Reversible Unfolding of the Severe Acute Respiratory Syndrome Coronavirus Main Protease in Guanidinium Chloride

Hui-Ping Chang, Chi-Yuan Chou, and Gu-Gang Chang
Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Taipei, Taiwan

ABSTRACT Chemical denaturant sensitivity of the dimeric main protease from severe acute respiratory syndrome (SARS) coronavirus to guanidinium chloride was examined in terms of fluorescence spectroscopy, circular dichroism, analytical ultracentrifuge, and enzyme activity change. The dimeric enzyme dissociated at guanidinium chloride concentration of <0.4 M, at which the enzymatic activity loss showed close correlation with the subunit dissociation. Further increase in guanidinium chloride induced a reversible biphasic unfolding of the enzyme. The unfolding of the C-terminal domain-truncated enzyme, on the other hand, followed a monophasic unfolding curve. Different mutants of the full-length protease (W31 and W207/W218), with tryptophanyl residue(s) mutated to phenylalanine at the C-terminal or N-terminal domain, respectively, were constructed. Unfolding curves of these mutants were monophasic but corresponded to the first and second phases of the protease, respectively. The unfolding intermediate of the protease thus represented a folded C-terminal domain but an unfolded N-terminal domain, which is enzymatically inactive due to loss of regulatory properties. The various enzyme forms were characterized in terms of hydrophobicity and size-and-shape distributions. We provide direct evidence for the functional role of C-terminal domain in stabilization of the catalytic N-terminal domain of SARS coronavirus main protease.

INTRODUCTION

Severe acute respiratory syndrome (SARS) was one of the most serious epidemic diseases in 2003 and is still a threat to the public health worldwide (1). A specific SARS coronavirus (CoV) was identified to be the contagious agent. Maturation of the SARS-CoV involves proteolytic processing of the virus polyproteins by the chymotrypsin-like main protease (3CLpro) and a papain-like protease (PLpro). SARS-CoV might invade cells by a protease-mediated endosomal pathway (2). Any knowledge about the structural stability of viral protease can be very rewarding for the development of antivirus drugs (3–6).

The 3CLpro from SARS-CoV was the first protein of this virus to have its three-dimensional structure solved (7,8). As had previously been shown for the 3CLpro of transmissible gastroenteritis (corona)virus (9), the SARS-CoV enzyme also exists as a dimer in solution and in crystals, with each subunit containing three structural domains (Fig. 1). The catalytic dyad His^{41}-Cys^{145} is located between the interfaces of domains I (residues 8–99) and domain II (residues 100–183), which constitute the chymotrypsin-like folding scaffold. This catalytic N-terminal domain is connected by a long loop (residues 184–199) to the helical domain III (residues 200–306). The functional role of the extra C-terminal domain III is less clear. Numerous experimental evidence shows the importance of N- and C-terminal residues in the dimerization (4,10–12). Structural studies indicate the close proximity of C-terminus and N-finger (N-terminal residues 1–7) (9) from subunit A to the active site of subunit B that led to the development of a novel autocleavage mechanism that can explain the dimeric nature of the mature enzyme (13).

Here we demonstrate that the SARS main protease provides us a very useful model for the folding studies of multisubunit, multidomain proteins. Novel folding intermediates were identified, which allow us to delineate unequivocally the functional role of domain III.

MATERIALS AND METHODS

Construction of 3CLpro expression vectors

The genes of the full-length 3CLpro and C-terminal truncated $3CL_{I+II}$ of the SARS-CoV main protease were amplified by polymerase chain reaction (PCR) with appropriate primers. The forward primers for full length was 5′-GGTGGTCATATGAGTGGTTTTAGG-3′, and the reverse primer was 5′-AACTCGAGGGTAACACCAGAG-3′. After digestion with *Bgl*II and *Xho*I, the PCR product was divided into two fragments, 168-bp and 747-bp. The 168-bp fragment was then digested with *Nde*I. Finally, the 168-bp *Nde*I-*Bgl*II and 747-bp *Bgl*II-*Xho*I fragments were coligated to the *Nde*I and *Xho*I sites of the vector pET-29a(+) (Novagen, Madison, WI). The $3CL_{I+II}$ was subsequently amplified by PCR, using the forward primer of 5′-TGAAGATCTGCTCATTCGCAA-3′, and reverse primer of 5′-AACTCGACTATGGTTGTGTCTG-3′. After digestion with *Bgl*II and *Xho*I, the PCR products were inserted into the *Bgl*II and *Xho*I sites of the pET-SARS main protease.

Submitted June 19, 2006, and accepted for publication November 2, 2006.

Address reprint requests to Hui-Ping Chang, Dept. of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, 155 Li-Nong St., Section 2, Taipei 112, Taiwan. E-mail: huiping_chp@hotmail.com.

Abbreviations used: SARS, severe acute respiratory syndrome; CoV, coronavirus; 3CLpro, 3-chymotrypsin-like SARS-CoV main protease; $3CL_{I+II}$, C-terminal domain-truncated SARS-CoV main protease; W31 (or more precisely W207F/W218F), double mutant of SARS-CoV 3CLpro with tryptophanyl residues at 207 and 218 mutated to phenylalanine; W207/W218 (or more precisely W31F), point mutant of SARS-CoV 3CLpro with tryptophanyl residue at position 31 mutated to phenylalanine; AEW, average emission wavelength; CD, circular dichroism; AUC, analytical ultracentrifugation; GdnCl, guanidinium chloride; f/f_o, frictional ratio; ANS, 1-anilino-8-naphthalene sulfonic acid.

