplementary Fig. S2). Next, the AUC results were fitted to the $A + B$ hetero-association model (Brown & Schuck, 2006) and a dissociation constant of 6.7 µ*M* was obtained. This confirmed that at the crystallization concentration, where both proteins are at 230 µ*M*, most of the PL^{pro} C112S mutant molecules should form a stable complex with Ub.

3.2. Overall structure of PL^{pro} in complex with Ub

The crystal structure of the SARS-CoV PL^{pro} C112S–Ub complex was determined at 1.4 Å resolution (Table 1 and

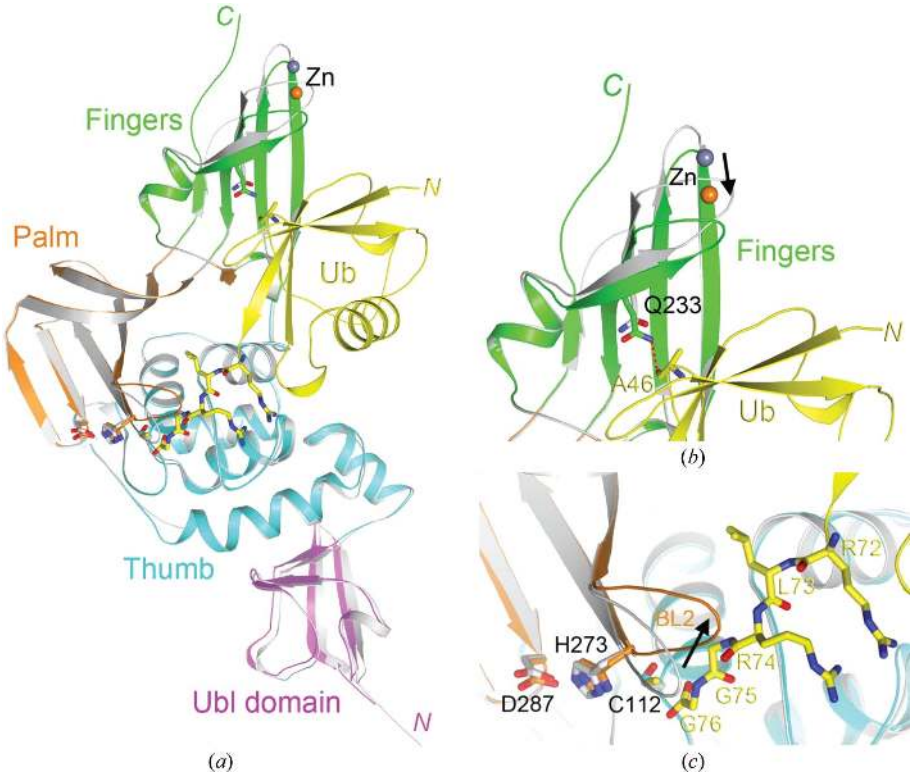

Figure 2
Structure of the SARS-CoV PL^{pro} C112S mutant in complex with Ub. (*a*) Overlay of the PL^{pro} C112S–Ub complex (colour) and free wild-type PL^{pro} (grey). The Ubl, thumb, palm and fingers domains of the PL^{pro} C112S mutant in complex with Ub are coloured magenta, cyan, orange and green, respectively, and the Ub is coloured yellow. The orange and grey spheres show the location of the zinc ion in the two structures. The movement of the zinc-binding site (*b*) and the β14–β15 (BL2) loop (*c*) are indicated by arrows. The residues are shown as sticks and hydrogen bonding is indicated by red dashed lines. All structure figures in this paper were produced using *PyMOL* (http://www.pymol.org).

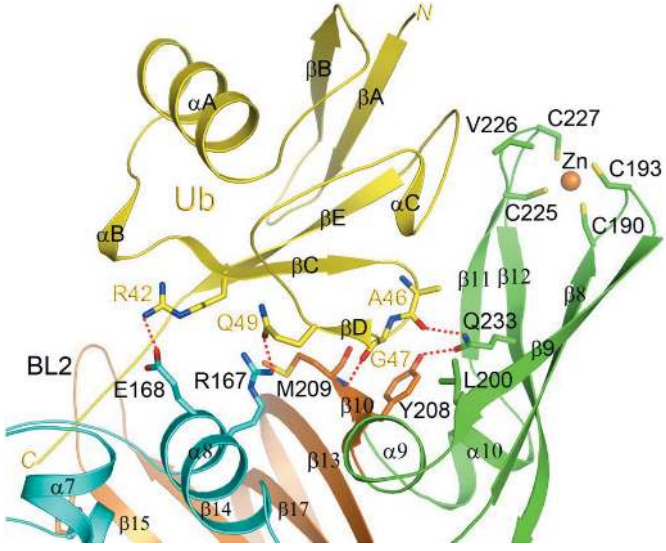

Figure 3
Interactions of the Ub core with SARS-CoV PL^{pro}. Schematic drawing showing the detailed interactions between PL^{pro} C112S (coloured as in Fig. 2) and the Ub core (yellow). Hydrogen-bonding and ion-pair interactions are indicated by red dashed lines.

