J. Biochem. 2010;148(3):349–358 doi:10.1093/jb/mvq071

JB THE JOURNAL OF BIOCHEMISTRY

# Essential covalent linkage between the chymotrypsin-like domain and the extra domain of the SARS-CoV main protease

*Received May 24, 2010; accepted June 24, 2010; published online June 29, 2010*

**Meng-Ying Tsai[1], Wei-Hsin Chang[1], Jin-Yi Liang[2], Long-Liu Lin[2], Gu-Gang Chang[1] and Hui-Ping Chang[2,*]**

[1]Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, 155 Li-Nong St., Section 2, Taipei 112; and [2]Department of Applied Chemistry, National Chia-Yi University, 300 Syuefu Road, Chiayi 600, Taiwan

*Hui-Ping Chang, Department of Applied Chemistry, National Chia-Yi University, 300 Syuefu Road, Chiayi 600, Taiwan. Tel: +886 5 271 7899, Fax: +886 5 271 7901, email: huiping_chp@hotmail.com

**The main protease of the coronavirus causing severe acute respiratory syndrome performs proteolytic processing of the viral polyproteins. The active form of the enzyme is a homodimer with each subunit consisting of three structural domains. Domains I and II, hosting the complete catalytic machinery, constitute the N-terminal chymotrypsin-like folding scaffold and connect to the extra C-terminal domain III by a long loop. Previously, the domain III-truncated enzyme was demonstrated to fold independently into an intact chymotrypsin-like fold, but it showed no enzyme activity. To further delineate the structure–function relationships of the domain III and the long loop, we generated some truncated and mutated $M^{pro}$ forms bearing various combinations of the loop with other structural parts of the enzyme. Their conformational and association properties were investigated in detail. Far-ultraviolet circular dichroism (CD) measurements revealed that these fragments could fold independently. The secondary, tertiary and quaternary structures of these mixtures were monitored by CD, fluorescence spectroscopy and analytical ultracentrifugation. However, no enzyme activity was observed for any mutant or mixtures. These observations indicate that the covalent linkage between the chymotrypsin like and the extra domain is essential for enzymatic activity of the main coronavirus protease and for the integrity of its quaternary structure.**

*Keywords*: SARS-CoV/$M^{pro}$/folding/protein stability/nicked protein.

*Abbreviations*: SARS, severe acute respiratory syndrome; CoV, coronavirus; $M^{pro}$, main protease; $M^{pro}_{(I+II)+loop}$, a C-terminal domain-truncated $M^{pro}$ with loop retained; $M^{pro}_{(I+II)}$, domain (I+II) of $M^{pro}$; $M^{pro}_{loop+III}$, a N-terminal domain-truncated $M^{pro}$ with loop retained; $M^{pro}_{(III)}$, domain III of $M^{pro}$; $M^{proW31}_{W207F/W218F}$, a double mutant of $M^{pro}$ with the tryptophanyl residues at positions 207 and 218 mutated to phenylalanine, the only unaltered tryptophanyl residue was W31[a]; $M^{proW207/W218}_{W31F}$, a point mutant of $M^{pro}$ with tryptophanyl residue at position 31 mutated to phenylalanine; $M^{pro}_{\Delta(297-306)}$, a deletion mutant of $M^{pro}$ with C-terminal amino acid residus 297–306 deleted; AEW, average emission wavelength; CD, circular dichroism; AUC, analytical ultracentrifugation; GdmCl, guanidinium chloride.

[a]For consistency we have adopted a systematic abbreviation nomenclature for all engineered enzymes: superscript font for unaltered residues and subscript for the engineered portion.

Coronaviruses (CoVs) are extremely common in animals but usually not pathogens for humans. The severe acute respiratory syndrome (SARS) CoV is an exception in this respect, which caused a board range of diseases and raised a worldwide health hazard in 2003. Maturation of the SARS-CoV involves highly complicated events of proteolytic processing of the polyproteins by the main protease ($M^{pro}$) and the papain-like protease (*1–5*). All $M^{pro}$ forms, including that from SARS-CoV, have a common homologous structural feature, irrespective of their origin (*1*). The development of wide-spectrum antiviral drugs against $M^{pro}$ is thus an attractive therapeutic aim.

SARS-CoV $M^{pro}$ is a homodimer in the crystalline form and a monomer-dimer equilibrium in solution (*2*, *6–7*). Available evidence indicates that the dimeric form is the enzymatically active form and that the monomer is inactive (*8*). Each subunit of the dimeric $M^{pro}$ consists of three structural domains [Fig. 1A, generated with Pymol (*9*)]. The N-terminal Domains I (residues 8–99) and II (residues 100–183) have an anti-parallel β-barrel structure with a chymotrypsin-like folding scaffold. The C-terminal domain III (residues 200–306) has five α-helices arranged into a globular cluster that is connected to Domain II by a long loop (residues 184–199). The SARS-CoV $M^{pro}$ has a catalytic Cys–His dyad, instead of the Ser–His–Asp triad of chymotrypsin. The active centre is located in the interface between Domains I and II. The functional role of domain III but not the loop that links domain III and domain (I + II) has been investigated intensively (*8*, *10*, *11*).

We have developed a domain III-truncated mutant to study the functional role of this extra C-terminal domain in SARS-CoV $M^{pro}$ (*10*). In chemically induced denaturation experiments, the free energy change of the C-terminal domain-truncated mutant

