Mechanism for controlling the monomer–dimer conversion of SARS coronavirus main protease

Cheng-Guo Wu,[a] Shu-Chun Cheng,[b] Shiang-Chuan Chen,[a] Juo-Yan Li,[a] Yi-Hsuan Fang,[a] Yau-Hung Chen[b] and Chi-Yuan Chou[a]*

[a]Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Taipei 112, Taiwan, and [b]Department of Chemistry, Tamkang University, Tamsui 251, Taiwan

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

The *Severe acute respiratory syndrome coronavirus* (SARS-CoV) main protease (M[pro]) cleaves two virion polyproteins (pp1a and pp1ab); this essential process represents an attractive target for the development of anti-SARS drugs. The functional unit of M[pro] is a homodimer and each subunit contains a His41/Cys145 catalytic dyad. Large amounts of biochemical and structural information are available on M[pro]; nevertheless, the mechanism by which monomeric M[pro] is converted into a dimer during maturation still remains poorly understood. Previous studies have suggested that a C-terminal residue, Arg298, interacts with Ser123 of the other monomer in the dimer, and mutation of Arg298 results in a monomeric structure with a collapsed substrate-binding pocket. Interestingly, the R298A mutant of M[pro] shows a reversible substrate-induced dimerization that is essential for catalysis. Here, the conformational change that occurs during substrate-induced dimerization is delineated by X-ray crystallography. A dimer with a mutual orientation of the monomers that differs from that of the wild-type protease is present in the asymmetric unit. The presence of a complete substrate-binding pocket and oxyanion hole in both protomers suggests that they are both catalytically active, while the two domain IIIs show minor reorganization. This structural information offers valuable insights into the molecular mechanism associated with substrate-induced dimerization and has important implications with respect to the maturation of the enzyme.

1. Introduction

Coronaviruses (CoVs) belong to the order *Nidovirales*, which are enveloped positive-stranded RNA viruses with a large genome of about 30 kb (Gorbalenya *et al.*, 2006). They include important pathogens of humans and other animals (Weiss & Navas-Martin, 2005). In late 2002, a novel CoV causing severe acute respiratory syndrome (SARS) with a 15% fatality rate emerged and spread to three continents in six months (World Health Organization; http://www.who.int/csr/sars/country/2003_08_15/en/). In the following few years, the discovery of two further species of human CoVs, NL-63 and HCoV-HKU1, as well as SARS-CoV-like strains in bats, confirmed the great diversity of CoVs (van der Hoek *et al.*, 2004; Lau *et al.*, 2005; Li *et al.*, 2005; Woo *et al.*, 2005). In September 2012, the World Health Organization was informed of several cases of acute respiratory syndrome with renal failure owing to infection with a novel CoV in the Middle East. The possible global health implications are still under critical evaluation. Nevertheless, this re-emphasizes the possibility of the future re-

emergence of SARS or a related disease. Therefore, studies to aid our understanding of these viruses and the development of novel antiviral inhibitors are both urgent and necessary.

The coronavirus nonstructural polyproteins (pp1a and pp1ab) are cleaved by two kinds of viral cysteine proteases: a main protease (M^{pro} or $3CL^{pro}$; EC 3.4.22.69) and a papain-like protease (EC 3.4.22.46; Snijder *et al.*, 2003). This process is considered to be a suitable antiviral target because cleavage is required for viral maturation (Wu *et al.*, 2006; Chou *et al.*, 2008; Zhu *et al.*, 2011). M^{pro} cleaves the polyproteins at 11 sites that contain the canonical L-Q-↓-(A/S/N) sequence (Hegyi & Ziebuhr, 2002). Catalysis by M^{pro} has been studied extensively over the years using both kinetic and mutagenesis approaches (Anand *et al.*, 2003; Chou *et al.*, 2004; Hsu, Kuo *et al.*, 2005; Lin *et al.*, 2008; Shi *et al.*, 2008; Chen *et al.*, 2008; Hu *et al.*, 2009; Cheng *et al.*, 2010). Structural information is also available on M^{pro} from SARS CoV and many other CoVs (Anand *et al.*, 2002; Yang *et al.*, 2003; Xue *et al.*, 2008; Zhao *et al.*, 2008). The

structure of the coronaviral M^{pro} consists of three domains and the catalytic dyad His/Cys is located at the interface between domains I and II. The first two domains have an antiparallel β-barrel structure that forms a folding scaffold similar to other viral chymotrypsin-like proteases (Anand *et al.*, 2002, 2003). Domain III is a five-helix fold that contributes to the dimerization of M^{pro} (Anand *et al.*, 2002, 2003). There is a very long loop (residues 176–200) between domains II and III. Recent studies have suggested that foldon unfolding of SARS-CoV M^{pro} domain III alone is able to mediate the interconversion between the monomer and a three-dimensional domain-swapped dimer under physiological conditions (Kang *et al.*, 2012); nevertheless, how the two domain IIIs meet and remain together until α-helix swapping takes place is still unknown.