Fig. 2). The crystal belonged to space group $P2_1$, with unit-cell parameters a = 47.5, b = 68.3, c = 68.4 Å, β = 95.7°. The current atomic model containing one PL^{pro} and one Ub molecule in the crystallographic asymmetric unit agrees well with the crystallographic data and the expected values of geometric parameters (Table 1). 98% of the residues are in the most favoured region of the Ramachandran plot and none are in the disallowed region.

The overall structure of PL^{pro} C112S in the Ub complex (Fig. 2*a*) is similar to that of free PL^{pro} (Ratia *et al.*, 2006). The r.m.s. distance between equivalent C$^{\alpha}$ atoms of the two structures is 0.56 Å. The zinc-binding motif in our structure is more bent toward Ub, and the Zn atom shows a 3.8 Å shift relative to that in free PL^{pro} (Fig. 2*b*). Similarly, the β14–β15 loop, which is called the BL2 loop in previous studies (Ratia *et al.*, 2006), also has a large movement toward the catalytic cleft (Fig. 2*c*). However, the catalytic triad does not show significant changes.

3.3. Interactions of PL^{pro} with the Ub core

The structure shows PL^{pro} in a noncovalent complex with Ub (Fig. 3 and 4). PL^{pro} interacts with several residues of the Ub core (residues *1–72*; throughout the paper Ub residues are given in italics and PL^{pro} residues in roman font) by its cupped-hand structure consisting of thumb, palm and finger domains, while the four C-terminal residues (*73–76*) of Ub are bound to the narrow active-site channel of PL^{pro}.

Among the nine PL^{pro} residues located within 4 Å of the Ub core, three residues, Leu200, Tyr208 and Val226, make hydrophobic contacts with the Ub core (Fig. 3). Four residues are engaged in polar interactions, including one ion pair between the side-chain guanidinium group of *Arg42* and the side-chain carboxylate group of Glu168 and three hydrogen bonds: from the main-chain amide of *Ala46* to the side chain of Gln233, from the main-chain carbonyl O atom of *Gly47* to the main-chain amide of Met209 and from the side chain of *Gln49* to the side chain of Arg167 (Fig. 3). Only one of these interactions is found in the fingers domain (Gln233), while the others are located in the palm (Met209) and thumb (Arg167 and Glu168) domains. In addition, there are 36 solvent molecules in the interface between PL^{pro} and the Ub core mediating interactions between the two proteins. These observations indicate that the PL^{pro}–Ub core interface is primarily hydrophilic in nature and that the Ub core may be loosely associated with PL^{pro}.

To identify the conserved PL^{pro} residues interacting with Ub, we compared the SARS-CoV PL^{pro} amino-acid sequence with those of PL^{pro} from five other CoVs (Fig. 5). The multiple sequence alignment revealed that the residues interacting with the Ub core were poorly conserved. However, the Glu168 residue is an aspartate residue in BCoV and two other human CoVs, while the equivalent residue in MERS-CoV is an arginine (Barretto *et al.*, 2005; Zaki *et al.*, 2012; Han *et al.*, 2005). To assess the functional importance of this residue, we characterized the peptide-cleavage and DUB activities of the E168A, E168D and E168R mutants. The kinetic assays indicated that the mutations do not significantly affect the proteolytic activity towards a peptide substrate (Table 2). On the other hand, the E168A and E168R mutants show a 25-fold decrease in DUB activity, while the E168D mutant has a 1.3-fold increase (Supplementary Fig. 3 and Table 2). AUC analysis also showed that the C112S/E168R double mutant cannot form a stable complex with Ub (Supplementary Fig. S4). These experimental data suggest that the ion pair between Glu168 (or Asp168) and *Arg42* is important for the recognition of the Ub core by PL^{pro} and for its DUB activity. Moreover, the decrease in the DUB activity of the E168R mutant also suggests that MERS-CoV PL^{pro} may not show significant DUB activity. Earlier studies provided evidence that SARS-CoV and MHV PL^{pro} can bind to IRF3, cause its deubiquitination and prevent its nuclear translocation (Zheng *et al.*, 2008), and can further reduce interferon induction and promote viral growth in infected cells (Zheng *et al.*, 2008). In contrast, without significant DUB activity, MERS-CoV PL^{pro} may not efficiently inhibit interferon induction; however, its effect on viral infection still needs to be further investigated.