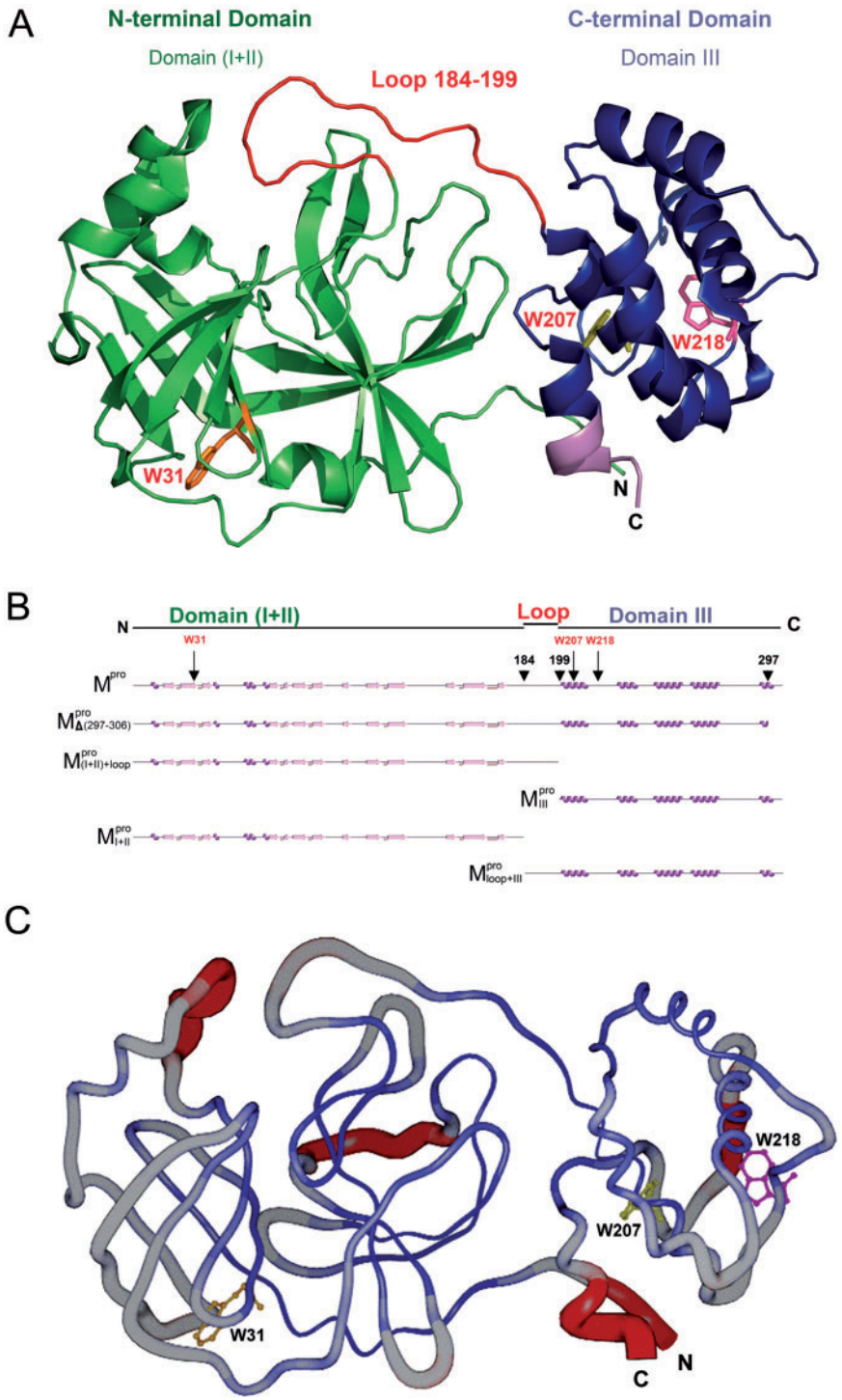


**Fig. 1 Structural features of the SARS-CoV $M^{pro}$.** (A) The three structural domains of SARS-CoV $M^{pro}$ (pdb code: 1uk3) are shown in ribbon representation. The N-terminal domain (residues 1–183) is shown in green, C-terminal domain (residues 200–306) in blue, the loop (residues 184–199) connecting the two domains in red, the deletion region (residues 297–306) of $M^{pro}_{\Delta(297-306)}$ in violet. The tryptophanyl residues $Trp^{31}$, $Trp^{207}$ and $Trp^{218}$ are shown in orange, yellow and pink, respectively, as bond model. Generate with Pymol. (B) Secondary structure features of the enzyme and the truncated regions of the various truncation mutants. (C) The structural flexibility of the enzyme is shown as a variable worm of the polypeptide chain based on the B factor of the pdb file (pdb code: 1uk3). The rigid structure is shown as a thinner worm in blue. The more flexible peripheral regions are shown in red. Only one of the subunits is shown. This figure was generated with SPOCK (http://quorum.tamu.edu).

was higher than the wild-type (WT) $M^{pro}$, which indicated that domain III might contribute to stabilization of the domain (I + II) (*10*). To further delineate the structure and function relationship of the domain III and the long loop, we generated more $M^{pro}$ mutants bearing various combinations of the loop connecting to other structural parts of the enzyme (Fig. 1B). The secondary, tertiary and quaternary structural

features were monitored by circular dichroism (CD), fluorescence spectrometry and analytical ultracentrifugation (AUC). Our results indicate that the covalent linkage between the catalytic N-terminal domain and the extra C-terminal domain is essential for both enzyme activity and for the integrity of the quaternary structure of SARS-CoV $M^{pro}$.

## Materials and Methods

### *Construction, expression and purification of various SARS-CoV $M^{pro}$*

The genes of the full-length, truncated, $M^{pro}_{(I+II)+loop}$ and deletion mutant, $M^{pro}_{\Delta(297-306)}$, of SARS-CoV $M^{pro}$ inserted into the vector pET29(+) (Novagen, Madison, WI, USA) had been described previously (*8*, *10*). The genes of the N-terminal truncated ($M^{pro}_{III}$, $M^{pro}_{loop+III}$) and C-terminal truncated ($M^{pro}_{(I+II)}$) mutants of SARS-CoV $M^{pro}$ were made by polymerase chain reaction (PCR) with appropriate primers using the full-length $M^{pro}$ as the template. The forward primers for $M^{pro}_{III}$, $M^{pro}_{loop+III}$ and $M^{pro}_{(I+II)}$ were 5′-GGTGGT**CATATG**ATAACAT TAAAT-3′, 5′-GGTGGT**CATATG**CCATTTGTTGAC-3′ and 5′-GGG**CCATGG**ATAGTGGTTTTAGGAAAATG-3′, respectively; and the reverse primers were 5′-AA**CTCGAG**GGTAACACC AGAG-3′, 5′-AA**CTCGAG**GGTAACACCAGAG-3′ and 5′-GGG**CTCGAG**ACCATAGAATTTACCTTC-3′, respectively. The *NdeI*, *NcoI* and *XhoI* sites are shown underlined and in bold. After digestion with *Nde*I and *Xho*I, each PCR product was ligated to the *Nde*I and *Xho*I sites of the pET-SARS-CoV $M^{pro}$ except$M^{pro}_{(I+II)}$, which was inserted into the *Nco*I and *Xho*I sites of the pET-28(a) (Novagen, Madison, WI, USA) after digestion with *Nco*I and *Xho*I.

Mutagenesis of full-length SARS-CoV proteases was performed using the QuikChange mutagenesis kit (Stratagene, La Jolla, CA, USA). For the tryptophan mutants, the forward primers for $M^{proW207/W218}_{W31F}$, $M^{proW31/W218}_{W207F}$ and $M^{proW31/W207}_{W218F}$ were 5′-CTCTTAA TGGATTG**TTC**TTGGATGACACAGT-3′, 5′-CCATAACA TTA AATGTTTTGGCA**TTC**CTGTATGCTGC-3′ and 5′-CAATGGT GATAGG**TTC**TT TCTTAATAGATTCACC-3′, respectively. The reverse primers were 5′-ACTGTGTC ATCCAA**GAA**CAATCCAT TAAGAG-3′, 5′-GCAGCATACAG**GAA**TGCCAAAACA TTTAA TGTTATGG-3′ and 5′-GGTGAATCTATTAAGAAA**GAA**CCTA TCACCATTG-3′, respectively. The primers as described above were used to mutate the $Trp^{31}$, $Trp^{207}$ and $Trp^{218}$ codons to phenylalanine codon by PCR. The mutational sites are shown underlined and in bold. The pET-$M^{pro}$ expression vector was used as template to construct pET-$M^{proW207/W218}_{W31F}$, pET-$M^{proW31/W207}_{W218F}$. For pET-$M^{proW31}_{(W207F/W218F)}$ double mutants, pET-$M^{pro(W31/W207)}_{W218F}$ was used as the template.

All nucleotide sequences were confirmed by autosequencing analysis. All genes coding for the enzymes were successfully expressed and the corresponding enzymes purified to apparent homogeneity by protocols described previously (*8*).

### *CD analysis*

Far-UV CD spectra were measured at 30°C with a Jasco J-815 spectropolarimeter under constant $N_2$ flush. The temperature was controlled by air-controlling Peltier device (PTC-423S/15, JASCO, Tokyo, Japan). The protein solution was scanned from 250 to 190 nm in cuvettes with a 0.1 mm path length at 0.1-nm intervals with a bandwidth of 1.0 nm. The photomultiplier absorbance was always below 600 V in the analysed region. Each scanning was repeated 10× and an average was reported. Data were corrected for the absorbance of buffer, and the results were expressed as molar ellipticity [Θ] in the units of degrees $cm^2$ $decimol^{-1}$.

Molar ellipticity [Θ] was calculated according to equation 1:

$$[\Theta] = \frac{\Theta}{10 \cdot C \cdot l} \quad (1)$$

Where $l$ represents the light path length (cm); $C$, the molar concentration of protein (mol/l) and Θ, the observed ellipticity (mdeg).

The secondary structural contents were estimated by the online DICHROWEB server (web site, http://www.cryst.bbk.ac.uk/cdweb/html/home.html) (*12*, *13*) using the implemented CDSSTR programme (*14–16*) (reference set 7). Spectral analysis was taken over the wavelength range from 190 to 240 nm.

Thermal denaturation experiments were performed by monitoring the ellipticity at 222 nm. The temperature was increased with a heating rate of 1°C/min from 30 to 90°C and the temperature at which half of the protein molecules were unfolded was recorded ($T_m$).