Mature M^{pro} is a stable homodimer in which the two subunits are arranged perpendicularly to each other (Yang *et al.*, 2003). Mutation or deletion of the N-finger (the first seven residues) and the C-terminus (residues 298–306) can lead to a
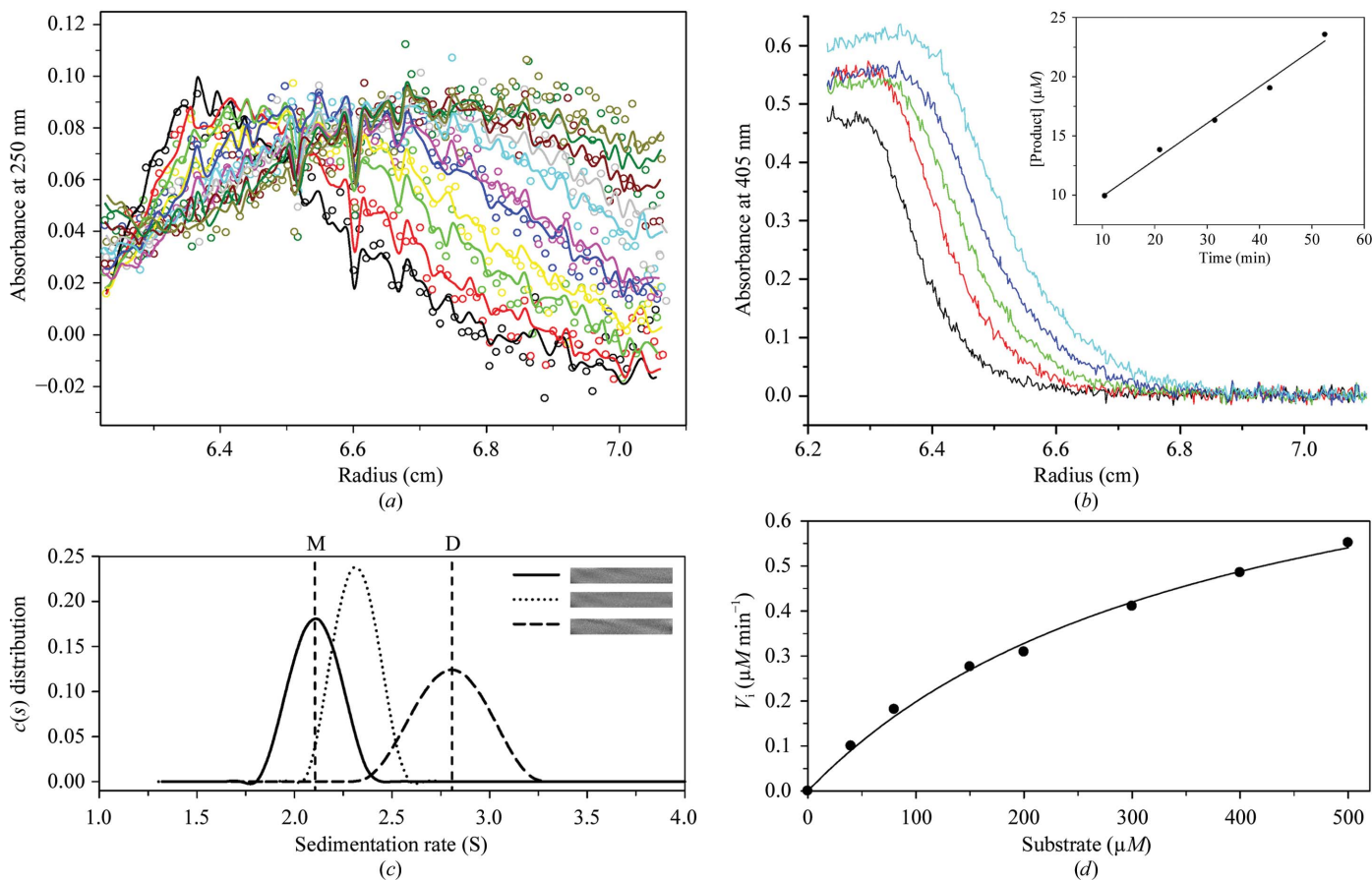

Figure 1
AEC pattern of the R298A mutant of SARS-CoV M^{pro}. The amount of protein used was 15 µl (1 mg ml^{-1}) and the total volume of the cell was 330 µl. (*a*) A typical trace of the absorbance at 250 nm of the R298A mutant during an experiment at a substrate concentration of 200 µ*M*. The symbols represent experimental data and the lines are the results fitted to the Lamm equation using the *SEDFIT* program (Chou *et al.*, 2011; Schuck, 2000). The best-fit distribution result is shown by dashed lines in (*c*). (*b*) The absorbance at 405 nm tracing the released product (pNA) after the first hour of the same experiment. The time interval between two successive spectra, from black to cyan, is 10 min. The inset plot shows the product at different times. The line indicates the best-fit result for the initial velocity calculation. (*c*) The continuous *c*(*s*) distributions of the M^{pro} R298A mutant from the best-fit analysis of the 250 nm results. The distributions in 10 m*M* phosphate buffer pH 7.6 are shown by solid lines, while those in the presence of peptide substrate at 40 and 200 µ*M* are shown as dotted and dashed lines, respectively. The two straight dashed lines indicate the positions of the monomer (M) and dimer (D). The residual bitmaps of the raw data and the best-fit results are shown as insets. (*d*) Plot of the initial velocities (from 405 nm results) *versus* substrate concentration. The line represents the best-fit results according to the Michaelis–Menten equation.

monomeric M^{pro} with little enzyme activity (Yang *et al.*, 2003; Chou *et al.*, 2004; Hsu, Chang *et al.*, 2005). M^{pro} containing additional N- and C-terminal segments of the polyprotein undergoes autoprocessing to yield the mature protease *in vitro* (Hsu, Kuo *et al.*, 2005). Inactive as a monomer, the binding of the peptide substrate or of an N-terminally and/or C-terminally elongated M^{pro} molecule is able to induce dimerization of M^{pro}, allowing catalysis (Cheng *et al.*, 2010; Chen *et al.*, 2010; Li *et al.*, 2010). The effect of substrate-induced dimerization is reversible and can be blocked by the mutation of a key residue, Glu166, which is responsible for the binding to Ser1 of the other protomer and is one of the residues recognizing Gln P1 of the substrate (Anand *et al.*, 2002; Yang *et al.*, 2003; Cheng *et al.*, 2010). In crystal structures of monomeric mutants of SARS-CoV M^{pro}, such as R298A or G11A, the oxyanion loop (Ser139–Cys145) is converted to a short 3$_{10}$-helix; this leads to complete collapse of the oxyanion hole, resulting in enzyme inactivation (Shi *et al.*, 2008; Hu *et al.*, 2009).

Despite the availability of a large amount of biochemical and structural information on M^{pro}, the mechanism by which monomeric M^{pro} is converted to a dimer during the maturation process is currently poorly understood. Here, we report the crystal structure of the SARS-CoV M^{pro} R298A mutant in the presence of peptide substrate. The structure reveals a functional dimeric form but with a minor change in the relative orientation of the two domain IIIs. Detailed exploration of this structure provides a better and more detailed understanding of the mechanisms that control the dimerization of coronaviral M^{pro}.