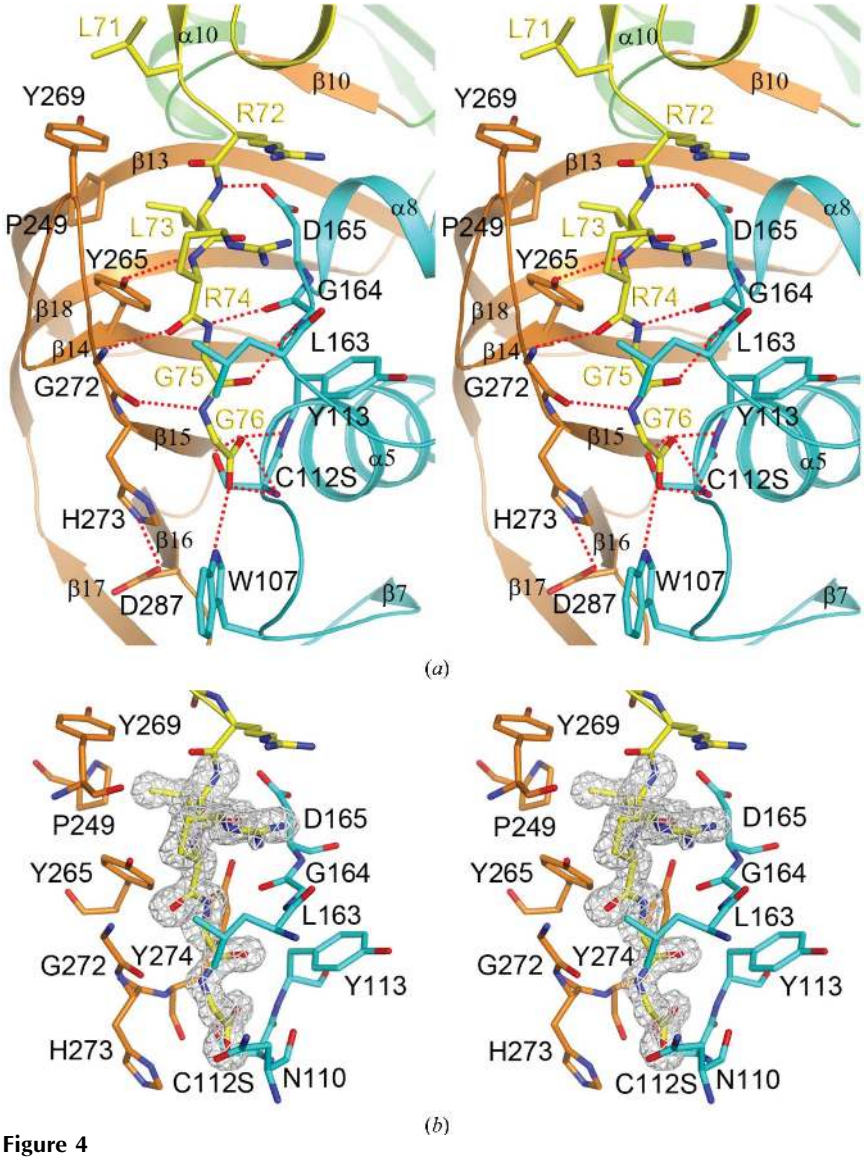

Figure 4
Binding mode of the C-terminus of Ub in the active site of SARS-CoV PL^{pro}. (*a*) Schematic drawing in stereo showing the detailed interactions between the active site of the PL^{pro} C112S mutant and the C-terminus of Ub. Hydrogen-bonding and ion-pair interactions are indicated by red dashed lines. (*b*) The OMIT $F_o - F_c$ electron-density map for the final four C-terminal residues of Ub at 1.4 Å resolution, contoured at 3.5σ. The interacting residues of PL^{pro} are shown as sticks.

3.4. Interactions of PL^{pro} with Ub C-terminal residues

The four Ub C-terminal residues *Leu73*-*Arg74*-*Gly75*-*Gly76* have clearly defined electron density and are located in the narrow active-site channel of PL^{pro} (Figs. 4*a* and 4*b*). In fact, these four residues are likely to mimic the P4–P1 residues of a substrate. Their backbone atoms form a network of hydrogen-bonding interactions with PL^{pro}, including *Leu73*–Asp165 side chain, *Arg74*-Tyr265 side chain, *Arg74*–Gly272, *Gly75*–Gly164 and *Gly76*–Gly272, and the C-terminal carboxylate of *Gly76* makes interactions with Trp107, C112S and Tyr113. In addition, *Gly75* (P2-mimicking) and *Gly76* (P1-mimicking) are located in a narrow and hydrophobic cavity consisting of Asn110, Tyr113, Tyr274 and Leu163 (Fig. 4*b*). Previous studies have confirmed that the Y274A mutant is inactive (Barretto *et al.*, 2005). Here, we mutated Leu163 to glutamine and found that the K_m and k_{cat} values of the L163Q mutant determined using the peptide-cleavage assay show a 1.6-fold increase and a ninefold decrease, respectively, compared with those of wild-type PL^{pro} (Table 2). The mutation also lead to an 80% decrease in DUB activity. These data suggest that in addition to hovering above the active site for substrate binding, the hydrophobic Leu163 may also serve to enhance the nucleophilicity of Cys112 (Ratia *et al.*, 2006).

The residue Tyr269 in the BL2 loop show a significant movement compared with that

Table 2
The kinetic parameters and DUB activity of SARS-CoV PL^{pro} and its mutants.