### *Spectrofluorimetric analysis*

Fluorescence spectra of the enzyme were monitored in a Perkin-Elmer LS-50B luminescence spectrometer at 30°C. All spectra were corrected for buffer absorption. The excitation wavelength was set at 295 nm and the fluorescence emission spectrum was scanned from 300 to 400 nm. When excited with 295 nm UV light, the protein intrinsic emission fluorescence was originated from tryptophanyl residue only. The maximal peak of the fluorescence spectrum and the change in fluorescence intensity were used in monitoring the unfolding processes of the enzyme. Both the red shift and the changes in fluorescence intensity were analysed together using the average emission wavelength (AEW) (<λ>) according to equation 2 (*17*):

$$\langle\lambda\rangle = \frac{\sum_{i=\lambda_1}^{\lambda_N} (F_i \cdot \lambda_i)}{\sum_{i=\lambda_1}^{\lambda_N} F_i} \quad (2)$$

in which $F_i$ is the fluorescence intensity at the specific emission wavelength ($\lambda_i$).

### *Enzyme activity assay of SARS-CoV $M^{pro}$ using fluorogenic substrate*

The enzyme activity measurements were performed in 30 mM Tris–HCl buffer (pH 7.7) at 30°C. Enhanced fluorescence due to cleavage of the internal quenched fluorogenic substrate peptide (*ortho*-aminobenzoic acid-TSAVLQSGFRK-2,4-dinitrophenyl amide) was monitored at 420 nm with excitation at 362 nm using a Perkin-Elmer LS-50B luminescence spectrometer.

### *AUC*

The molar mass and sedimentation coefficient of the proteases were analysed by a sedimentation velocity experiment. Sedimentation velocity was performed in a Beckman-Coulter XL-A analytical ultracentrifuge (Fullerton, CA, USA) with an An50Ti rotor at 20°C and 42,000 rpm in 12-mm double-sector *Epon* charcoal-filled centerpieces. Prior to the experiments, the sample was diluted to various protein concentrations with 30 mM Tris–HCl. The UV absorption of the cells at 280 nm was scanned in a continuous mode with time interval of 8 min and a step size of 0.003 cm. The partial specific volume of the enzyme, solvent density and viscosity were calculated by the software SEDNTERP (http://www.jphilo.mailway.com/). All samples were visually checked for clarity after ultracentrifugation to make sure that there was no indication of precipitation. The recorded scans at different time points were collected and fitted to a continuous size distribution model by the SEDFIT programme (*18*, *19*). The observed sedimentation profiles of a continuous size distribution [$c(s)$] can be calculated from equation 3.

$$a(r,t) = \int c(s)L(s,D,r,t)\mathrm{d}s + \varepsilon \quad (3)$$

where $a(r,t)$ denotes the experimentally observed signal, $L(s,D,r,t)$ denotes the solution of the Lamm equation for a single species, and $\varepsilon$ is the noise component.

### *Reversible unfolding/refolding of the enzyme in GdmCl*

The folding experiments of all the full-length WT, truncated and mutated SARS-CoV $M^{pro}$ were examined in different concentrations of GdmCl in Tris–HCl buffer (30 mM, pH 7.7) at 30°C.

The unfolding of the enzyme was monitored by protein fluorescence change. The unfolding was found to reach equilibrium in transient. A 10-min incubation, both unfolding and refolding, was used in the experiments in the present study.

#### *Protein unfolding data analysis*

Since unfolding/refolding of our sample is a reversible process, the unfolding data were treated with the following thermodynamic models, the two-state (Scheme I) and three-state unfolding model (Scheme II), by global fitting of the data to equations described previously (*10*, *20*):

$$N \overset{K_{N-U}}{\longleftrightarrow} U$$

(Scheme I)

$$N \overset{K_{N-I}}{\longleftrightarrow} I \overset{K_{I-U}}{\longleftrightarrow} U$$

(Scheme II)

## Results

### *Engineered variants of SARS-CoV M$^{pro}$ for probing linker function*

We used molecular biology techniques to create SARS-CoV $M^{pro}$ variants, which allowed us to study the structural role of the linker between domain (I+II) and domain III in $M^{pro}$. We created four linker variants: $M^{pro}_{III}$, $M^{pro}_{loop+III}$, $M^{pro}_{(I+II)}$ and $M^{pro}_{(I+II)+loop}$. Among them, $M^{pro}_{loop+III}$ and $M^{pro}_{(I+II)+loop}$ have the loop covalently linked to other parts of the enzyme, but $M^{pro}_{III}$ and $M^{pro}_{(I+II)}$ lack the linker. The yields of the purified protein were usually in the range of 5–15 mg/ml from 200 ml of cell culture medium. The purity of the enzymes was determined to be >97% by SDS-PAGE (Supplementary Fig. S1).

AUC analysis verified the expected decreases in molar mass of $M^{pro}_{III}$ (13.2 kDa), $M^{pro}_{loop+III}$ (16.98 kDa), $M^{pro}_{(I+II)}$ (22.13 kDa) and $M^{pro}_{(I+II)+loop}$ (23.23 kDa) versus 35 kDa for the monomeric mutant, $M^{pro}_{\Delta(297-306)}$ and 64.26 kDa for the dimeric full-length WT $M^{pro}$ and mutants. All these experimentally determined molar mass values were in agreement with the theoretical values calculated from amino acid sequences.

### *Essential role of the covalent linker between domain (I+II) and domain III*

We prepared three kinds of solutions comprising the entire enzyme (nicked $M^{pro}$): $M^{pro}_{(I+II)+loop}/M^{pro}_{III}$ (Mixture 1), $M^{pro}_{(I+II)}/M^{pro}_{loop+III}$ (Mixture 2) and $M^{pro}_{(I+II)}/M^{pro}_{III}$ (Mixture 3), all at a molar ratio of 1:1 (146 μM). The conformational feature of this nicked $M^{pro}$ and its constituent fragments were evaluated by CD spectroscopy as shown in Fig. 2. Results indicated that the enzymes have a well-defined secondary structure, except for that of $M^{pro}_{(I+II)}$, which had a low CD signal, as anticipated. The secondary structure of the enzymes was estimated using the DICHROWEB server (*12*, *13*). To estimate the secondary structure of $M^{pro}_{(I+II)}$, a 1.5-fold excess of enzyme was employed. The CDSSTR analysis is shown in Table I. The normalized root mean square deviation (NRMSD) values of the data fitting for WT, mutants and various truncations were all <0.05 showing excellent goodness-of-fit. The calculated helical and β-strand contents of the recombinant WT were 0.23 and 0.24, respectively, consistent with the data derived from the crystal structure (Table I).

The helical content of the monomeric $M^{pro}_{\Delta(297-306)}$ was 0.17, a value significantly different from the 0.26 of the crystal structure of a monomeric $M^{pro}$ (PDB code: 2QCY) (*21*). However, this is predictable because a seventh helix was deleted in $M^{pro}_{\Delta(297-306)}$ (Fig. 1B).

Both $M^{pro}_{III}$ and $M^{pro}_{loop+III}$ displayed far-ultraviolet (Far-UV) CD spectra with the features of α-helical polypeptides, with minimal values of negative ellipticity at around 208 and 222 nm. In contrast, $M^{pro}_{(I+II)}$ and $M^{pro}_{(I+II)+loop}$ were characteristic for low α-helix and high β-strand contents, as expected. In aqueous solution and at neutral pH, the $M^{pro}_{III}$ and $M^{pro}_{loop+III}$ adopted ~48 and 71% helicity, respectively. On the other hand, $M^{pro}_{(I+II)}$ and $M^{pro}_{(I+II)+loop}$ assumed mainly β-strand structures of 33 and 37%, respectively. These results suggest that the overall structure was not destroyed by truncations. This result was also observed in Shi's study, they found that each fragment (I + II and III) can fold independently (*3*). As shown in Table I, the secondary structural characteristics of Mixtures 1 or 3 are more similar to $M^{pro}_{\Delta(297-306)}$ than to the WT. In contrast, a considerable difference in the far-UV CD spectra was observed in Mixture 2 (Fig. 2). These CD data indicate that there are secondary structure differences between the nicked and intact proteins.