2. Materials and methods

2.1. Protein expression and purification

The R298A mutant of SARS-CoV M^{pro} inserted into the pET-28a(+) vector (Cheng *et al.*, 2010) was expressed in *Escherichia coli* BL21 (DE3) cells. Cultures were grown in LB medium at 310 K for 4 h and were then induced with 0.4 m*M* isopropyl β-D-1-thiogalactopyranoside; this was followed by overnight incubation at 293 K. After centrifugation at 6000*g* at 277 K for 10 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 lysed by sonication. The crude extract was then centrifuged at 12 000*g* at 277 K for 25 min to remove the insoluble pellet. The supernatant was incubated with 1 ml Ni–NTA beads (Qiagen, Hilden, Germany) at 277 K for 1 h and loaded into an empty column. After flowthrough and washing with washing buffer (20 m*M* Tris pH 8.5, 250 m*M* NaCl, 8 m*M* imidazole, 2 m*M* β-mercaptoethanol), 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 resulting protein fraction was then loaded onto an S-300 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 collected fractions was analyzed by SDS–PAGE. Fractions containing the M^{pro} protein were pooled and concentrated to

Table 1
Summary of crystallographic information for M^{pro} R298A (pH 8.0).

Values in parentheses are for the highest resolution shell.

Space group	*P*1
Unit-cell parameters (Å, °)	*a* = 55.0, *b* = 59.4, *c* = 59.8, α = 71.3, β = 73.4, γ = 72.3
Resolution (Å)	30–2.09 (2.20–2.09)
R_{merge}† (%)	3.5 (22.0)
$\langle I/\sigma(I) \rangle$	21.5 (3.2)
Completeness (%)	94.7 (82.6)
No. of reflections	35837 (4672)
Multiplicity	2.2 (2.0)
R factor‡ (%)	19.0
Free *R* factor§ (%)	24.1
R.m.s.d., bond lengths (Å)	0.009
R.m.s.d., bond angles (°)	1.3

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, 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.

30 mg ml^{-1} using an Amicon Ultra-4 30 kDa centrifugal filter (Millipore). The typical yield of protein was 5–10 mg per litre of cell culture.

2.2. Active enzyme centrifugation

Active enzyme centrifugation (AEC) can be used to observe quaternary-structural changes and catalytic activity simultaneously (Chou *et al.*, 2011). The analytical ultracentrifugation experiments were performed with an XL-A analytical ultracentrifuge (Beckman, Fullerton, California, USA) using an An-50 Ti rotor (Cheng *et al.*, 2010). A commercially available double-sector Vinogard-type bandforming centrepiece (Beckman, Fullerton, California, USA) was used for AEC (Chou *et al.*, 2011; Chou & Tong, 2011). Briefly, 15 µl of the R298A mutant of M^{pro} (1 mg ml^{-1}) was added to the small well above the sample sector. After the cell had been assembled, 330 µl peptide substrate (TSAVLQ-pNA from GL Biochem, Shanghai, People's Republic of China) dissolved in D$_2$O at 0, 40 or 200 µ*M* was loaded into the bulksample sector space. Centrifugation was then carried out at 42 000 rev min^{-1}. Absorbance at 250 nm was chosen to allow detection of the protein band, while absorbance at 405 nm was used to monitor the catalytic release of the product, pNA. The spectrum was recorded continuously using a time interval of 600 s per scan and a step size of 0.003 cm. A typical trace of the 250 nm spectral results is shown in Fig. 1(*a*). The data set was then fitted to a continuous *c*(*s*) distribution model using the *SEDFIT* program (Schuck, 2000; Brown & Schuck, 2006). The signals at 405 nm were used to calculate the initial velocities and were then fitted to the Michaelis–Menten equation, from which the kinetic parameters K_m and k_{cat} were determined (Chou *et al.*, 2011).

2.3. Protein crystallization

Crystals of the R298A mutant were obtained at 295 K by the sitting-drop vapour-diffusion method. The protein solution was set up at 15 mg ml^{-1} and included 1 m*M* TSAVLQ-pNA.

The reservoir solution consisted of 0.1 *M* Tris pH 8.0, 30%(*w*/*v*) PEG 300, 5%(*w*/*v*) PEG 1000. Large but poorly diffracting crystals appeared in 3 d and were used for micro-seeding. Single crystals of cubic shape and with dimensions of 0.2–0.3 mm were obtained in less than a week. All crystals were cryoprotected in reservoir solution with 1 m*M* TSAVLQ-pNA and were then flash-cooled in liquid nitrogen. Crystal soaking in 2 m*M* TSAVLQ-pNA failed to improve the electron density of the substrate. Soaking at higher peptide concentrations was impossible owing to low solubility.

2.4. Data collection, structure determination and refinement

X-ray diffraction data were collected at 100 K on SPXF beamline 13B1 at the National Synchrotron Radiation Research Center (Taiwan) using an ADSC Quantum-315r CCD (X-ray wavelength 0.976 Å). The diffraction images were processed and scaled using the *HKL*-2000 package (Otwinowski & Minor, 1997). The crystal belonged to space group *P*1, with unit-cell parameters *a* = 55.0, *b* = 59.4, *c* = 59.8 Å, α = 71.3, β = 73.4, γ = 72.3°. There are two M^{pro} molecules in the asymmetric unit. The structure was solved by the molecular-replacement method with the program *Phaser* (McCoy *et al.*, 2007) using the structure of wild-type M^{pro} (PDB entry 1uk4) as the search model (Yang *et al.*, 2003). Manual rebuilding of the structure model was performed with *Coot* (Emsley & Cowtan, 2004). Structure refinement was carried out using the program *REFMAC* (Murshudov *et al.*, 2011). Data-processing and refinement statistics are summarized in Table 1. The crystal structure has been deposited in the Protein Data Bank (PDB entry 4hi3).