0006-3495/07/02/1374/10 $2.00

doi: 10.1529/biophysj.106.091736

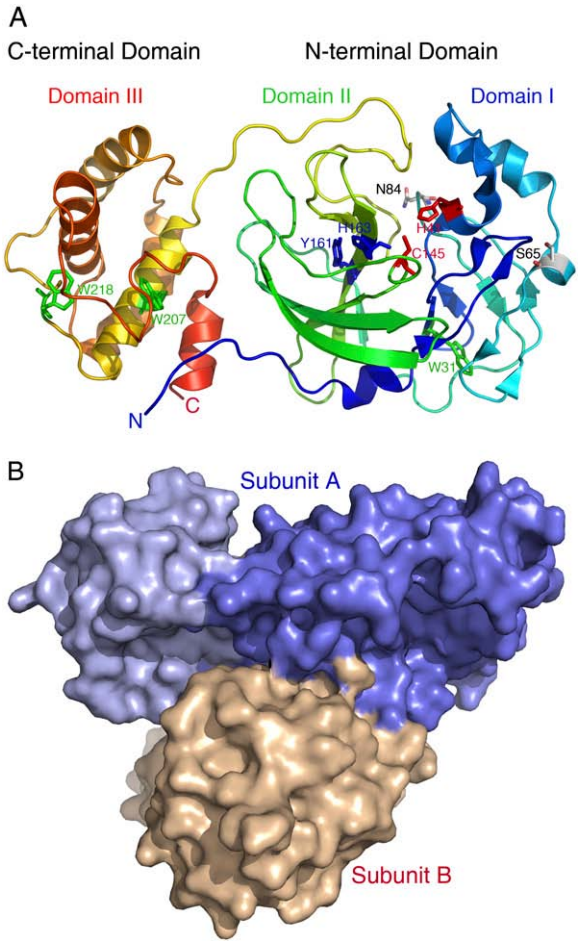

FIGURE 1 Structural features of SARS-CoV main protease. (*A*) The three structural domains of SARS-CoV main protease are shown in rainbow color (Protein Data Bank code: 1Z1I). The catalytic dyad His^{41}-Cys^{145} between domains I and II are highlighted with red side chain. The important residues Tyr^{161} and His^{163} in the substrate-binding pocket are shown in blue side chain. The tryptophanyl residues Trp^{31}, Trp^{207}, and Trp^{218} are shown in green side chain. Other two residues Ser^{65} and Asn^{84} located at surface are also shown. (*B*) Surface model showing the unsymmetrical dimer association mode. N-terminal domain and C-terminal domain of subunit *A* are in blue and light blue, respectively. Generate with Pymol (37).

Site-directed mutagenesis (14) was used to construct pET-H41A, pET-S65A, pET-N84A, pET-C145A, pET-Y161A, and pET-H163A vectors. The forward primers for H41A, S65A, N84A, C145A, Y161A, and H163A were 5′-CTGTCCAAGA**GCG**GTCATTTGCAC-3′, 5′-CAAATCCAACCAT-**GCG**TTTCTTGTTC-3′, 5′-CATTCTATGCAA**GCG**TGTCTGCTTAG-3′, 5′-CTTAATGGATCA**GCG**GGTAGTGTTG-3′, 5′-GTCTTTCTGC**GCG**-ATGCATCATATG-3′, and 5′-CTGCTATATG**GCG**CATATGGAGC-3′, respectively; and the reverse primers were 5′-GTGCAAATGAC**CGC**-TCTTGGACAG-3′, 5′-GAACAAGAAA**CGC**ATGGTTGGATTTG-3′, 5′-CTAAGCAGACA**CGC**TTGCATAGAATG-3′, 5′-CAACACTACC**CG-C**TGATCCATTAAG-3′, 5′-CATATGATGCAT**CGC**GCAGAAAGAC-3′, and 5′-GCTCCATATG**CGC**CATATAGCAG-3′, respectively. The pET-3CL expression vectors were used as templates and the primers as described above were used to mutate the His^{41}, Ser^{65}, Asn^{84}, Cys^{145}, Tyr^{161}, and His^{163} codons to alanine codon by PCR. The mutational sites are shown underlined and in bold.

For the tryptophan mutants of 3CLpro, the forward primers for W31F, W207F, and W218F were 5′-CTCTTAATGGATTG**TTC**TTGGATGACA-CAGT-3′, 5′-CCATAACATTAAATGTTTTGGCA**TTC**CTGTATGCT-GC-3′, and 5′-CAATGGTGATAGG**TTC**TTTCTTAATAGATTCACC-3′, respectively. The reverse primers were 5′-ACTGTGTCATCCAA**GAA-**CAATCCATTAAGAG-3′, 5′-GCAGCATACAG**GAA**TGCCAAAACA-TTTAATGTTATGG-3′, and 5′-GGTGAATCTATTAAAGAA**GAA**CC-TATCACCATTG-3′, respectively. For W207F/W218F (W31) double mutant, W207F was used as the template.

All nucleotide sequences were confirmed by autosequencing analysis. A sequence comparison bearing the mutation positions at the active site region is shown in supplemental Fig. 1S (Supplementary Material). The wild-type (WT) enzyme and all mutants were successfully expressed and purified to apparent homogeneity by protocols described previously (11). Typical CD spectra of the recombinant 3CLpro are shown in Supplementary Material (Fig. 2S).

Enzyme activity assay of SARS-CoV main protease using fluorogenic substrate

The enzyme kinetic measurements were performed in 30 mM Tris-Cl buffer (pH 7.7) or phosphate-buffered saline (pH 7.6) 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.