The thermal stabilities of the recombinant SARS-CoV $M^{pro}$ and the mixtures were also examined by monitoring the continuous changes in the ellipticity at 222 nm during the temperature range from 30°C to 90°C (Table I). The enzyme precipitated at high temperature and the thermal denaturation was irreversible. The melting profiles are shown in Supplementary Fig. 2S. While two distinct transitions were observed around 49.2 and 59.2°C in the melting profile of the Mixture 1; 49.5 and 62.8°C in the Mixture 2, 46.5 and 58.6°C in the Mixture 3, only one large transition was observed around 53.9°C in the profile of WT. Nevertheless, it can be concluded that the transition proceeds in a three-state fashion with indication of independent behaviour by either of the domains in the mixtures. Among the truncated SARS-CoV $M^{pro}$, the C terminally truncated protein had a significantly lower $T_m$ (50.2 for $M^{pro}_{(I+II)}$ and 52.2 for $M^{pro}_{(I+II)+loop}$) than did the N terminally truncated one (57.7°C and 79.2°C for $M^{pro}_{III}$ and 61.2°C for $M^{pro}_{loop+III}$). The observation is an indication that the catalytic domain is intrinsically less stable and that it derives a significant fraction of its stabilization energy from its interactions with domain III through the linker. The $T_m$ value for WT was 53.9°C (~54), which is a little higher than

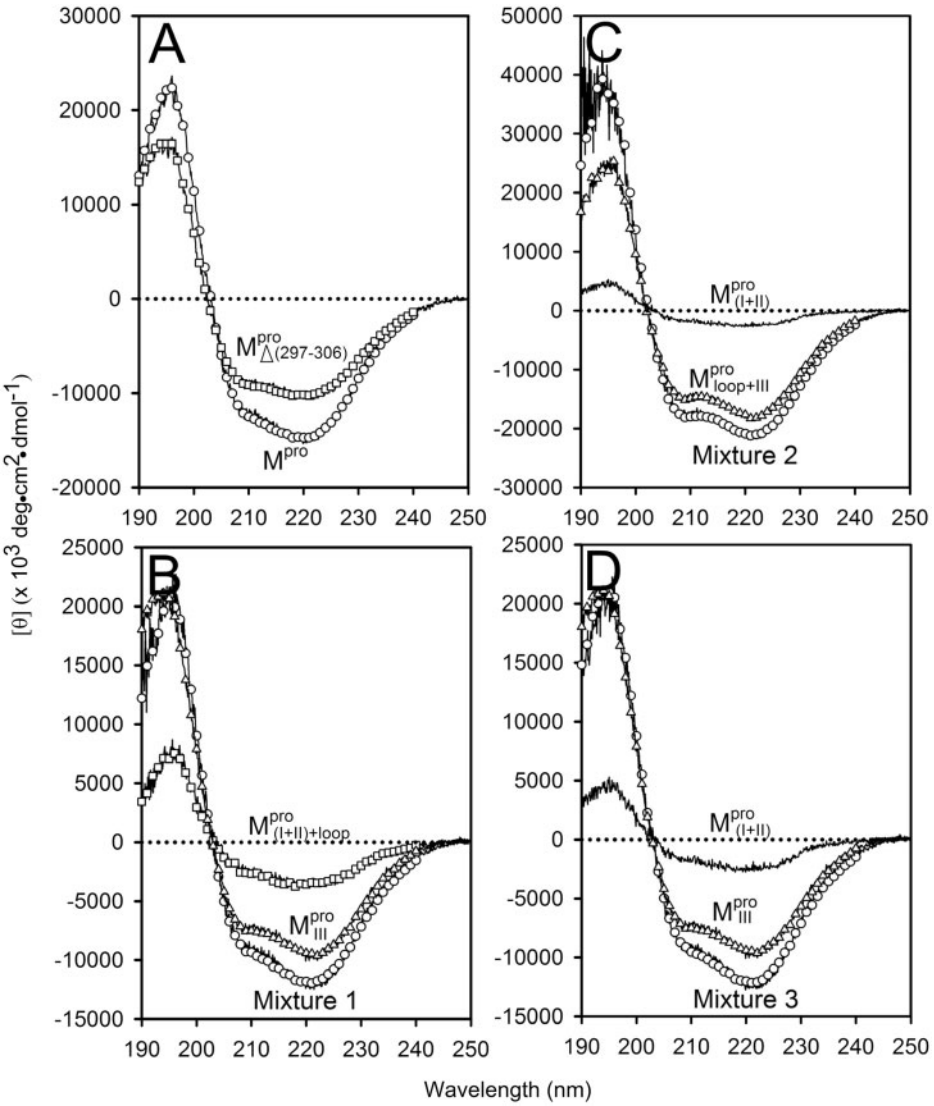


**Fig. 2 Far-UV CD spectra of $M^{pro}$, mixtures and each constituting fragments.** Spectra were recorded at 30°C in 30 mM Tris–HCl buffer (pH 7.7) at protein or fragment concentration of 146 μM except for WT (95.6 μM) and $M^{pro}_{\Delta(297-306)}$ (106 μM). (A) $M^{pro}_{\Delta(297-306)}$ and $M^{pro}$; (B) $M^{pro}_{(I+II)+loop}$, $M^{pro}_{III}$ and Mixture 1 ($M^{pro}_{(I+II)+loop}/M^{pro}_{III}$); (C) $M^{pro}_{(I+II)}$, $M^{pro}_{loop+III}$ and Mixture 2 ($M^{pro}_{(I+II)}/M^{pro}_{loop+III}$); (D) $M^{pro}_{(I+II)}$, $M^{pro}_{III}$ and Mixture 3 ($M^{pro}_{(I+II)}/M^{pro}_{III}$).

that of $M^{pro}_{(I+II)}$ (50.2°C) and $M^{pro}_{(I+II)+loop}$ (52.2°C). These results are consistent with the previous report that the C-terminal domain is involved in stabilizing the N-terminal catalytic domain of SARS-CoV $M^{pro}$ (*10*). Two transitions with the midpoint temperatures at 57.7 and 79.2°C upon thermal unfolding were also demonstrated in $M^{pro}_{III}$. Moreover, one transition temperature for $M^{pro}_{\Delta(297-306)}$ was determined to be 68.2°C, which is higher than that of WT (53.9°C). Further characterization of these transitions is needed to address what structure is attributed to the unfolding, when narrowed down to the area of a single α-helix in $M^{pro}_{III}$ and $M^{pro}_{\Delta(297-306)}$.

### *Reversible unfolding of N- and C-terminal-truncated SARS-CoV $M^{pro}$ in guanidinium chloride*

Previously, we found that the guanidinium chloride (GdmCl)-induced unfolding of WT $M^{pro}$ in 30 mM Tris–HCl was completely reversible (*10*). Similar results were obtained in all engineered SARS-CoV $M^{pro}$ used in the present study and representative data are shown in Fig. 3. The C-terminal-truncated $M^{pro}$ ($M^{pro}_{(I+II)}$, $M^{pro}_{(I+II)+loop}$), with only one tryptophanyl residue, remained intact and showed less fluorescence in the folded state (Fig. 3A and B) than the N-terminal truncated $M^{pro}$ ($M^{pro}_{III}$, $M^{pro}_{loop+III}$), which has two tryptophanyl residues (Fig. 3C and D). Domains (I+II) and III were found to fold independently. GdmCl-induced unfolding was thus very suitable for conformational stability studies of the enzyme. We therefore attempted to probe the structural role of the loop of $M^{pro}$.