3. Results and discussion

3.1. Substrate-induced dimerization of SARS-CoV M^{pro}

To explore the influence of substrate binding on the dimerization of M^{pro}, we performed AEC experiments on the R298A mutant of M^{pro} with or without the peptide substrate. Figs. 1(*a*) and 1(*b*) show typical absorbance traces at 250 and 405 nm of the R298A mutant during the experiment at a substrate concentration of 200 µ*M*. After fitting the 250 nm signals to the continuous

size-distribution model, it was obvious that the R298A mutant was monomeric (2.1 S) in the absence of substrate (Fig. 1*c*, solid lines), whereas it was dimeric (2.8 S) at a substrate concentration of 200 µ*M* (Fig. 1*c*, dashed lines). Somewhat surprisingly, at a substrate concentration of 40 µ*M* the major species (2.3 S) of the R298A mutant was located between the monomer and the dimer (Fig. 1*c*, dotted lines). According to previous studies, these results suggested that the R298A mutant is a rapid self-association protein, similar to wild-type M^{pro} (Cheng *et al.*, 2010). These observations using AEC
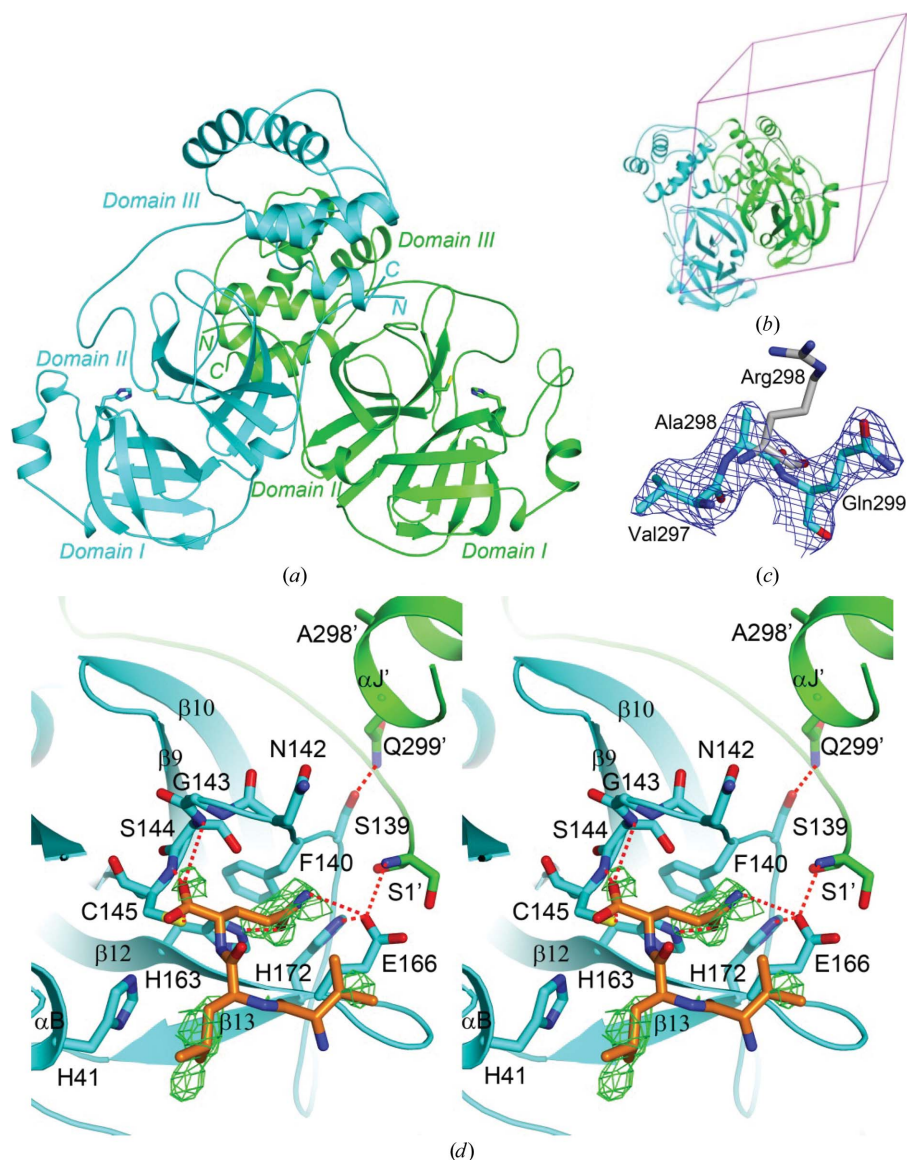

Figure 2
Dimeric structure of the R298A mutant of SARS-CoV M^{pro}. (*a*) The overall structure of the dimeric M^{pro} R298A mutant. The two protomers are coloured cyan and green, respectively. (*b*) Crystal packing of the R298A mutant. Two molecules form a biological dimer in the unit cell. (*c*) Final $2F_o - F_c$ electron density contoured at 1.0σ for residues 297–299 of the R298A mutant. The Arg298 residue in wild-type Mpro is also shown (grey). (*d*) Schematic drawing in stereoview showing the detailed interactions at the active site of the R298A mutant. The residual electron density around the active site is contoured at 2.5σ (green mesh) and the C atoms of the modelled P3–P1 (Val-Leu-Gln) residues are coloured orange. Hydrogen-bonding interactions are indicated by red dashed lines. All structure figures were produced with *PyMOL* (http://www.pymol.org).

confirmed the substrate-induced dimerization of the R298A mutant. Moreover, after calculating the initial velocities from signals at 405 nm and then fitting them to the Michaelis–Menten equation (inset in Figs. 1b and 1d), an apparent K_m of 380 μM and an apparent k_{cat} of 0.012 s^{-1} were calculated. These values are very close to those for wild-type M^{pro} from AEC analysis (Chou et $al.$, 2011). This confirmed that the dimeric R298A protein is functionally the same as that of wild-type M^{pro}. However, based on the crystal structure of the monomeric R298A mutant (Shi et $al.$, 2008), the transition