	Peptide cleavage			Deubiquitination
Protein	K_m† (m*M*)	k_{cat}† (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})	Relative activity‡
Wild type	0.73 ± 0.08	0.79 ± 0.04	1.08 ± 0.13	1
Ub core related				
E168A	1.26 ± 0.13	0.29 ± 0.01	0.23 ± 0.03	0.04
E168D	1.10 ± 0.11	0.61 ± 0.03	0.55 ± 0.06	1.32
E168R	1.01 ± 0.20	0.36 ± 0.03	0.36 ± 0.08	0.04
Ub C-terminus related				
L163Q	1.20 ± 0.05	0.09 ± 0.002	0.08 ± 0.004	0.20
D165A	3.33 ± 0.40	0.04 ± 0.002	0.01 ± 0.001	<0.01
Y265A	2.48 ± 0.73	0.01 ± 0.003	0.004 ± 0.001	0.01
Y265F	1.35 ± 0.30	0.61 ± 0.08	0.45 ± 0.12	0.43

† The kinetic data of wild-type PL^{pro} and the mutants were fitted to the Michaelis–Menten equation. All of the assays were repeated several times to ensure reproducibility. ‡ A fixed concentration of Ub-AFC (0.5 μ*M*) was used. The relative activity was corrected by normalizing the initial velocity of wild-type PL^{pro} to unity and treating all others with respect to this standardized value.

of the free PL^{pro} (Supplementary Fig. S5). It results in the side chain of *Leu71* (P6-mimicking) making a hydrophobic contact with Tyr269, and the side chain of *Leu73* (P4-mimicking) pointing towards a hydrophobic pocket consisting of residues Pro249, Tyr265 and Tyr269 (Fig. 4*a*). However, in contrast to USP–Ub complexes (Renatus *et al.*, 2006) and to the predicted PL^{pro}–Ub binding model (Ratia *et al.*, 2006), the side-chain guanidinium group of *Arg72* (P5-mimicking) does not interact with PL^{pro}. Overall, our observations confirm the substrate specificity of SARS-CoV PL^{pro} at the P1, P2, P4 and P6 positions, which are important for optimal substrate recognition and cleavage (Han *et al.*, 2005).

Previous studies predicted that the oxyanion is within hydrogen-bonding distance of the Trp107 side chain and the Cys112 backbone amide (Ratia *et al.*, 2006). In the present structure, the electron density at the *Gly76* terminus shows a carboxylate group (Fig. 4*b*). One of the O atoms of the *Gly76* carboxylate is hydrogen-bonded to the backbone amide of C112S and the indole N atom of Trp107, with distances of 2.84 and 2.79 Å, respectively, while another O atom is close to the side-chain hydroxyl group (2.46 Å) and the main-chain amide (3.02 Å) of C112S and the backbone amide of Tyr113 (2.96 Å; Fig. 4*a*). Of these two O atoms, that pointing towards the PL^{pro} is more buried and suitable as the oxyanion during catalysis. This suggests that the main-chain amides of residues 112 and 113 may form the oxyanion hole. Similarly, the unusually short distance between the carboxyl O atom of *Gly76* and the hydroxyl group of C112S suggests that it may be a low-barrier hydrogen bond, as found in nonaqueous active sites of enzymes (Cleland *et al.*, 1998). In comparison to the active site of the USP2–Ub complex, the deprotonated S^{γ} atom of the Cys nucleophile is rotated by 120° and points away from Ub *Gly76* because of electrostatic repulsion between the two negatively charged moieties (Renatus *et al.*, 2006). The charge–charge repulsion may destabilize the protease–ligand complex. The difference between the hydroxyl and sulfhydryl groups and the smaller Ub contact of PL^{pro} (see below) are able to explain why the C112S mutant can stably bind to Ub while wild-type PL^{pro} cannot.