The kinetic parameters were obtained by assaying the enzyme activity at various peptide substrate concentrations around its Michaelis constant (K_m) (11). The initial velocities (v_o) were measured and fitted to the Michaelis-Menten equation (Eq. 1):

$$v_o = \frac{k_{cat}[E]_t[S]}{K_m + [S]}, \tag{1}$$

where k_{cat} is the catalytic constant; $[E]_t$, is the enzyme concentration, and $[S]$, is the substrate concentration.

Reversible unfolding/refolding of the enzyme in GdnCl

WT (3CLpro) or domain III-deleted ($3CL_{I+II}$) SARS-CoV main protease was unfolded with different concentrations of GdnCl in Tris-Cl buffer (30 mM, pH 7.7) at 30°C for 10 min. The unfolding of the enzyme was monitored by fluorescence, CD, sedimentation velocity, and enzyme activity loss. The unfolding was found to reach equilibrium in transient. A 10-min incubation, both unfolding and refolding, was used in the experiments in this study.

Spectrofluorimetric analysis

Fluorescence spectra of the enzyme were monitored in a Perkin-Elmer LS-50B luminescence spectrometer at 30°C, and all spectra were corrected for buffer absorption. The excitation wavelength was set at 280 nm (295 nm for tryptophan mutants, W31 and W207/W218), and the fluorescence emission spectrum was scanned from 300 to 400 nm. 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 analyzed together using the average emission wavelength (AEW) ($\langle\lambda\rangle$) according to Eq. 2 (15):

$$\langle\lambda\rangle = \frac{\sum_{i=\lambda_1}^{\lambda_N}(F_i \times \lambda_i)}{\sum_{i=\lambda_1}^{\lambda_N} F_i}, \tag{2}$$

in which F_i is the fluorescence intensity at the specific emission wavelength (λ_i).

Exposure of the enzyme hydrophobic areas was examined by mixing the enzyme solution with 1-anilino-8-naphthalene sulfonic acid (ANS) and the fluorescence emission spectrum between 400 and 600 nm was recorded with excitation at 395 nm.

Circular dichroism analysis

CD spectra were measured at 30°C with a Jasco J-810 spectropolarimeter under constant N_2 flush. The enzyme solution was scanned from 300 to 190 nm. When the analysis was completed, the buffer was immediately scanned to obtain the baseline. Unfolded enzyme solution was filtered through a 0.22-mm filter (Whatman, UK) before analysis. Denaturing agent included in the buffer was used to monitor the baseline for the unfolded enzyme.

Mean residue ellipticity [Θ] was calculated according to Eq. 3:

$$[\Theta] = \frac{M \times \Theta}{100 \times C \times l}, \tag{3}$$

where l represents the light path length (dm); M, the average molecular weight of amino acids; C, the concentration of protein (mg/ml); and Θ, the observed ellipticity.

Analytical ultracentrifugation

Sedimentation velocity was performed in a Beckman-Coulter XL-A analytical ultracentrifuge with an An50Ti rotor at 20°C and 42,000 rpm in 12-mm double-sector Epon charcoal-filled centerpieces. The ultraviolet (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 free 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 due to unfolding of the protein. Multiple scans at different time points were analyzed with the SEDFIT program (16,17).

The size-and-shape distributions of the enzyme in the presence of GdnCl were analyzed with a new feature of SEDFIT (17) that allowed full exploration of the protein conformational changes. The sedimentation velocity data were fitted using a two-dimensional distribution with respect to frictional ratio $c(s, f/f_o)$ according to Lamm equation (17,18) (Eq. 4):

$$a(r,t) = \int\int c(s,f/f_o)\chi(s,D(s,f/f_o),r,t)dsd(f/f_o), \tag{4}$$

with $a(r,t)$ denoting the observed optical signal at radius r and time t; $\chi(s,D,r,t)$, the solution of the Lamm equation; and $D(s,f/f_o)$, the dependence of diffusion coefficient (D) on sedimentation coefficient (s) and frictional ratio (f/f_o), where

$$\frac{\partial\chi}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[rD\frac{\partial\chi}{\partial r} - s\omega^2 r^2\chi\right] \tag{5}$$

$$D(s,f/f_o) = \frac{\sqrt{2}}{18\pi}\frac{kT}{\sqrt{s}}\frac{1}{\sqrt{(\eta f/f_o)^3}}\sqrt{\frac{1-\bar{\nu}\rho}{\bar{\nu}}}, \tag{6}$$

with ω denoting angular velocity; k, Boltzmann constant; T, absolute temperature; $\bar{\nu}$, enzyme partial specific volume; η, buffer viscosity; and ρ, buffer density.

All two-dimensional distributions were solved and normalized to a confidence level of $p = 0.95$ by maximum entropy and a resolution N of 200 with sedimentation coefficients between 0.1 and 20 S. The anhydrous friction ratio is from 1.0 to 2.0 or 3.5 at a resolution of 10.

Unfolding data analysis

Because unfolding/refolding is a reversible process (see Fig. 2), the unfolding data were treated with the following thermodynamic models by global fitting of the data to Eq. 7 or Eq. 8:

The two-state unfolding model (Scheme I) was described by Eq. 7 (19):

$$N \xleftrightarrow{K_{N-U}} U \qquad \text{(Scheme I)}$$

$$y_{obs} = \frac{y_N + y_U \cdot e^{-\left(\frac{\Delta G_{(H_2O)N\to U} - m_{N\to U}[GdnCl]}{RT}\right)}}{1 + e^{-\left(\frac{\Delta G_{(H_2O)N\to U} - m_{N\to U}[GdnCl]}{RT}\right)}}. \tag{7}$$

The three-state unfolding model (Scheme II) was described by Eq. 8 (20):

$$N \xleftrightarrow{K_{N-I}} I \xleftrightarrow{K_{I-U}} U \qquad \text{(Scheme II)}$$