from monomer to dimer should be impossible because of the dramatic rotation of domain III and the formation of a short 3$_{10}$-helix from an active-site loop; these changes result in the catalytic machinery being frozen in a collapsed state. Furthermore, recent studies suggested that the R298E mutant maintained N-terminal autocleavage activity comparable to that of wild-type M^{pro}, although it did cause complete dimer dissociation and disruption of trans-cleavage activity (Chen et $al.$, 2010). This indicated that N-terminal autocleavage of M^{pro} is not dependent on a 'mature' dimeric protease, but on an
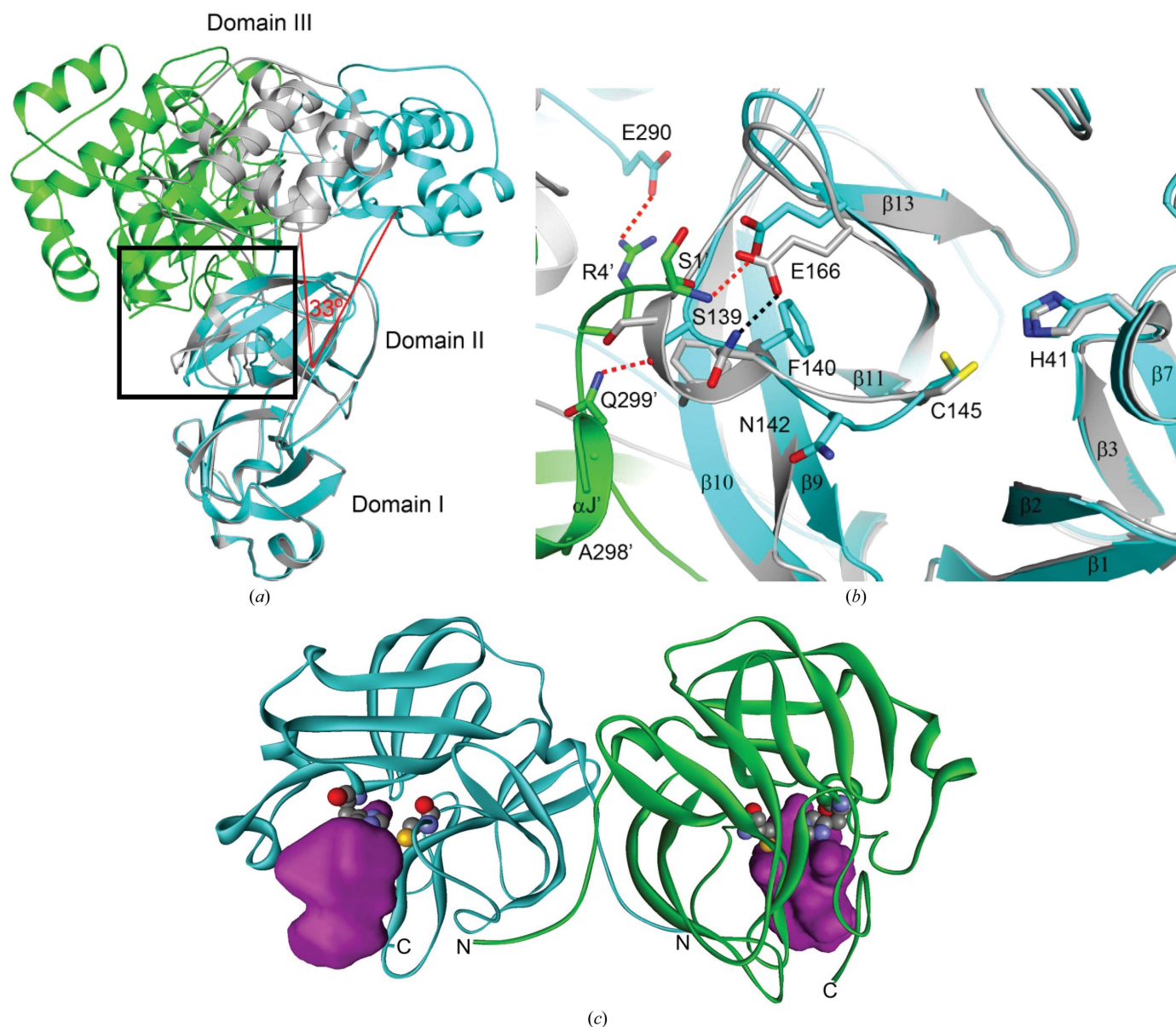

Figure 3
Comparisons with the monomeric structure of the R298A mutant of M^{pro}. (a) Overlay of the current structure of the R298A dimer (in cyan and green) with that of the R298A monomer (grey; Shi et $al.$, 2008). The orientation of domain III shows a 33° change compared with that in the monomer. The red lines represent the positions of residues 193–200 in the two structures. The region in the box is enlarged in (b). (b) Detailed interactions at the dimer interface of the R298A mutant. The red dashed lines indicate hydrogen bonds associated with the dimer, while black dashed lines indicate hydrogen bonds in the monomer. (c) Substrate-binding cavities mapped onto the current structure. Both protomers show a large and deep cavity (magenta surface representation) near the catalytic dyad, which is essentially the same as that in wild-type M^{pro} at pH 7.6, although wild-type M^{pro} at pH 6.0 shows one active protomer and one inactive protomer (Yang et $al.$, 2003). The cavity was calculated using DS $Modeling$ 1.7 (Accelrys) and was drawn using $Discovery$ $Studio$ $Visualizer$ 2.5 (Accelrys).

'intermediate' or 'loose' dimer (Chen *et al.*, 2010). On the other hand, we have expressed and purified an R298A mutant with an N-terminal extension similar to that of the R298E mutant (Chen *et al.*, 2010). However, the mutant lost its trans-cleavage activity (less than 0.8%) and failed to form a dimer even in the presence of peptide substrates (Supplementary Fig. 1[1]). Efforts to crystallize the N-terminally extended R298A mutant were also unsuccessful. To delineate the mechanism for the conversion of M^{pro} from an inactive monomer to a functional dimer, we next determined the crystal structure of the dimeric R298A mutant at a resolution of 2.09 Å (Table 1). It is important to note that we used an R298A mutant with an authentic N-terminus and a C-terminus with eight extra residues (LEH_6) for convenience in purification. According to previous studies (Chen *et al.*, 2010), in the presence of peptide substrates the structure should be more like a mature dimer with trans-cleavage activity, not a premature dimer with an N-terminal extension.