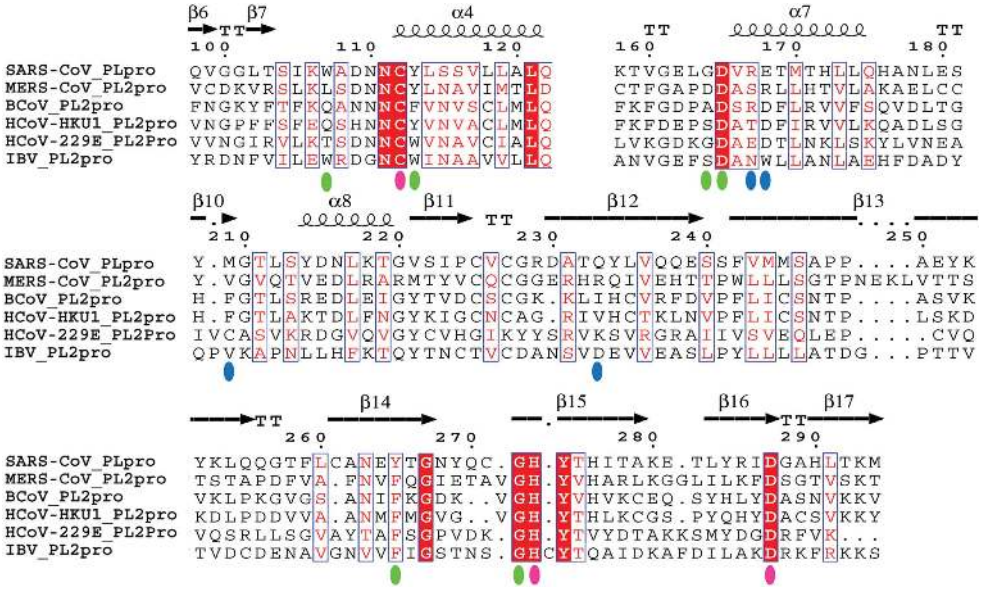

Figure 5
Sequence alignment of various type 2 coronaviral PL^{pro}s. Group 1, HCoV229E; group 2a, BCoV and HCoV-HKU1; group 2b, SARS-CoV and MERS-CoV; group 3, IBV. The blue and green ovals indicate residues making polar contact with the Ub core and C-terminus, respectively, while the magenta ovals indicate the catalytic triad. The red boxes with white characters indicate residues with strict identity, while blue boxes with red characters show residues with similarity in a group. Accession numbers are as follows: SARS-CoV, AY291451; MERS-CoV, NC019843.2; BCoV, NC012948.1; HCoV-HKU1, DQ415909.1; HCoV-229E, JX503060.1; IBV, JQ088078.1.

The BL2 loop of many DUBs, such as HAUSP, USP2 and USP14, serves a regulatory role in catalysis (Hu *et al.*, 2002, 2005; Renatus *et al.*, 2006). Here, we provide direct structural evidence that the conformation of the BL2 loop (β14–β15 loop) of PL^{pro} is also influenced by substrate binding (Fig. 2*c* and Supplementary Fig. S5). In the presence of substrate, residue Gly272 in the BL2 loop interacts with *Arg74* (P3-mimicking) and *Gly76* (P1-mimicking) (Fig. 4*a*). Furthermore, another residue in this loop, Tyr269, makes hydrophobic interactions with *Leu71* (P6-mimicking) and *Leu73* (P4-mimicking). All of these interactions bring the BL2 loop closer to the C-terminus of Ub. On the other hand, a closed BL2 loop in free PL^{pro} may make it inactive. A series of 'BL2 loop' bound inhibitors, such as GRL0617, whose binding site is close to Asp165 and Tyr274, have been

shown to make the BL2 loop more closed, with an IC_{50} of 0.3–0.6 μ*M* (Ratia *et al.*, 2008; Ghosh *et al.*, 2010; Lee *et al.*, 2013).

3.5. Asp165 and Tyr265 are also important for PL^{pro} catalysis

Besides Leu163 and Glu168, we mutated two other residues, Asp165 and Tyr265, which are highly conserved in CoV PL^{pro} (Fig. 5; Barretto *et al.*, 2005; Han *et al.*, 2005; Zaki *et al.*, 2012) and important for binding to the *Leu73* (P4-mimicking) and *Arg74* (P3-mimicking) backbone (Fig. 4*a*). Our kinetic data and deubiquitination assay show that the D165A mutant is essentially inactive, with a 108-fold decrease in k_{cat}/K_m, as a result of a 20-fold decrease in k_{cat} and a 4.6-fold increase in K_m (Table 2). The increase in the K_m value of the D165A mutant indicates that the direct interaction between the Asp165 side chain and the backbone of *Leu73* (P4) is important for substrate binding of PL^{pro}.

Compared with wild-type PL^{pro}, the Y265A mutant showed only 1% of the DUB activity and a 79-fold decrease in k_{cat} and a 3.4-fold increase in K_m (Table 2). This mutant was reported to be inactive in an earlier study (Barretto *et al.*, 2005). In comparison, the Y265F mutant showed only a 2.4-fold decrease in proteolytic activity based on k_{cat}/K_m and 43% of the DUB activity (Table 2). This indicates the importance of a phenyl ring at this position, while the hydrogen bond to the backbone of Ub *Arg74* (P3-mimicking) may be dispensable.

3.6. Comparison with other USP–Ub complexes

PL^{pro} shows a similar topology to several cellular DUBs (Ratia *et al.*, 2006). Here, we compare the structure of the PL^{pro}–Ub complex with those of the USP2–Ub (PDB entry 2hd5; Renatus *et al.*, 2006), USP14–Ub aldehyde (PDB entry 2ayo; Hu *et al.*, 2005) and HAUSP–Ub aldehyde (PDB entry 1nbf; Hu *et al.*, 2002) complexes (Supplementary Fig. S6). Interestingly, PL^{pro} has shorter fingers and lacks the BL1 loop (Supplementary Fig. 6*a*). Only approximately 1000 Å^2 of the surface area of PL^{pro} (640 Å^2 by the Ub core and 360 Å^2 by the Ub C-terminus) is buried in the interface, which is smaller than in the USP2–Ub (1900 Å^2; Supplementary Fig. S6*b*), USP14–Ub aldehyde (1500 Å^2; Supplementary Fig. S6*c*) and HAUSP–Ub aldehyde (1700 Å^2) complexes. Less contact of PL^{pro} with Ub and the charge–charge repulsion of Cys112 with *Gly76* may both influence Ub binding. It may explain why wild-type PL^{pro} and Ub are unable to form a stable complex (Supplementary Fig. S2), which may aid in product release. In contrast, a more stable USP2–Ub complex leads to product inhibition by Ub, with a K_i of 2.8 μ*M* (Renatus *et al.*, 2006). Similarly, the current structure of PL^{pro} C112S–Ub was captured in a noncovalent enzyme–product complex. As many