The AEW, representing the changes in both fluorescence wavelength and fluorescence intensity, was used to estimate the stability parameters of the unfolding/refolding process (Table II). SARS-CoV $M^{pro}$ has three Trp residues, $Trp^{31}$ in domain I and $Trp^{207}/Trp^{218}$

in domain III. The fluorescence of these Trp mutants showed the conformational changes of individual domains. With both $Trp^{207}$ and $Trp^{218}$ at domain III mutated to phenylalanine, $M^{proW31}_{(W207F/W218F)}$ could only report the conformational changes in the N-terminal domain (Fig. 1A). $M^{pro(W207/W218)}_{W31F}$, a single mutant with $Trp^{31}$ mutated to phenylalanine, revealed structural information of the C-terminal domain (Fig. 1A). All these unfoldings were completely reversible and showed the independent folding of domains (I+II) and III (Fig. 3). Both the N- and C-terminal domains followed monophasic unfolding curves. However these did not coincide, as demonstrated in $M^{proW31}_{(W207F/W218F)}$ and $M^{pro(W207/W218)}_{W31F}$ (Fig. 4A). It was obvious that there were differential conformational stabilities for domains (I+II) and III. The unfolding curves of the N-terminal domains, with or without a loop, $M^{pro}_{(I+II)+loop}$ and $M^{pro}_{(I+II)}$, and the C-terminal domains with or without loops, $M^{pro}_{loop+III}$ and $M^{pro}_{III}$, are shown in Fig. 4B. These data indicate that the conformational changes in the C-terminal domain-truncated enzymes ($M^{pro}_{(I+II)+loop}$ and $M^{pro}_{(I+II)}$) and in the N-terminal domain-truncated enzymes ($M^{pro}_{loop+III}$ and $M^{pro}_{III}$) were similar to $M^{proW31}_{(W207F/W218F)}$ and $M^{pro(W207/W218)}_{W31F}$, respectively. It was evident that the N-terminal domain of $M^{pro}$ underwent unfolding ($[GdmCl]_{0.5} \sim 1.5$ M) earlier than the C-terminal domain ($[GdmCl]_{0.5} \sim 4$ M). According to the stability parameters shown in Table II, the long loop was able to stabilize the truncated protease in GdmCl, especially for the C-terminal domain.

**Table I. Structural characteristic of SARS-CoV full-length, N- and C-terminal-truncated $M^{pro}$ and complex in crystal and solution.**

| Protease | Quaternary structure | α-Helix (%) | β-strand (%) | Random coil (%) | Normalized NRMSD[a] | $T_m$[b] |
|---|---|---|---|---|---|---|
| 1uk3[c] | Dimer | 24 | 29 | 47 | | |
| 2qcy[c] | Monomer | 26 | 28 | 46 | | |
| $M^{pro}$ | Dimer | 23[d] | 24 | 53 | 0.032 | 53.9 |
| $M^{pro}_{\Delta(297-306)}$ | Monomer | 17 | 29 | 54 | 0.027 | 68.2 |
| $M^{pro}_{(I+II)}$ [e] | Monomer | 8 | 33 | 59 | 0.022 | 50.2 |
| $M^{pro}_{(I+II)+loop}$ | Monomer | 5 | 37 | 58 | 0.045 | 52.5 |
| $M^{pro}_{III}$ | Monomer | 48 | 15 | 37 | 0.021 | 57.7 |
| $M^{pro}_{loop+III}$ | Monomer | 71 | 8 | 21 | 0.009 | 61.2 |
| $M^{pro}_{(I+II)+loop}$ | | 19 | 26 | 55 | 0.035 | 49.2 |
| $M^{pro}_{III}$(Mixture1) | | | | | | 59.2 |
| $M^{pro}_{(I+II)}$ | | 41 | 15 | 44 | 0.039 | 49.5 |
| $M^{pro}_{loop+III}$(Mixture2) | | | | | | 62.8 |
| $M^{pro}_{(I+II)}$ | | 19 | 27 | 54 | 0.031 | 46.5 |
| $M^{pro}_{III}$(Mixture3) | | | | | | 58.6 |

[a]Normalized root mean SD of the secondary structure fitting results.
[b]Melting temperature determined by the CD spectropolarimeter.
[c]Crystal structure determined by X-ray crystallography.
[d]Percentage calculated by CDSSTR programme.
[e]The data is from 171.36 μM $M^{pro}_{(I+II)}$, which is more than the other enzyme, CD analysis is only reliable under this condition.

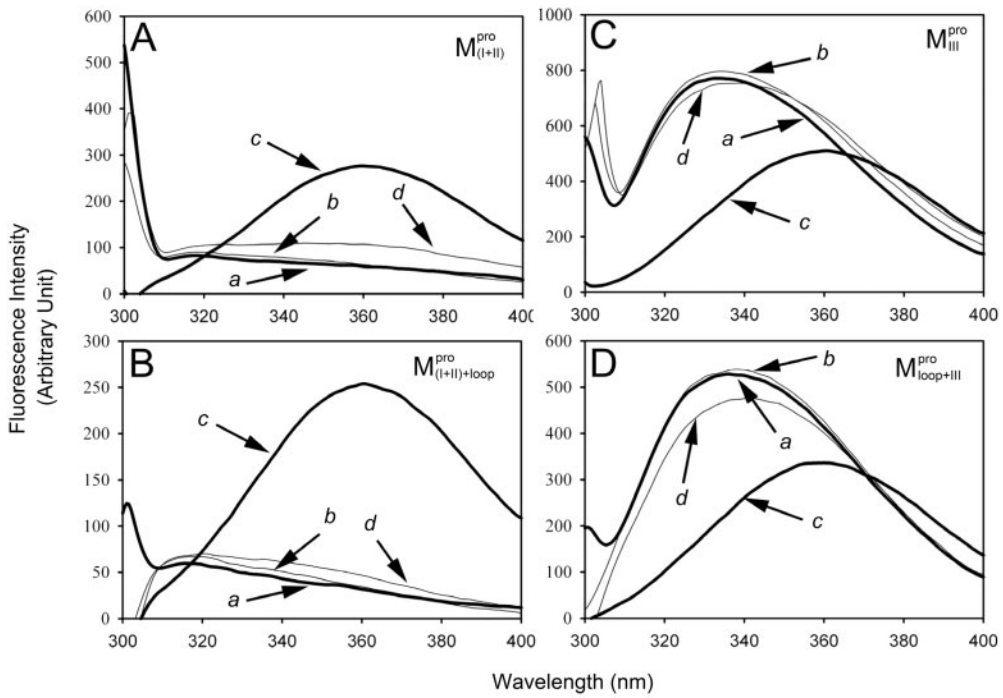


**Fig. 3 Equilibrium unfolding/refolding of the N- and C-terminal-truncated SARS-CoV $M^{pro}$ in GdmCl.** The enzyme (50.6 μg/ml for $M^{pro}_{(I+II)}$, 48.7 μg/ml for $M^{pro}_{(I+II)+loop}$, 2.7 μg/ml for $M^{pro}_{III}$ and 6.8 μg/ml for $M^{pro}_{loop+III}$) in 30 mM Tris–HCl buffer (pH 7.7) was excited with 295 nm UV light and the protein intrinsic fluorescence emission spectra were monitored at 30°C. (A) $M^{pro}_{(I+II)}$; (B) $M^{pro}_{(I+II)+loop}$; (C) $M^{pro}_{III}$; (D) $M^{pro}_{loop+III}$. Curves *a*, Native enzyme in 30 mM Tris–HCl buffer; *b*, Native enzyme in the same buffer containing 0.6 M GdmCl; *c*, Unfolded enzyme in 6 M GdmCl buffer and *d*, Refolded enzyme in 0.6 M GdmCl buffer. In *d*, the protein was unfolded in 6 M GdmCl buffer for 10 min and then diluted 10-fold with the buffer. All samples were treated with GdmCl buffer or buffer-only simultaneously for the same duration. The final enzyme concentrations were the same for all spectra.