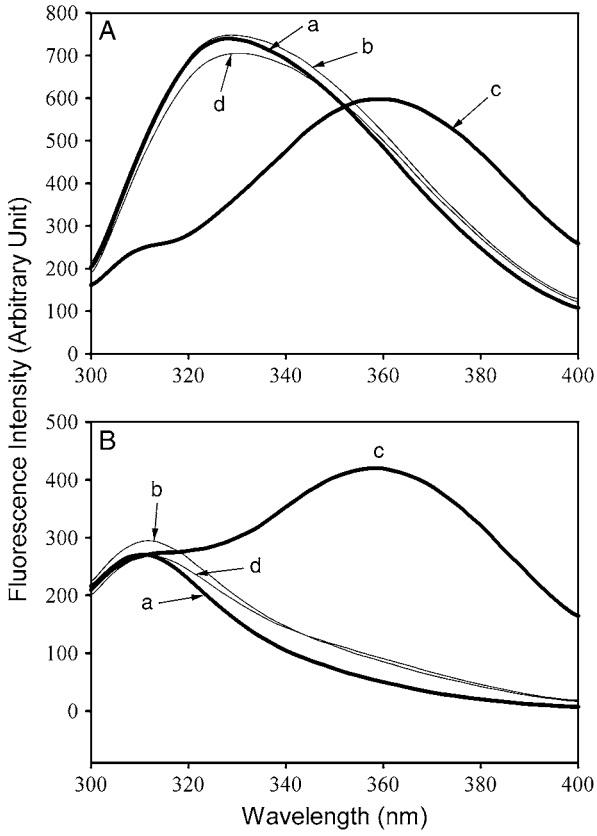

FIGURE 2 Equilibrium unfolding/refolding of the SARS-CoV main protease in GdnCl. The enzyme (5 μg/ml for 3CLpro and 15 μg/ml for $3CL_{I+II}$) in 30 mM Tris-Cl buffer (pH 7.7) was excited with 280 nm UV light and the protein intrinsic fluorescence emission spectra were monitored at 30°C. (*A*) 3CLpro; (*B*) $3CL_{I+II}$. Curves are: (*a*) native enzyme in Tris-Cl buffer; (*b*) native enzyme in the same buffer containing 0.6 M GdnCl; (*c*) unfolded enzyme in 6 M GdnCl-buffer; (*d*) refolded enzyme in 0.6 M GdnCl-buffer. In curve *d*, the protein was unfolded in 6 M GdnCl-buffer for 10 min and then diluted 10-fold with the buffer. All samples were treated with GdnCl-buffer or buffer-only simultaneously for the same duration. The final enzyme concentrations were the same for all spectra.

$$y_{obs} = \frac{y_N + y_I \cdot e^{-\left(\frac{\Delta G_{(H_2O)N\rightarrow I} - m_{N\rightarrow I}[GdnCl]}{RT}\right)} + y_U \cdot e^{-\left(\frac{\Delta G_{(H_2O)N\rightarrow I} - m_{N\rightarrow I}[GdnCl]}{RT}\right)} \cdot e^{-\left(\frac{\Delta G_{(H_2O)I\rightarrow U} - m_{I\rightarrow U}[GdnCl]}{RT}\right)}}{1 + e^{-\left(\frac{\Delta G_{(H_2O)N\rightarrow I} - m_{N\rightarrow I}[GdnCl]}{RT}\right)} + e^{-\left(\frac{\Delta G_{(H_2O)N\rightarrow I} - m_{N\rightarrow I}[GdnCl]}{RT}\right)} \cdot e^{-\left(\frac{\Delta G_{(H_2O)I\rightarrow U} - m_{I\rightarrow U}[GdnCl]}{RT}\right)}}, \quad (8)$$

where y_{obs} is the observed biophysical signal; y_N, y_I, and y_U are the calculated signals of the native, intermediate, and unfolded states, respectively. [GdnCl] is the GdnCl concentration, and $\Delta G_{N\text{-}I}$ and $\Delta G_{I\text{-}U}$ are the free energy changes for the N↔I and I↔U processes, respectively; $m_{N\text{-}I}$ and $m_{I\text{-}U}$ are the sensitivities of the respective unfolding process to denaturant concentration.

RESULTS

Active site region of SARS-CoV main protease

The original crystal structure of SARS-CoV main protease revealed that the dimeric enzyme is half-sited at pH 6.0, in which one of the protomer appears to be active and the other inactive (7). In a recent article, Chen et al. (21) confirmed that the dimeric SARS-CoV main protease is asymmetric and only one protomer is active at a time. They proposed a novel association-activation-catalysis-dissociation mechanism for the enzyme activity control. We checked whether dissociation is the major reason for enzyme inactivation. The following residues were selected for analysis: His^{41} and Cys^{145}, the catalytic dyad; His^{163}, the substrate-binding residue; Tyr^{161}, residue for holding His^{163} in place. Some surface residues at domain I, Ser^{65} and Asn^{84}, were also selected as control. In the Ramachandran plot (22), Asn^{84} is in the disallowed region of the unbound 3CLpro but in a generously allowed region of a 3CLpro-azapeptide inhibitor complex (23). Table 1 summarizes the kinetic parameters of these mutants. Mutation at catalytic dyad residues (H41A and C145A) abolishes almost the complete biological function while the enzyme exists exclusively as dimer. Assuming that the rate-limiting step of the hydrolytic activity of 3CLpro, like chymotrypsin, is at the covalent acylation-deacylation step, then the K_m value for the substrate reflects an apparent binding affinity for the enzyme-substrate complex. However, mutation at S1 site (H163A) of the substrate-binding pocket did not change the substrate binding appreciably as anticipated. The major effect was on the k_{cat} value. H163A possessed ~0.1% of enzyme activity compared with WT. In contrast, the catalytic dyad mutants, H41A and C145A, had only 10^{-6} relative activity left.