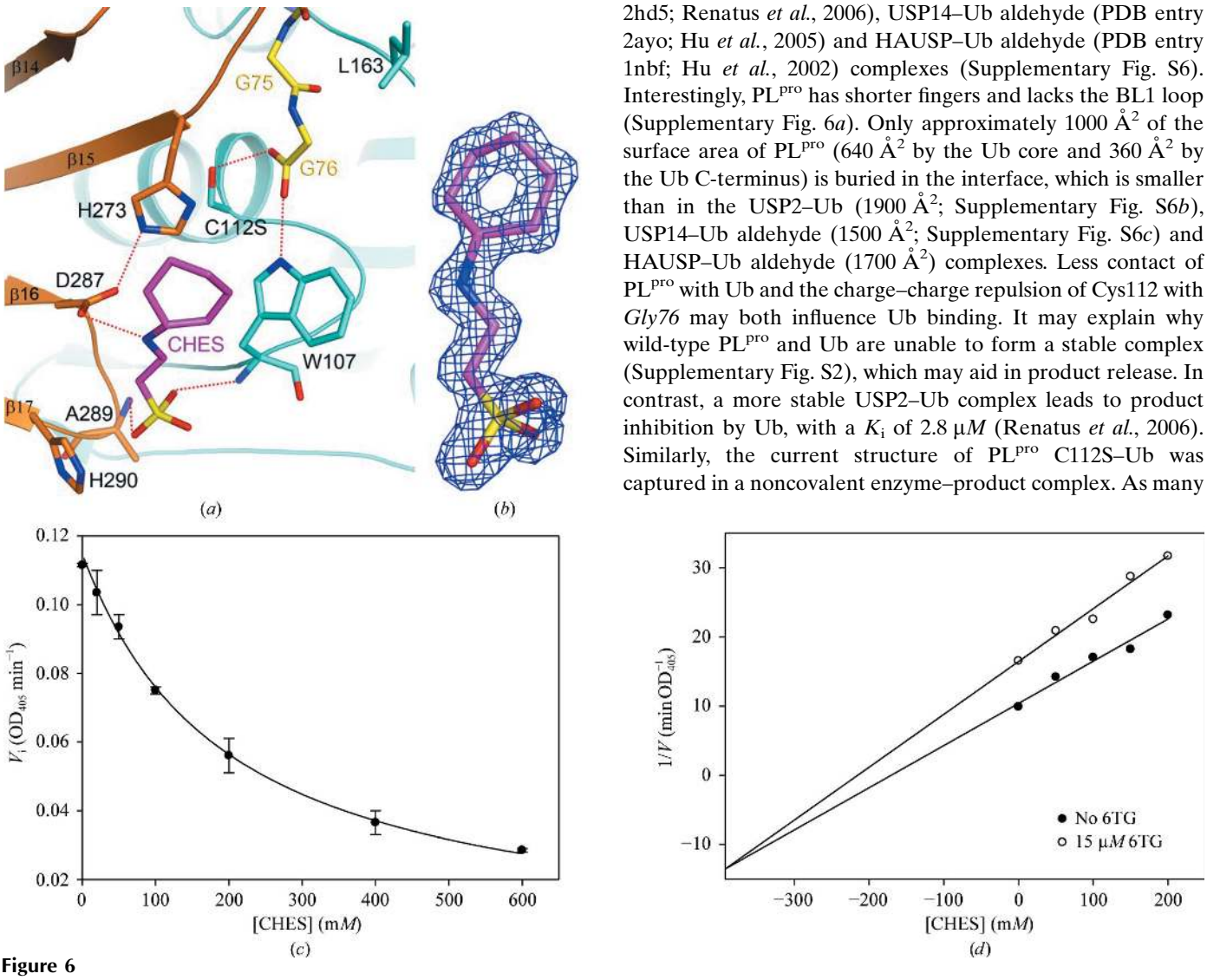

Figure 6
The interaction of CHES with the PL^{pro} active site and its inhibitory effect. (*a*) Hydrogen-bonding (red dashed lines) and hydrophobic interactions between CHES (magenta) and PL^{pro} (orange for palm domain and cyan for thumb domain). (*b*) The OMIT $F_o - F_c$ electron density for CHES at 1.4 Å resolution, contoured at 3.5σ. (*c*) Inhibition of PL^{pro} by CHES. The experiments were performed in triplicate and the dots and bars represent the mean and errors, respectively. The curve is the best fit of the data for IC_{50} calculation. (*d*) Two inhibition patterns for CHES and 6TG. The points are the reciprocal of the experimental velocities, and the lines are the best fit of the data to (1). According to the results, the two inhibitors antagonized each other's binding to PL^{pro} (α = 2.3).

cysteine proteases are normally observed with their catalytic cysteine forming the catalytic triad, mutation of the cysteine to serine may be a good strategy to delineate the structures of these enzyme–product complexes.