3.2. Overall structure of the R298A mutant of SARS-CoV M^{pro} in the presence of peptide substrate

The original goal of our experiment was to determine the binding modes of peptide substrates to the R298A mutant. We therefore included high concentrations of peptide substrate (1 mM; threefold higher than the K_m) in the cocrystallization conditions. The refined atomic model agreed well with the crystallographic data and the expected bond angles and bond lengths (Table 1). About 91% of the residues are in the most favoured region of the Ramachandran plot; none are in the disallowed region.

Consistent with our expectations and the AEC results, two essentially identical monomers of the R298A mutant exist in the asymmetric unit, with an r.m.s.d. of 1.1 Å between their C^α atoms (Figs. 2a and 2b). There is a minor change in orientation between the two copies of domain III and therefore the r.m.s.d. value decreases to 0.28 Å when domain III is excluded. The overall structure of the dimeric R298A mutant is similar to other structures of SARS-CoV M^{pro} reported previously. For example, the r.m.s.d. between equivalent C^α atoms of this dimeric structure and the structure of wild-type M^{pro} at pH 8.0 is 0.59 Å (Yang *et al.*, 2003). The r.m.s.d. is 0.41 Å when the structure is compared with that of wild-type M^{pro} in complex with a peptide aldehyde inhibitor (Zhu *et al.*, 2011). However, the position of domain III is different when these structures are compared (see below) and has been excluded from the comparisons.

Our crystallographic analysis of the R298A mutant shows that there is no evidence for any electron density beyond the C^β atom of this side chain (Fig. 2c). This is a confirmation that our protein carries the R298A mutation and that only mutant protein was crystallized in our experiments. Unfortunately, we cannot observe the complete electron density of the peptide substrate, even though there is some residual electron density

near the His41/Cys145 catalytic dyad of subunit A (green mesh in Fig. 2d). To evaluate the possible substrate-binding mode in this structure, we generated a model with the P3–P1 substrate residues (Val-Leu-Gln) based on the structure of the C145A mutant (Hsu, Kuo *et al.*, 2005). In the model, the side chains of P1 Gln and P2 Leu are able to fill the residual electron density and an O atom of the carboxyl group of P1 Gln is located in the oxyanion hole (Fig. 2d). In contrast, in subunit B there is no residual electron density at the same position, suggesting that the statuses of the two active sites in the dimer may not be identical, although the amide N atoms of Gly143 and Cys145 in each subunit are oriented into the oxyanion hole. In addition, most of the residues in the active site, including Phe140, His163, Glu166, His172 and the catalytic dyad His41/Cys145, are in the same positions as those of wild-type M^{pro} (Yang *et al.*, 2003). This seems to confirm that the catalytic machinery of the R298A dimer is functional.

3.3. Comparison with the monomeric structure of the R298A mutant

The mutual arrangement of the domains of the R298A mutant is changed dramatically and therefore we compared our dimer structure with that of the monomer (PDB entry 2qcy; Shi *et al.*, 2008) based on superposition of domains I and II. As shown in Fig. 3(a), the most obvious conformational change is a 33° rotation of domain III. This rotation allows steric hindrance between the two copies of domain III during dimerization to be avoided. The variability of the mutual domain arrangement suggests that conversion from the monomeric form to the dimeric form is feasible. Nevertheless, the dimerization of some other monomeric mutations, such as G11A and S139A, still needs to be confirmed, although they show a similar overall structure to the R298A monomer (Chen *et al.*, 2008; Hu *et al.*, 2009).

The structural comparison also shows that in the present structure the short 3_{10}-helix of residues Ser139-Phe140-Leu141 is disrupted and adopts a loop conformation similar to that of the wild type (Fig. 3b; see below). The unfolding of this helix enables the insertion of the N-finger of the other subunit, which is further stabilized by the interaction of Glu166 with Ser1′ (primed residue numbers indicate subunit B; Fig. 3b). Most importantly, the key stacking interaction between the rings of Phe140 and His163, which is important to prevent His163 from being protonated (Yang *et al.*, 2003), are maintained (Fig. 2d). This ensures that His163 can efficiently interact with the side chain of P1 Gln. In contrast to the collapsed substrate-binding site in the R298A monomer (Shi *et al.*, 2008), there is a large and complete substrate-binding pocket in each subunit (Fig. 3c). This further suggests that both subunits of the R298A dimer are catalytically active.

In addition, at the dimer interface of the present structure, ion pairs and hydrogen bonds, such as Arg4′–Glu290 and Ser139–Gln299′, can also be seen, which is a similar situation to that of the wild-type structure (Fig. 3b; see below). All of these observations confirm that the dimerization of R298A is very similar to that of wild-type M^{pro}, with only minor differ-