An interesting finding was the Y161A mutant, which has only 7×10^{-5} relative enzymatic activity left. Y161A dissociated to an unstable monomer that is prone to polymerize (Supplementary Material, Fig. 3S).

TABLE 1 Kinetic parameters of the SARS-CoV main protease mutants

Protease	Quaternary structure	K_m (μM)	k_{cat} (10^3 s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
Experiment I‡				
3CLpro	Dimer	10.5 ± 3.1	85.5 ± 7.1	8.1 ± 2.5
H41A*	Dimer	–	0.0002	–
C145A*	Dimer	–	0.0002	–
H163A	Dimer	17.5 ± 4.2	0.14 ± 0.01	0.008 ± 0.002
Y161A*	Monomer/ polymer	–	0.006	–
S65A	Dimer	32.0 ± 9.9	96.2 ± 14.3	3.0 ± 1.0
N84A	Dimer	9.9 ± 2.3	64.1 ± 11.5	6.5 ± 1.9
$3CL_{I+II}$	Monomer	ND†	ND†	ND†
Experiment II§				
3CLpro	Dimer	14.5 ± 4.5	5.2 ± 0.7	0.4 ± 0.1
W31	Dimer	18.7 ± 7.8	5.9 ± 1.1	0.3 ± 0.1
W207/W218	Dimer	14.5 ± 4.3	11.0 ± 1.3	0.8 ± 0.3

*The enzyme activities for these mutants were too small to get reliable kinetic parameters, only relative activity was reported.
†Not detectable.
‡Determined in phosphate-buffered saline (pH 7.6).
§Determined in 30 mM Tris-Cl buffer (pH 7.7).

Reversible unfolding of SARS-CoV main protease in GdnCl

SARS-CoV 3CLpro was denatured at high temperature. The thermal denaturation was irreversible. White coagulations occurred in the cuvettes when the enzyme at 0.3–0.4 mg/ml in phosphate-buffered saline (pH 7.6) was heated to 60°C for 30 min (data not shown). This result was also observed in other studies (24,25). Thermal unfolding is thus not suitable for further analysis of the unfolding process of 3CLpro.

The fluorescence emission spectrum of the native enzyme shows a maximum at 328 nm (Fig. 2 *A*, curve *a*). The enzyme was very sensitive to GdnCl. In the presence of 6 M of GdnCl, the enzyme was completely unfolded in 10 min and the maximum emission fluorescence was shifted to 358 nm (Fig. 2 *A*, curve *c*). The unfolding was found to be completely reversible. A simple 10-fold dilution instantaneously induced refolding of the enzyme to a yield of 92–95% as judged by the fluorescence spectrum recovery (Fig. 2 *A*, curve *d*).

Similar results were obtained for the C-terminal domain-III-truncated enzyme ($3CL_{I+II}$) in which only one buried tryptophanyl residue remained and thus very little fluorescence was detected in the folded state (Fig. 2 *B*, curve *a*). The GdnCl-induced unfolding is thus very suitable for conformational stability studies of the enzyme.

Unfolding curves of SARS-CoV main protease in various concentrations of GdnCl

Unfolding of the enzyme in different concentrations of GdnCl is shown in Fig. 3. When the integrated fluorescence area was calculated at various GdnCl concentrations, a

monophasic unfolding process seems sufficient for describing the data (Fig. 3 *C*). However, there is a strong denaturant-dependent pretransition slope, which might indicate exposure of hydrophobic areas (26). Similar results were obtained with CD spectra (Fig. 3 *D*). More interesting results are shown in Fig. 3 *B*, which shows clearly a biphasic red shifting of the fluorescence emission peak from 328 to 358 nm upon addition of GdnCl. The average emission wavelength (Fig. 3 *A*) that registered changes in both fluorescence wavelength and fluorescence intensity was used to calculate the thermodynamic parameters of the unfolding/refolding process (Table 2).

The folded enzyme started to unfold beyond ~0.4 M GdnCl and reached an unfolding intermediate with $[GdnCl]_{0.5,\text{N-I}}$ at 1.1 M. The second unfolding phase started at ~3 M denaturant with $[GdnCl]_{0.5,\text{I-U}} = 4.0$ M, corresponding to free energy change (ΔG) of 1.22 ± 0.16 kcal/mol, and 10.88 ± 0.27 kcal/mol, respectively, for the $N \leftrightarrow I$ and $I \leftrightarrow U$ processes.