3.7. CHES in the active site of PL^{pro}

Besides Ub, there is a CHES molecule with clear electron density in the active site of our structure (Figs. 6*a* and 6*b*). The cyclohexyl C atoms of CHES lie between the imidazole group of His273, one of the triad residues, and the indole group of Trp107. Its ammonium group is ion-paired with the side-chain carboxylate of Asp287, the third member of the catalytic triad, while its sulfonic acid makes hydrogen bonds to the main-chain amides of Trp107 and Ala287. The bound position of CHES suggests that it may occupy part of the S1′ and S2′ sites and interfere with the catalysis of PL^{pro}. Our kinetic studies showed an IC_{50} value of 188 m*M*, indicating that the inhibitory activity is weak (Fig. 6*c*). Nonetheless, this binding site is likely to accommodate other compounds. We have previously reported that two thiopurine analogues, 6-mercaptopurine (6MP) and 6TG, are competitive and reversible inhibitors of SARS-CoV PL^{pro}, with a K_i of 15 μ*M* (Chou *et al.*, 2008). Docking calculations suggested that 6MP and 6TG may interact with the catalytic triad of PL^{pro} and some other USPs (Chen *et al.*, 2009). To characterize the mutual effect of these thiopurine analogues and CHES, multiple inhibition studies were performed (Fig. 6*d*). The experiments were carried out by measuring the initial reaction velocities of PL^{pro} at various concentrations of 6TG and CHES. The substrate concentrations were held constant at subsaturating levels, which allows both inhibitors to bind to the enzyme (Benson *et al.*, 2008). By fitting the data to (1), the results showed that the lines intersect below the *x* axis in the Yonetani–Theorell plot (Yonetani & Theorell, 1964) and the α value is 2.3. This indicates that 6TG and CHES antagonize each other's binding, suggesting that they may interact with PL^{pro} at the same or at close regions (Copeland, 2000).

4. Conclusion

The crystal structure of the SARS-CoV PL^{pro} C112S–Ub complex gives the first molecular insights into Ub binding by coronaviral PL^{pro}. The Ub core binds to the palm, thumb and fingers domains of PL^{pro}, while its final four C-terminal residues bind into a narrow channel by a network of hydrogen bonds and reach towards the active site. Mutagenesis studies further confirmed that several PL^{pro} residues that recognize the Ub core or C-terminal residues, such as Glu168, Leu163, Asp165 and Tyr265, are essential or required for its DUB or proteolytic activities, respectively. One of the mutations, E168R, which can be used to mimic MERS-CoV PL^{pro}, shows minor DUB activity. Furthermore, in the present structure there is a CHES molecule close to the catalytic triad, showing a weak inhibitory effect. Our structure provides key insights into the molecular mechanism of catalysis of this family of enzymes as well as an inhibitory approach using a novel binding site.

We thank Gu-Gang Chang and Liang Tong for helpful discussions. This research was supported by grants from the National Science Council, Taiwan (98-2320-B-010-026-MY3 and 101-2320-B-010-061) to CYC. We also thank NYMU for financial support (Aim for Top University Plan from Ministry of Education). Portions of this research were carried out at the National Synchrotron Radiation Research Center, a national user facility supported by the National Science Council of Taiwan, ROC. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Core Facility Program for Biotechnology.

References

Barretto, N., Jukneliene, D., Ratia, K., Chen, Z., Mesecar, A. D. & Baker, S. C. (2005). *J. Virol.* **79**, 15189–15198.
Benson, B. K., Meades, G. Jr, Grove, A. & Waldrop, G. L. (2008). *Protein Sci.* **17**, 34–42.
Brown, P. H. & Schuck, P. (2006). *Biophys. J.* **90**, 4651–4661.
Chen, X., Chou, C.-Y. & Chang, G.-G. (2009). *Antivir. Chem. Chemother.* **19**, 151–156.
Cheng, S.-C., Chang, G.-G. & Chou, C.-Y. (2010). *Biophys. J.* **98**, 1327–1336.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Chou, Y. W., Cheng, S.-C., Lai, H.-Y. & Chou, C.-Y. (2012). *Arch. Biochem. Biophys.* **520**, 74–80.
Chou, C.-Y., Chien, C.-H., Han, Y.-S., Prebanda, M. T., Hsieh, H.-P., Turk, B., Chang, G.-G. & Chen, X. (2008). *Biochem. Pharmacol.* **75**, 1601–1609.
Chou, C.-Y., Hsieh, Y.-H. & Chang, G.-G. (2011). *Methods*, **54**, 76–82.
Cleland, W. W., Frey, P. A. & Gerlt, J. A. (1998). *J. Biol. Chem.* **273**, 25529–25532.
Clementz, M. A., Chen, Z., Banach, B. S., Wang, Y., Sun, L., Ratia, K., Baez-Santos, Y. M., Wang, J., Takayama, J., Ghosh, A. K., Li, K., Mesecar, A. D. & Baker, S. C. (2010). *J. Virol.* **84**, 4619–4629.
Copeland, R. (2000). *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, 2nd ed. New York: Wiley.
Dam, J., Velikovsky, C. A., Mariuzza, R. A., Urbanke, C. & Schuck, P. (2005). *Biophys. J.* **89**, 619–634.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fouchier, R. A., Hartwig, N. G., Bestebroer, T. M., Niemeyer, B., de Jong, J. C., Simon, J. H. & Osterhaus, A. D. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 6212–6216.
Ghosh, A. K., Takayama, J., Rao, K. V., Ratia, K., Chaudhuri, R., Mulhearn, D. C., Lee, H., Nichols, D. B., Baliji, S., Baker, S. C., Johnson, M. E. & Mesecar, A. D. (2010). *J. Med. Chem.* **53**, 4968–4979.
Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. & Snijder, E. J. (2006). *Virus Res.* **117**, 17–37.
Han, Y.-S., Chang, G.-G., Juo, C.-G., Lee, H.-J., Yeh, S.-H., Hsu, J. T.-A. & Chen, X. (2005). *Biochemistry*, **44**, 10349–10359.
Hsieh, Y.-H. & Chou, C.-Y. (2011). *J. Biomed. Sci.* **18**, 4.
Hu, M., Li, P., Li, M., Li, W., Yao, T., Wu, J.-W., Gu, W., Cohen, R. E. & Shi, Y. (2002). *Cell*, **111**, 1041–1054.
Hu, M., Li, P., Song, L., Jeffrey, P. D., Chenova, T. A., Wilkinson, K. D., Cohen, R. E. & Shi, Y. (2005). *EMBO J.* **24**, 3747–3756.
Lau, S. K. P., Woo, P. C. Y., Li, K. S. M., Huang, Y., Tsoi, H.-W., Wong, B. H. L., Wong, S. S. Y., Leung, S.-Y., Chan, K.-H. & Yuen, K.-Y. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 14040–14045.