**Table II. Stability parameters of full-length, N- and C-terminal-truncated SARS-CoV $M^{pro}$ and complex.**

| Protease | $\Delta G_{N\text{-}I}$ (kcal mol$^{-1}$) | $m_{N\text{-}I}$ (kcal mol$^{-1}$ M$^{-1}$) | $[GdmCl]_{0.5,N\text{-}I}$ (M) | $\Delta G_{I\text{-}U}$ (kcal mol$^{-1}$) | $m_{I\text{-}U}$ (kcal mol$^{-1}$ M$^{-1}$) | $[GdmCl]_{0.5,I\text{-}U}$ (M) |
|---|---|---|---|---|---|---|
| $M^{pro}$ | 2.07 ± 0.33 | 1.66 ± 0.23 | 1.3 ± 0.3 | 11.58 ± 0.58 | 2.79 ± 0.14 | 4.2 ± 0.3 |
| $M^{proW31}_{(W207F/W218F)}$ | 2.22 ± 0.43 | 1.62 ± 0.27 | 1.4 ± 0.4 | – | – | – |
| $M^{pro(W207/W218)}_{W31F}$ | – | – | – | 12.08 ± 0.66 | 2.87 ± 0.16 | 4.2 ± 0.2 |
| $M^{pro}_{(I+II)+loop}$ | 2.03 ± 0.52 | 0.98 ± 0.33 | 2.1 ± 0.8 | 8.63 ± 0.60 | 2.13 ± 0.14 | 4.1 ± 0.3 |
| $M^{pro}_{III}$(Mixture 1) | | | | | | |
| $M^{pro}_{(I+II)}$ | 3.81 ± 0.95 | 2.62 ± 0.64 | 1.5 ± 0.5 | 9.92 ± 0.33 | 2.36 ± 0.08 | 4.2 ± 0.2 |
| $M^{pro}_{loop+III}$(Mixture 2) | | | | | | |
| $M^{pro}_{(I+II)}$ | 4.32 ± 1.74 | 3.19 ± 1.26 | 1.4 ± 0.8 | 6.86 ± 0.45 | 1.80 ± 0.12 | 3.8 ± 0.4 |
| $M^{pro}_{III}$(Mixture 3) | | | | | | |
| $M^{pro}_{(I+II)+loop}$ | 3.41 ± 0.78 | 2.29 ± 0.52 | 1.5 ± 0.5 | 9.79 ± 0.41 | 2.34 ± 0.10 | 4.2 ± 0.3 |
| $M^{pro}_{loop+III}$(Mixture 4) | | | | | | |
| $M^{pro}_{(I+II)}$ | 5.28 ± 0.35 | 3.87 ± 0.25 | 1.4 ± 0.1 | | | |
| $M^{pro}_{(I+II)+loop}$ | 5.62 ± 1.04 | 3.97 ± 0.69 | 1.4 ± 0.3 | | | |
| $M^{pro}_{III}$ | | | | 5.99 ± 0.15 | 1.71 ± 0.04 | 3.5 ± 0.1 |
| $M^{pro}_{loop+III}$ | | | | 9.07 ± 0.30 | 2.23 ± 0.07 | 4.1 ± 0.1 |

### *Unfolding and enzyme activity of the mixtures of the SARS-CoV $M^{pro}$*

To test whether the covalent linkage between domains (I + II) and III would be essential for the structural integrity and enzymatic activity of $M^{pro}$, we designed four nicked proteins. In the presence of GdmCl, all mixtures had similar unfolding curves with the WT $M^{pro}$, which followed a biphasic-unfolding curve (Fig. 5). All these mixtures had a free energy of the first phase that was lower than the C-terminal domain-truncated enzyme (Table II). These results indicated that there were intimate interactions between N- and C-terminal domains and that the C-terminal domain III could stabilize the N-terminal domain. However, the presence of an N-terminal domain decreased the stability of the C-terminal domain III. The stability parameters of Mixture 3 with no loops were similar to the truncated enzyme. Thus, the absence of the loop weakened the interaction between domains (I+II) and III.

The enzyme activities of WT $M^{pro}$,$M^{pro}_{(I+II)+loop}$, $M^{pro}_{(I+II)}$ and the mixtures were measured using fluorescence-based peptide cleavage assays. The fluorescence enhancement of WT $M^{pro}$ was collinear with time (Supplementary Fig. S3). However, $M^{pro}_{(I+II)+loop}$, $M^{pro}_{(I+II)}$ and all nicked enzymes were enzymatically inactive, even at a 10-fold concentration excess of the WT $M^{pro}$ (Supplementary Fig. S3).

### *Quaternary structures of the full-length, N- and C-terminal-truncated SARS-CoV $M^{pro}$ and various mixtures*

The quaternary structures of SARS-CoV $M^{pro}$ were examined using AUC and the sedimentation velocity data were analysed using the continuous size distribution model with the SEDFIT programme (Fig. 6A). In experiments, no aggregation was observed at sedimentation coefficients of up to 20 S. The excellent matching of the experimental data points and the curve fitting results (Supplementary Fig. 4SA), the randomly distributed residual values (Supplementary Fig. 4SB), and the homogeneous bitmap picture (Supplementary Fig. 4SC) of the SARS-CoV $M^{pro}$ all indicated that a highly reliable model for the sedimentation velocity experiments was obtained and the AUC was an excellent biophysical probe for accessing the quaternary structure of the enzyme. The WT $M^{pro}$ existed as a dimeric form in solution but the $M^{pro}_{\Delta(297-306)}$, N- and C-terminal domain-truncated $M^{pro}$ forms were monomeric (Table I and Fig. 6A). A mixture of N- and C-terminal domain-truncated $M^{pro}$ forms remained monomeric in solution (Fig. 6A). As dimerization is the prerequisite for enzyme activity of the $M^{pro}$, the absence of any proteolytic activity is easily explained.

We examined the possibility of using WT $M^{pro}$ as a template to induce the dimerization of mixtures (Fig. 6B). The WT proteases sedimented at 4.4 S with a molar mass of 66 kDa, in agreement with the molecular weight of the dimer (67.6 kDa) calculated from its amino acid sequence. There was a peak between 1.6 S and 2.1 S, corresponding to the truncated proteases. The WT $M^{pro}$ could not induce dimer formation of the mixture. Thus, the covalent linkage between domains (I + II) and domain III was essential for the structural integrity of SARS-CoV $M^{pro}$.

## Discussion

Complementation of protein fragments (reconstitution experiments) is an efficient method for inspecting the requirements for protein function and folding pathways (*22–23*). The function of the protein may be retained in the reconstituted complex depending on the number of cuts and where in the sequence they are made. Fragment complementing systems have been used to probe the molecular features required for fragments to associate and fold into a stable native-like protein complex (*24*) and to elucidate autonomous folding units. We applied this concept in a broad way to a protein that contains multiple domains. Our