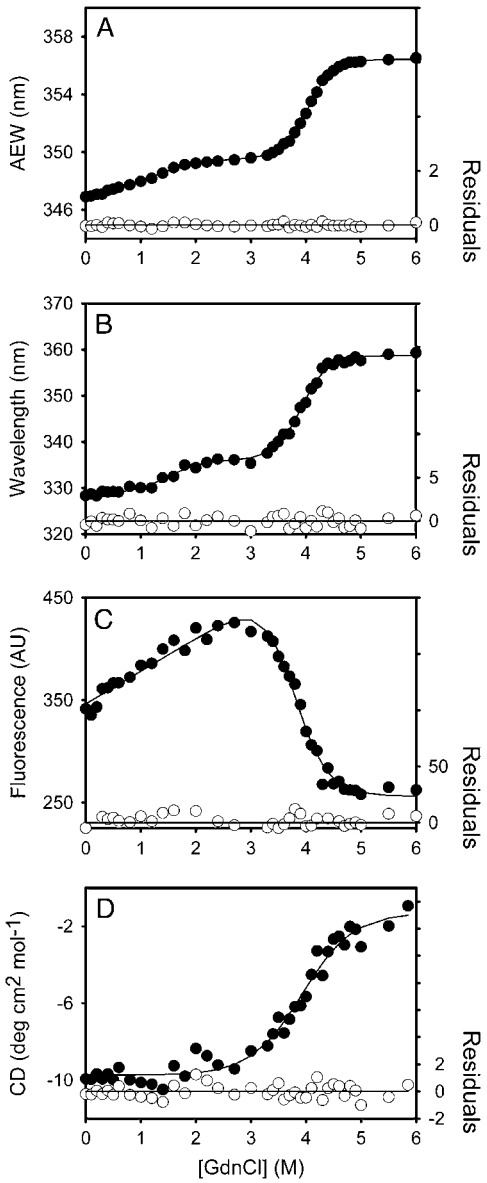

FIGURE 3 Unfolding curves of the full-length (3CLpro) SARS-CoV main protease in different concentrations of GdnCl. The enzyme, 3CLpro, 5 μg/ml (*A–C*) and 0.35 mg/ml (*D*) in 30 mM Tris-Cl (pH 7.7) was mixed with various concentrations of GdnCl. The fluorescence (*A–C*) or CD spectrum (*D*) was recorded after 10 min incubation at 30°C. (*A*) Average fluorescence emission wavelength (*AEW*) (Eq. 2). (*B*) Fluorescence emission wavelength shift. (*C*) Integrated fluorescence area. (*D*) Circular dichroism changes at 222 nm versus GdnCl concentration. Solid symbols are experimental data and the smooth curves are computer-generated best-fitted lines according to a three-state unfolding model (*A* and *B*) (Eq. 8) or a two-state unfolding model (*C* and *D*) (Eq. 7). The fitting residuals are also shown under each panel (*open circles*).

Quaternary structural changes of SARS-CoV main protease in the presence of GdnCl

The above results indicate clearly a biphasic unfolding curve for the dimeric enzyme. It is ultimately important to differentiate between dissociation versus unfolding under GdnCl. Analytical ultracentrifugation was used to examine this point. Fig. 4 *A* shows a typical sedimentation velocity profile for the enzyme in the presence of 0.4 M GdnCl. As shown in Fig. 4, *B* and *C*, the enzyme already dissociated to monomers at 0.4 M GdnCl. The major species of 4.6 ± 0.2 S corresponds to a dimer in the absence of GdnCl (Fig. 4 *C*), whereas in 0.4 M GdnCl the major peak corresponds to a monomer with a value of 2.9 ± 0.2 S. An equivalent continuous distribution $C(s,M)$ plot (not shown) indicated the dissociation of dimeric 3CLpro with molecular mass of 61 ± 17 kDa in buffer to a monomer with molecular mass of 33 ± 6 kDa in 0.4 M GdnCl. The biphasic-unfolding phenomenon is thus not a dissociation-unfolding process. Fig. 3, *A* and *B*, then represents a folding/unfolding process of the dissociated monomer.

Correlation of quaternary structural changes of SARS-CoV main protease with enzymatic activity

We then examined the enzymatic activity change upon GdnCl unfolding. The results shown in Fig. 5 indicate the close correlation of enzymatic activity loss with dimer dissociation. It should be noted that the protein concentrations were different in these experiments. The amount of enzyme used in activity assay (5 μg/ml) is much smaller than that used in AUC experiment (0.5–0.85 mg/ml). However, this still suggests that dissociation induces enzyme inactivation.

Possibility of unfolding intermediate as a molten globule state

After dissociation, the conformational changes of 3CLpro were clearly demonstrated by the increase in frictional ratio of the enzyme upon addition of GdnCl (Fig. 5). Analytical ultracentrifugation provides very detailed conformational information about the enzyme structural changes. Significant unfolded intermediate (15%) with frictional ratio of 2.67 was detected at 2 M GdnCl. This unfolded intermediate form was

TABLE 2 Stability parameters of SARS-CoV full-length and C-terminal truncated main proteases

Protease	$\Delta G_{N\text{-}I}$ (kcal mol^{-1})	$m_{N\text{-}I}$ (kcal mol^{-1} M^{-1})	$[GdnCl]_{0.5,N\text{-}I}$ (M)	$\Delta G_{I\text{-}U}$ (kcal mol^{-1})	$m_{I\text{-}U}$ (kcal mol^{-1} M^{-1})	$[GdnCl]_{0.5,I\text{-}U}$ (M)
3CLpro*	1.22 ± 0.16	1.15 ± 0.11	1.1 ± 0.2	10.88 ± 0.27	2.70 ± 0.07	4.0 ± 0.1
W31†	1.39 ± 0.46	0.91 ± 0.22	1.5 ± 0.6	–	–	–
W207/W218†	–	–	–	10.25 ± 0.47	2.49 ± 0.12	4.1 ± 0.3
$3CL_{I+II}$ †	7.27 ± 0.35	4.94 ± 0.23	1.5 ± 0.1			

*The unfolding data were fitted to a three-state unfolding model according to Eq. 8.
†The unfolding data were fitted to a two-state unfolding model according to Eq. 7.

not detected at 0.4 M or 4 M GdnCl, in which the enzyme existed exclusively as folded monomers or unfolded monomers, respectively.

The above observed unfolding intermediate may represent a molten globule state with exposure of the interior hydrophobic area, which will have pronounced binding ability to ANS and have a higher propensity to aggregate (27). However, no precipitation was observed for 3CLpro at any of the GdnCl concentrations. Neither was polymerization detected by AUC. ANS binding ability of the GdnCl-unfolded enzyme remained unchanged (Supplementary Material, Fig. 4S). This probably excludes a gross conformational change of the enzyme and the major intermediate form detected at 1–4 M GdnCl may not represent a partially unfolded molten globule state. The remaining possibility is then differential domain stability. The unfolding of $3CL_{I+II}$ was then examined.