Lee, H., Cao, S., Hevener, K. E., Truong, L., Gatuz, J. L., Patel, K., Ghosh, A. K. & Johnson, M. E. (2013). *ChemMedChem*, **8**, 1361–1372.
Li, S.-W., Yang, T.-C., Wan, L., Lin, Y.-J., Tsai, F.-J., Lai, C.-C. & Lin, C.-W. (2012). *Proteomics*, **12**, 3193–3205.
Li, W. *et al.* (2005). *Science*, **310**, 676–679.
Lindner, H. A., Fotouhi-Ardakani, N., Lytvyn, V., Lachance, P., Sulea, T. & Ménard, R. (2005). *J. Virol.* **79**, 15199–15208.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nijman, S. M. B., Luna-Vargas, M. P. A., Velds, A., Brummelkamp, T. R., Dirac, A. M. G., Sixma, T. K. & Bernards, R. (2005). *Cell*, **123**, 773–786.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pyrc, K., Jebbink, M. F., Berkhout, B. & van der Hoek, L. (2004). *Virol. J.* **1**, 7.
Ratia, K., Pegan, S., Takayama, J., Sleeman, K., Coughlin, M., Baliji, S., Chaudhuri, R., Fu, W., Prabhakar, B. S., Johnson, M. E., Baker, S. C., Ghosh, A. K. & Mesecar, A. D. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 16119–16124.
Ratia, K., Saikatendu, K. S., Santarsiero, B. D., Barretto, N., Baker, S. C., Stevens, R. C. & Mesecar, A. D. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 5717–5722.
Renatus, M., Parrado, S. G., D'Arcy, A., Eidhoff, U., Gerhartz, B., Hassiepen, U., Pierrat, B., Riedl, R., Vinzenz, D., Worpenberg, S. & Kroemer, M. (2006). *Structure*, **14**, 1293–1302.
Schuck, P. (2000). *Biophys. J.* **78**, 1606–1619.
Tan, Y.-J., Lim, S. G. & Hong, W. (2005). *Antiviral Res.* **65**, 69–78.
Weiss, S. R. & Navas-Martin, S. (2005). *Microbiol. Mol. Biol. Rev.* **69**, 635–664.
Woo, P. C. Y., Lau, S. K. P., Chu, C.-M., Chan, K.-H., Tsoi, H.-W., Huang, Y., Wong, B. H. L., Poon, R. W. S., Cai, J. J., Luk, W.-K., Poon, L. L. M., Wong, S. S. Y., Guan, Y., Peiris, J. S. M. & Yuen, K.-Y. (2005). *J. Virol.* **79**, 884–895.
World Health Organization (2003). *Summary Table of SARS Cases by Country, 1 November 2002–7 August 2003.* http://www.who.int/csr/sars/country/2003_08_15/en/. Geneva: World Health Organization.
World Health Organization (2012). *Novel Coronavirus Infection in the United Kingdom.* http://www.who.int/csr/don/2012_09_23/en/. Geneva: World Health Organization.
Yonetani, T. & Theorell, H. (1964). *Arch. Biochem. Biophys.* **106**, 243–251.
Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. & Fouchier, R. A. (2012). *N. Engl. J. Med.* **367**, 1814–1820.
Zheng, D., Chen, G., Guo, B., Cheng, G. & Tang, H. (2008). *Cell Res.* **18**, 1105–1113.