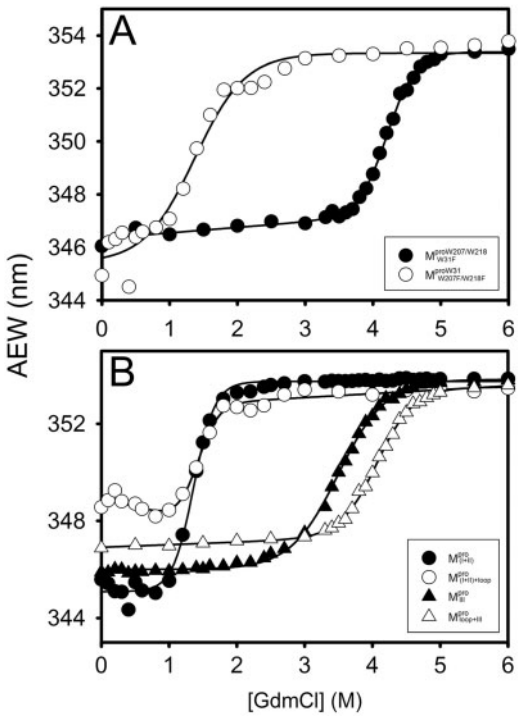


**Fig. 4 Unfolding curves of the full-length Trp-mutated, N- and C-terminal domain-truncated SARS-CoV $M^{pro}$.** (A) Unfolding of the full-length Trp-mutated enzyme. $M^{proW31}_{(W207F/W218F)}$ (open circle) in 0.54 μM; $M^{proW207/W218}_{W31F}$ (filled circle) in 0.27 μM. $M^{proW31}_{(W207F/W218F)}$ represents the structural change of N-terminal domain. $M^{proW207/W218}_{W31F}$ represents the structural change of C-terminal domain. (B) Unfolding of the N- and C-truncated protease bearing loop or not. $M^{pro}_{(I+II)}$ (filled circle) and $M^{pro}_{(I+II)+loop}$ (open circle) in 0.54 μM; $M^{pro}_{III}$ (filled triangle) and $M^{pro}_{loop+III}$ (open triangle) in 0.27 μM. The protein in 30 mM Tris–Cl (pH 7.7) was incubated with various concentrations of GdmCl for 10 min. The protein solution was excited at 295 nm and the emission fluorescence was recorded at 30°C. AEW (Equation 2) of each protein was used to present the unfolding data. Symbols are experimental data and the smooth curves are computer-generated best-fitted lines according to a two-state unfolding model. Truncated mutant had similar curve to the full-length mutated $M^{pro}$. $M^{pro}_{(I+II)}$ and $M^{pro}_{(I+II)+loop}$ were similar to $M^{proW31}_{(W207F/W218F)}$ while $M^{pro}_{III}$ and $M^{pro}_{loop+III}$ were similar to $M^{proW207/W218}_{W31F}$.

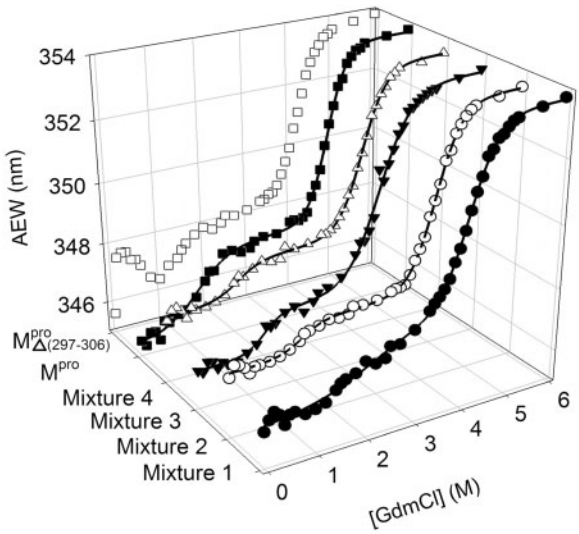


**Fig. 5 Unfolding curves of the WT, deletion mutants and mixtures of the $M^{pro}$.** Mixture 1 ($M^{pro}_{(I+II)+loop}/M^{pro}_{III}$) (filled circle), Mixture 2 ($M^{pro}_{(I+II)}/M^{pro}_{loop+III}$) (open circle), Mixture 3 ($M^{pro}_{(I+II)}/M^{pro}_{III}$) (inverted filled triangle), Mixture 4 ($M^{pro}_{(I+II)+loop}/M^{pro}_{loop+III}$) (open triangle), WT (filled square) and $M^{pro}_{\Delta(297-306)}$ (open square). The enzyme in 30 mM Tris–HCl buffer (pH 7.7) was incubated with various concentrations of GdmCl for 10 min. The enzyme solution was excited at 295 nm and the emission fluorescence was recorded at 30°C. AEW of the enzyme was used to analyse the unfolding data. All mixtures had similar unfolding curve to WT $M^{pro}$. For direct comparison purpose, the duplicate presentation of data shown in two-dimensional space is provided (Supplementary Fig. 5S).

previous studies indicated that the C-terminal domain-truncated $M^{pro}$ could fold independently into an intact chymotrypsin-like fold, but it showed no enzyme activity (*10*). Our other experimental data showed that the extra C-terminal domain played a pivotal role in controlling dimerization and enzyme activity (*25*). The present results on $M^{pro}$ fragments further defined our earlier studies that domain III acts to stabilize the catalytic N-terminal domain (I+II). Furthermore, the different SARS-CoV $M^{pro}$-complementing fragment systems described herein were used to probe the structural role of the loop. We confirmed that the covalent linkage between domains (I+II) and (III) is essential, which suggests that the long loop might stabilize the C-terminal domain. It is likely that some critical structural stabilizing interactions are conserved in the loop-retaining fragments with respect to those present in the natural and longer SARS-CoV fragment. Because of this, $M^{pro}_{(I+II)+loop}$ displays a higher content of β-sheets, which is its major secondary structure, than does $M^{pro}_{(I+II)}$, and $M^{pro}_{loop+III}$ instead displays more helicity than $M^{pro}_{III}$. Examining crystal structures of $M^{pro}$ reveals that residues around 187 have interactions with residues around 40 and 54 of N-terminal domain, and residues around 198 have interactions with residues around 240 of C-terminal domain. Thus, the truncations we did very likely affected the folding and stability of the domains. This may well explain the differences, we observed by CD and the difference in stability for domains with and without the linker loop.

Spectroscopic studies by far-UV CD and reversible unfolding analysis in GdmCl indicate that fragments $M^{pro}_{(I+II)}$, $M^{pro}_{(I+II)+loop}$, $M^{pro}_{III}$ and $M^{pro}_{loop+III}$ adopted folded states in aqueous solution at neutral pH and therefore appeared to possess correct tertiary structures. The mixtures of various combinations of these fragments (nicked SARS CoV $M^{pro}$) revealed structural features similar to those of the intact protein. However, some differences between the nicked and intact species were found from the AUC studies. These negative results seem to base on the mixing of the individual domains which are non-interacting. Little convinced result to show that the component in these mixtures could interact and form the structure like WT $M^{pro}$. The nicked proteins attained significant secondary and tertiary structures, as did the native enzyme. However, they seemed to lack specific domain–domain complementation (or interactions), thereby precluding a final folding step to assume the catalytic competent