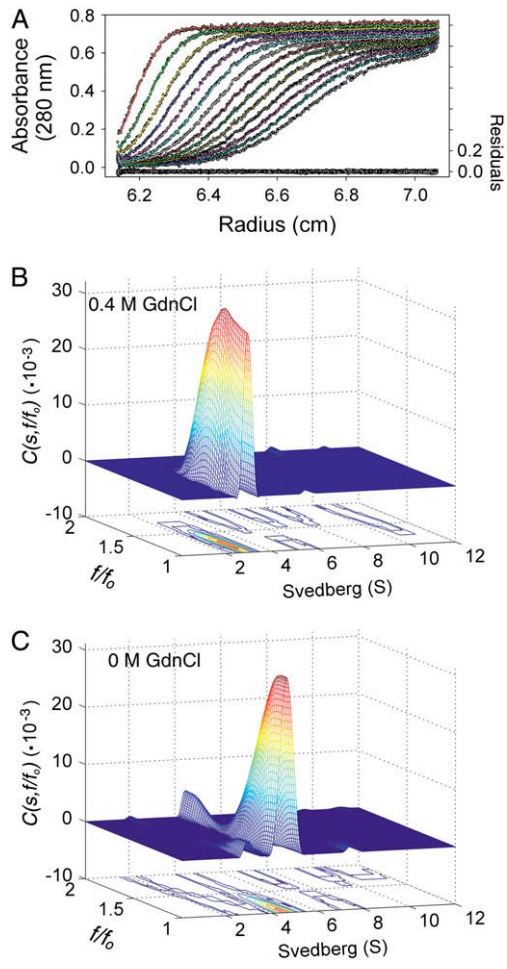

FIGURE 4 Quaternary structural changes of SARS-CoV main protease in GdnCl. Continuous sedimentation analysis of the SARS-CoV main protease in the presence and absence of GdnCl. (*A*) Sedimentation of the enzyme in 30 mM Tris-Cl (pH 7.7) containing 0.4 M GdnCl. The sedimentation velocity data were analyzed with SEDFIT (17). The fitting residuals are shown under the panel. (*B*) $C(s, f/f_o)$ distribution of the enzyme in 0.4 M GdnCl. (*C*) $C(s, f/f_o)$ distribution of the enzyme in the absence of GdnCl.

Unfolding curves of the C-terminal domain-truncated SARS-CoV main protease ($3CL_{I+II}$)

In the presence of GdnCl, $3CL_{I+II}$ follows the simple monophasic-unfolding curve (Fig. 6), which is also completely reversible (Fig. 2 *B*). Compared with the unfolding curve of the full-length 3CLpro, it is very likely that the biphasic unfolding of 3CLpro represents differential conformational stability of domain (I + II) and domain III.

This helps us to differentiate domain stability of 3CLpro and assign the functional role of domain III. However, it is still difficult to assign the corresponding domain to the biphasic unfolding. To provide a structural basis for the biphasic unfolding curves shown in Fig. 3, *A* and *B*, we constructed two mutants of 3CLpro, each with tryptophanyl residue(s) mutated at different domain and the fluorescence

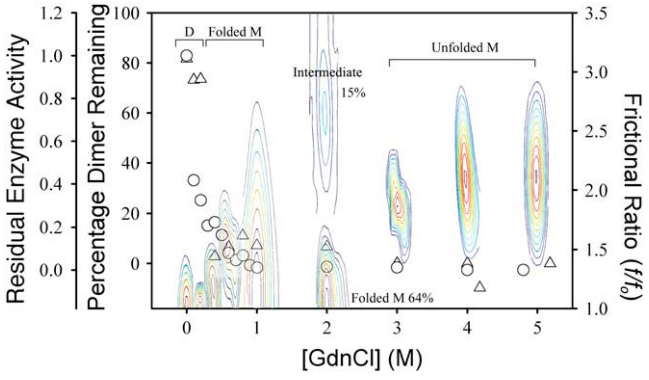

FIGURE 5 GdnCl-induced dissociation and unfolding of SARS-CoV main protease. (*Open circles*) Remaining enzyme activity; (*open triangles*) remaining dimer percentage. Unfolding of the enzyme was demonstrated by the contour of continuous sedimentation coefficient and frictional ratio $C(s, f/f_o)$ with the most concentrated region shown in red. The *x* axis of these contours has no physical meaning.