ences. Owing to the lack of a monomeric wild-type M^pro structure, the monomeric and dimeric R298A structures provide valuable insights into the dimerization process of the enzyme, which undergoes a dramatic mutual arrangement of the domains (Hsu, Kuo *et al.*, 2005; Chen *et al.*, 2008, 2010; Cheng *et al.*, 2010; Li *et al.*, 2010).
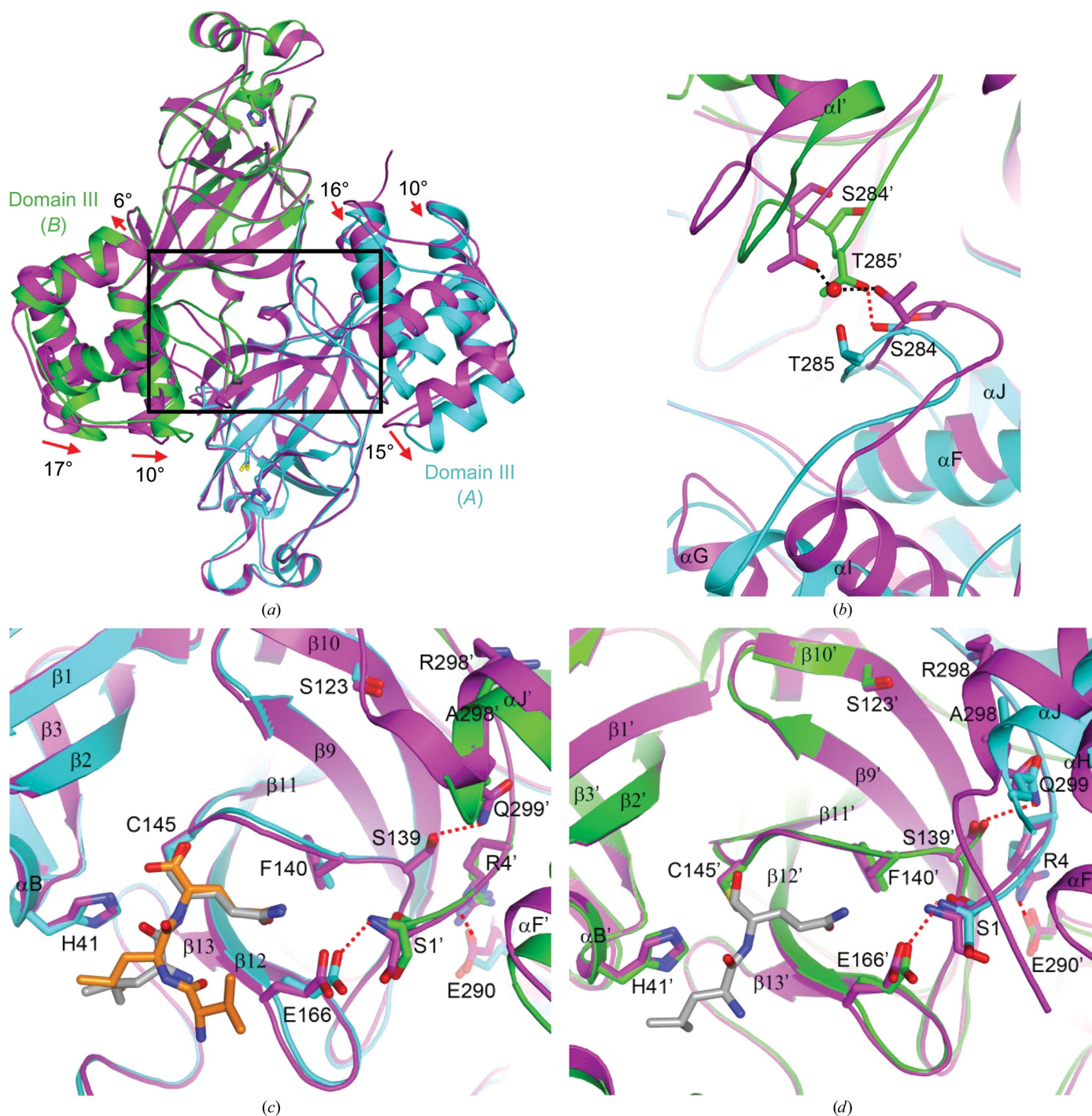

Figure 4
Comparisons with the structure of wild-type M^pro. (*a*) An overlay of the current structure of the R298A dimer (cyan and green) with that of wild-type M^pro in complex with peptidyl aldehyde inhibitor (magenta; PDB entry 3snd; Zhu *et al.*, 2011). The red arrows show the orientation change affecting the two domain IIIs. The region in the box is enlarged in (*b*). (*b*) The hydrogen-bonding interaction between the two domain IIIs. The red dashed line indicates the hydrogen bond between Ser284 and Thr285′ in the R298A mutant, while the black dashed lines show hydrogen bonds between the two Thr285 residues and the associated water (red sphere) in wild-type M^pro. (*c*) and (*d*) show the detailed interactions of the two protomers near the active sites of subunit *A* (cyan) and subunit *B* (green), respectively. The red dashed lines indicate hydrogen bonds and ion pairs for the R298A mutant, while wild-type M^pro (magenta) shows the same interactions (PDB entry 3snd; Zhu *et al.*, 2011). Overlay of the P2–P1 residues in the two structures (orange in R298A structures and grey in wild-type M^pro, respectively) confirmed that one O atom of the carboxyl group of P1 Gln was located in the oxyanion hole.

3.4. Reorganization of the dimer in the R298A mutant

The R298A mutant shows an apparent change in the relative orientation of domain III. When we compared our structure with two wild-type structures (Yang *et al.*, 2003; Zhu *et al.*, 2011), namely the free enzyme at pH 8.0 (PDB entry 1uk2) and the peptide aldehyde inhibitor complex (PDB entry 3snd), a relative shift of 5–24° could be observed for the two copies of domain III in the dimer compared with the free enzyme, while there was a shift of 6–17° compared with the inhibitor complex (Fig. 4*a*). Interestingly, we found that there is a hydrogen-bonding interaction between residues Ser284 and Thr285′ in the αI–αJ loop, while there is a water molecule between the two Thr285 residues in the structure of wild-type M^pro (Fig. 4*b*). The direct contact between the two αI–αJ loops leads to a shift of the αI and αJ helices, further causing the change in the orientation of the whole domain III. This observation also suggests that the two copies of the folded domain III are able to bind to each other by this interaction at the initial stage and to wait until the swapping of the α-helix to form a three-dimensional domain-swapped dimer or a more stable and super-active octameric M^pro (Zhang *et al.*, 2010; Kang *et al.*, 2012).

Furthermore, in the present structure each domain III is even closer to domains I and II of the other subunit, although subunit *A* shows a larger change in orientation than subunit *B*. At the dimer interface of the R298A mutant, in addition to two intermolecular Arg4–Glu290 ion pairs (Arg4–Glu290′ and Arg4′–Glu290), there are two Ser139···Gln299 hydrogen bonds (Ser139···Gln299′ and Ser139′···Gln299) and two Ser1–Glu166 ion pairs (Ser1–Glu166′ and Ser1′–Glu166); this contrasts with some of the structures of wild-type M^pro (PDB entries 1uk2 and 1uk4), which only show one pair of each. A similar assembly can also be observed in the structures of authentic wild-type M^pro (PDB entry 2h2z) and its complex with a peptide aldehyde inhibitor (PDB entry 3snd), although in these two structures the Arg298 residues do not interact with Ser123 of the other subunit (Figs. 4*c* and 4*d*; Xue *et al.*, 2007; Zhu *et al.*, 2011). This suggests that M^pro may not require all of the possible intermolecular interactions for dimerization, especially in the presence of substrates or peptidyl inhibitors. Moreover, previous studies have suggested that Glu166 plays a pivotal role in connecting the substrate-binding site to the dimer interface (Yang *et al.*, 2003; Chen *et al.*, 2008; Cheng *et al.*, 2010). In our structure, the interactions between the main-chain amide of Ser1 with the carboxyl group of the Glu166 side chain of the other subunit provide direct evidence to explain why the mutation at Glu166 blocks substrate-induced dimerization of M^pro (Cheng *et al.*, 2010). Remarkably, mutation of Arg298, which should be detrimental to dimerization, may be compensated by these interactions, while most of the other residues in the active site show only small changes.