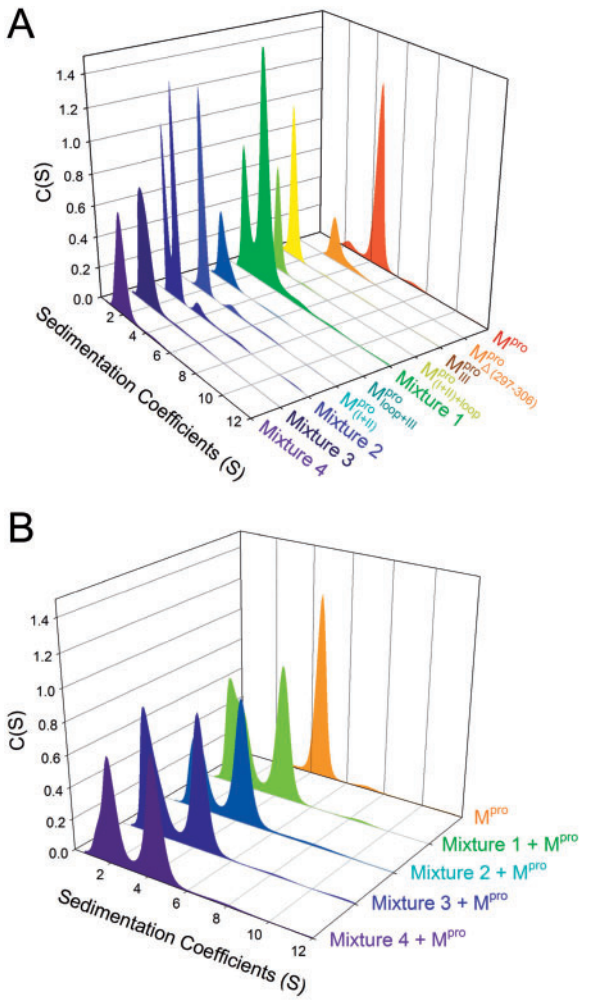


**Fig. 6 Quaternary structural analysis of the full-length, N- and C-terminal-truncated SARS-CoV $M^{pro}$ and non-covalent complexes.** Continuous sedimentation of the protein in 30 mM Tris–Cl (pH 7.7) was monitored with analytical ultracentrifuge. The sedimentation velocity data were analysed with SEDFIT. (A) The continuous sedimentation coefficient distribution of the protein [23.85 μM for $M^{pro}$, $M^{pro}_{\Delta(297-306)}$, $M^{pro}_{(I+II)}$, Mixture 2 ($M^{pro}_{(I+II)}/M^{pro}_{loop+III}$), Mixture 3 ($M^{pro}_{(I+II)}/M^{pro}_{III}$) and Mixture 4 ($M^{pro}_{(I+II)+loop}/M^{pro}_{loop+III}$), 22.8 μM for $M^{pro}_{(I+II)+loop}$, $M^{pro}_{III}$, $M^{pro}_{loop+III}$ and Mixture 1 ($M^{pro}_{(I+II)+loop}/M^{pro}_{III}$)]. (B) Quaternary structure of the complex in the presence of WT $M^{pro}$. All enzyme preparations were in 23.85 μM concentration except for Mixture 1 in 22.8 μM. For direct comparison purpose, the duplicate presentation of data shown in two-dimensional space is provided (Supplementary Fig. 6S).

conformation, as in the native $M^{pro}$. Therefore, it appears that the individual N- and C-terminal domains can fold independently, but cannot associate to form the enzymatic active dimer. This is consistent with current experimental observations that only the dimeric form of SARS-CoV $M^{pro}$ possesses biological function and that the dissociated monomer is enzymatically inactive (*26–28*). Shi *et al.* (*3*) have reported that the isolated extra domain III existed as a dimer even at a very low protein concentration and showed no concentration dependence by dynamic light scattering and size-exclusion FPLC analyses. A more recent paper reports that the C-terminal domain of SARS-CoV $M^{pro}$ can form a 3D domain swapped dimer in the crystalline form (*29*). However, under our AUC experimental conditions, the $M^{pro}_{III}$ and $M^{pro}_{loop+III}$ seem to exist majorly as a monomeric form in solution (Fig. 6A).

Taniuchi *et al.* (*22*) developed the concept of permissible sites of cleavage of a protein for producing fragment complementing systems and deduced that cleavages can occur at exposed, flexible loops but not in a typical secondary structure at the level of chain segments (Fig. 1C). In the present case, the cleaved peptide bonds are close to the adjacent essential structural domains. Therefore, the 183–184 and 199–200 peptide bond regions might not be permissible sites of cleavage. It would be intriguing to test chain fission at the middle of this loop.

In summary, because the C- and N-terminal domains of $M^{pro}$ can fold independently and show a completely reversible unfolding and refolding process in GdmCl, SARS-CoV $M^{pro}$ provides an ideal model system to delineate complicated protein folding in a feasible stepwise pathway. We demonstrated here that the mixtures of $M^{pro}$ also showed a reversible biphasic unfolding curve, similar to the intact WT form. Nevertheless, there was no enzyme activity found for any mixture. In addition, no dimer formation between any truncated protease was observed with or without the presence of WT $M^{pro}$ as a template. These results indicate that the covalent linkage between the chymotrypsin-like domain and the extra domain must be essential for enzyme activity and for the quaternary structural integrity of the SARS-CoV $M^{pro}$.

## Supplementary Data

Supplementary Data are available at *JB* Online.

### Funding

This work was supported by National Science Council, Republic of China (NSC 96-2113-M-415-010).

### Conflict of interest

None decalred.

## References

1. 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
2. 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. USA* **100**, 13190–13195
3. 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
4. Liang, P.H. (2006) Characterization and inhibition of SARS-coronavirus main protease. *Curr. Top. Med. Chem.* **6**, 361–376
5. Ratia, K., Saikatendu, K.S., Santarsiero, B.D., Barretto, N., Baker, S.C., Stevens, R.C., and Mesecar, A.D. (2006)

Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. *Proc. Natl Acad. Sci. USA* **103**, 5717–5722
6. Bartlam, M., Yang, H., and Rao, Z. (2005) Structural insights into SARS coronavirus proteins. *Curr. Opin. Struct. Biol.* **15**, 664–672
7. Anand, K., Palm, G.J., Mesters, J.R., Siddell, S.G., Ziebuhr, J., and Hilgenfeld, R. (2002) Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J.* **21**, 3213–3224
8. Lin, P.Y., Chou, C.Y., Chang, H.C., Hsu, W.C., and Chang, G.G. (2008) Correlation between dissociation and catalysis of SARS-CoV main protease. *Arch. Biochem. Biophys.* **472**, 34–42
9. DeLano, W.L. (2002) *The Pymol Molecular Graphic System*., DeLano Scientific, San Carlos, CA, USA
10. Chang, H.P., Chou, C.Y., and Chang, G.G. (2007) Reversible unfolding of the severe acute respiratory syndrome coronavirus main protease in guanidinium chloride. *Biophys. J.* **92**, 1374–1383
11. 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., and Rao, Z. (2008) Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J. Virol.* **82**, 2515–2527
12. Lobley, A., Whitmore, L., and Wallace, B.A. (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* **18**, 211–212
13. Whitmore, L. and Wallace, B.A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668–W673
14. Compton, L.A. and Johnson, W.C. Jr. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.* **155**, 155–167
15. Manavalan, P. and Johnson, W.C. Jr. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* **167**, 76–85
16. Sreerama, N. and Woody, R.W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260
17. Royer, C.A., Mann, C.J., and Matthews, C.R. (1993) Resolution of the fluorescence equilibrium unfolding profile of trp aporepressor using single tryptophan mutants. *Protein Sci.* **2**, 1844–1852
18. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619
19. Brown, P.H. and Schuck, P. (2006) Macromolecular Size-and-Shape Distributions by Sedimentation Velocity Analytical Ultracentrifugation. *Biophys. J.* **90**, 4651–4661
20. Pace, C.N. (1990) Measuring and increasing protein stability. *Trends Biotechnol.* **8**, 93–98
21. 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
22. Taniuchi, H., Parr, G.R., and Juillerat, M.A. (1986) Complementation in folding and fragment exchange. *Methods Enzymol.* **131**, 185–217
23. de Prat-Gay, G. (1996) Association of complementary fragments and the elucidation of protein folding pathways. *Protein Eng.* **9**, 843–847
24. Musi, V., Spolaore, B., Picotti, P., Zambonin, M., De Filippis, V., and Fontana, A. (2004) Nicked apomyoglobin: a noncovalent complex of two polypeptide fragments comprising the entire protein chain. *Biochemistry* **43**, 6230–6240
25. Hsu, W.C., Chang, H.C., Chou, C.Y., Tsai, P.J., Lin, P.I., and Chang, G.G. (2005) Critical assessment of important regions in the subunit association and catalytic action of the severe acute respiratory syndrome coronavirus main protease. *J. Biol. Chem.* **280**, 22741–22748
26. Chou, C.Y., Chang, H.C., Hsu, W.C., Lin, T.Z., Lin, C.H., and Chang, G.G. (2004) Quaternary structure of the severe acute respiratory syndrome (SARS) coronavirus main protease. *Biochemistry* **43**, 14958–14970
27. 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
28. Graziano, V., McGrath, W.J., Yang, L., and Mangel, W.F. (2006) SARS CoV main proteinase: The monomer-dimer equilibrium dissociation constant. *Biochemistry* **45**, 14632–14641
29. Zhong, N., Zhang, S., Xue, F., Kang, X., Zou, P., Chen, J., Liang, C., Rao, Z., Jin, C., Lou, Z., and Xia, B. (2009) C-terminal domain of SARS-CoV main protease can form a 3D domain-swapped dimer. *Protein Sci.* **18**, 839–844