4. Conclusion

Our studies show that SARS-CoV M^pro, the dimerization of which is important for its catalytic activity, is able to tolerate large orientation changes, especially involving domain III. Mutation of Arg298, when introduced at the dimer interface, disturbs the dimerization; nevertheless, in the presence of peptide substrate the dimerization is able to be induced or rescued by intermolecular hydrogen-bond (Ser139···Gln299) and ion-pair (Ser1–Glu166) interactions. Based on the existence of a complete substrate-binding pocket and a complete oxyanion hole, we suggested that the dimer is still catalytically active, even though there are conformational rearrangements of the two copies of domain III in the dimer. AEC experiments confirmed that the kinetic parameters of the R298A mutant are similar to those of wild-type M^pro. The present studies provide valuable insights into the mechanisms that control the monomer–dimer switch during the maturation process of M^pro.

References

Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J. & Hilgenfeld, R. (2002). *EMBO J.* **21**, 3213–3224.

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R. (2003). *Science*, **300**, 1763–1767.

Brown, P. H. & Schuck, P. (2006). *Biophys. J.* **90**, 4651–4661.

Chen, S., Hu, T., Zhang, J., Chen, J., Chen, K., Ding, J., Jiang, H. & Shen, X. (2008). *J. Biol. Chem.* **283**, 554–564.

Chen, S., Jonas, F., Shen, C. & Hilgenfeld, R. (2010). *Protein Cell*, **1**, 59–74.

Cheng, S.-C., Chang, G.-G. & Chou, C.-Y. (2010). *Biophys. J.* **98**, 1327–1336.

Chou, C.-Y., Chang, H.-C., Hsu, W.-C., Lin, T.-Z., Lin, C.-H. & Chang, G.-G. (2004). *Biochemistry*, **43**, 14958–14970.

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.

Chou, C.-Y. & Tong, L. (2011). *J. Biol. Chem.* **286**, 24417–24425.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. & Snijder, E. J. (2006). *Virus Res.* **117**, 17–37.

Hegyi, A. & Ziebuhr, J. (2002). *J. Gen. Virol.* **83**, 595–599.

Hoek, L. van der, Pyrc, K., Jebbink, M. F., Vermeulen-Oost, W., Berkhout, R. J., Wolthers, K. C., Wertheim-van Dillen, P. M., Kaandorp, J., Spaargaren, J. & Berkhout, B. (2004). *Nature Med.* **10**, 368–373.

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.-J. & Liang, P.-H. (2005). *J. Biol. Chem.* **280**, 31257–31266.

Hsu, W.-C., Chang, H.-C., Chou, C.-Y., Tsai, P.-J., Lin, P.-I. & Chang, G.-G. (2005). *J. Biol. Chem.* **280**, 22741–22748.

Hu, T., Zhang, Y., Li, L., Wang, K., Chen, S., Chen, J., Ding, J., Jiang, H. & Shen, X. (2009). *Virology*, **388**, 324–334.

Kang, X., Zhong, N., Zou, P., Zhang, S., Jin, C. & Xia, B. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 14900–14905.

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.

Li, C., Qi, Y., Teng, X., Yang, Z., Wei, P., Zhang, C., Tan, L., Zhou, L., Liu, Y. & Lai, L. (2010). *J. Biol. Chem.* **285**, 28134–28140.

Li, W. *et al.* (2005). *Science*, **310**, 676–679.

Lin, P.-Y., Chou, C.-Y., Chang, H.-C., Hsu, W.-C. & Chang, G.-G. (2008). *Arch. Biochem. Biophys.* **472**, 34–42.

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.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Schuck, P. (2000). *Biophys. J.* **78**, 1606–1619.

Shi, J., Sivaraman, J. & Song, J. (2008). *J. Virol.* **82**, 4620–4629.

Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L. M., Guan, Y., Rozanov, M., Spaan, W. J. M. & Gorbalenya, A. E. (2003). *J. Mol. Biol.* **331**, 991–1004.

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.

Wu, C.-Y., King, K.-Y., Kuo, C.-J., Fang, J.-M., Wu, Y.-T., Ho, M.-Y., Liao, C.-L., Shie, J.-J., Liang, P.-H. & Wong, C.-H. (2006). *Chem. Biol.* **13**, 261–268.

Xue, X., Yang, H., Shen, W., Zhao, Q., Li, J., Yang, K., Chen, C., Jin, Y., Bartlam, M. & Rao, Z. (2007). *J. Mol. Biol.* **366**, 965–975.

Xue, X., Yu, H., Yang, H., Xue, F., Wu, Z., Shen, W., Li, J., Zhou, Z., Ding, Y., Zhao, Q., Zhang, X. C., Liao, M., Bartlam, M. & Rao, Z. (2008). *J. Virol.* **82**, 2515–2527.

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. & Rao, Z. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 13190–13195.

Zhang, S., Zhong, N., Xue, F., Kang, X., Ren, X., Chen, J., Jin, C., Lou, Z. & Xia, B. (2010). *Protein Cell*, **1**, 371–383.

Zhao, Q., Li, S., Xue, F., Zou, Y., Chen, C., Bartlam, M. & Rao, Z. (2008). *J. Virol.* **82**, 8647–8655.

Zhu, L., George, S., Schmidt, M. F., Al-Gharabli, S. I., Rademann, J. & Hilgenfeld, R. (2011). *Antiviral Res.* **92**, 204–212.