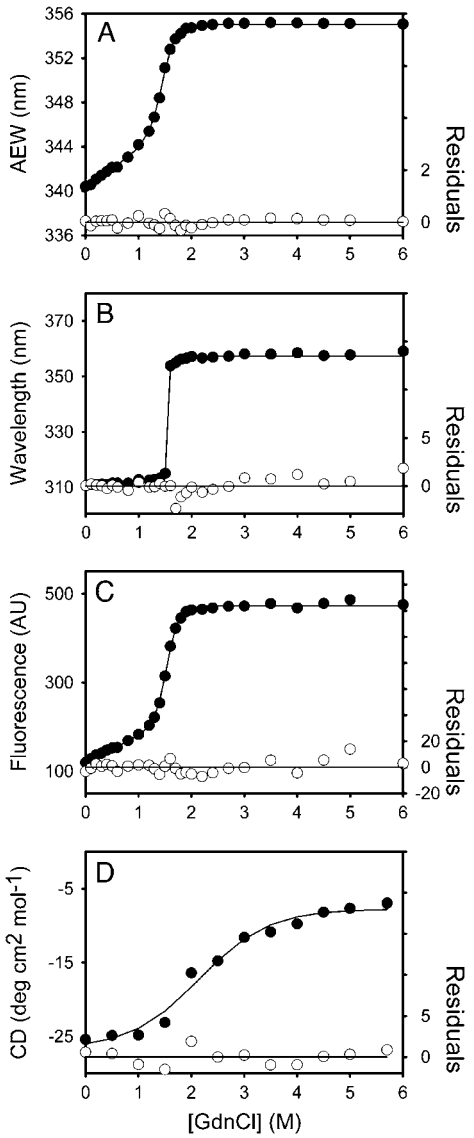

FIGURE 6 Unfolding curves of the C-terminal domain-truncated mutant ($3CL_{I+II}$) SARS-CoV main protease in different concentrations of GdnCl. The $3CL_{I+II}$ enzyme (13.3 μg/ml in panels *A–C*, and 0.18 mg/ml in panel *D*) in 30 mM Tris-Cl (pH 7.7) was mixed with various concentrations of GdnCl. The fluorescence (*A–C*) or CD spectrum (*D*) was recorded after 10 min incubation at 30°C. (*A*) Average fluorescence emission wavelength (*AEW*) (Eq. 2). (*B*) Fluorescence emission wavelength shift. (*C*) Integrated fluorescence area. (*D*) Circular dichroism changes at 222 nm versus GdnCl concentration. Solid symbols are experimental data and the smooth curves are computer-generated best-fitted lines according to a two-state unfolding model (Eq. 7). The fitting residuals are shown under each panel (*open circles*).

of these mutants now reports the structural changes of individual domain.

Unfolding curves of the tryptophanyl residue(s)-mutated SARS-CoV main protease

3CLpro has a total of three tryptophanyl residues (Trp^{31} at domain I, Trp^{207} and Trp^{218} at domain III, domain II has no Trp) (Fig. 1 *A*). W31, a double mutant with both Trp^{207} and Trp^{218} of domain III mutated to phenylalanine and only one tryptophanyl residue (Trp^{31}) remained. When excited with 295 nm UV light, the protein intrinsic emission fluorescence came from the tryptophanyl residue only. Thus W31 will only report the structural information of the N-terminal domain. Similarly, W207/W218, a single W31F mutant, will register the conformational change of the C-terminal domain.

Mutation at these tryptophanyl residues to phenylalanine did not change the kinetic properties of the enzyme. The k_{cat} and K_m values of W31 or W207/W218 were in the same order as in WT. The AUC data indicated that these Trp mutants still existed as dimer in solution (Table 1). When W31 and W207/W218 were unfolded in GdnCl, the fluorescence signal showed monophasic in both cases but distinct unfolding curves of these mutants were obtained (Fig. 7). It is clear that the N-terminal domain of 3CLpro (W31) is unfolded ($[GdnCl]_{0.5} \sim 1$ M) before the C-terminal domain (W207/W218) ($[GdnCl]_{0.5} \sim 4$ M). The unfolding parameters of $3CL_{I+II}$ thus represented the first phase of the 3CLpro (Table 2).

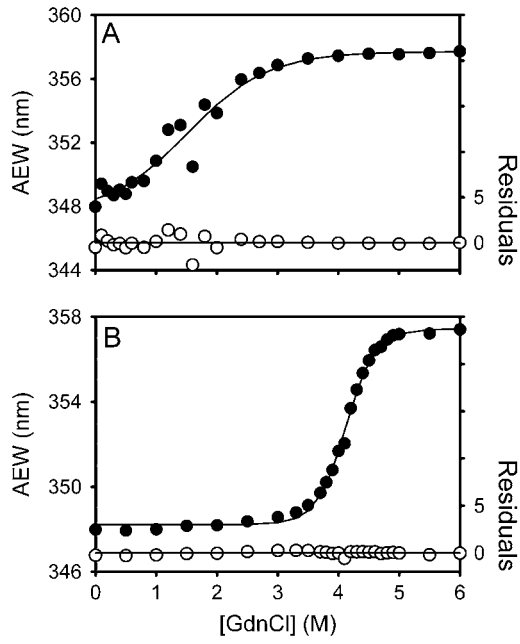

FIGURE 7 Unfolding curves of the tryptophanyl residue(s)-mutated SARS-CoV main protease in different concentrations of GdnCl. The mutant 3CLpro, W31 (40 μg/ml) (*A*) or W207/W218 (18 μg/ml) (*B*), in 30 mM Tris-Cl (pH 7.7) was mixed with various concentrations of GdnCl. The enzyme solution was excited at 295 nm, and the emission fluorescence was recorded at 30°C. Average fluorescence emission wavelength (*AEW*) (Eq. 2) of each mutant was used to present the unfolding data. (*A*) W31 (or more precisely W207F/W218F) represents the structural change of N-terminal domain. (*B*) W207/W218 (or more precisely W31F) represents the structural change of C-terminal domain. Solid symbols are experimental data and the smooth curves are computer-generated best-fitted lines according to a two-state unfolding model (Eq. 7). The fitting residuals are shown under each panel (*open circles*).

DISCUSSION

Some of the active site residues of SARS-CoV and related coronavirus main proteases have been explored previously. Thus Cys^{145} and His^{163} have been mutated in 3CLpro of transmissible gastroenteritis (corona) virus (9); His^{41}, Cys^{145}, Tyr^{161} in feline infectious peritonitis virus (28); and Cys^{145} in SARS-CoV (10,13,29). Table 1 summarizes the results obtained in this study. As expected, the almost complete loss of enzymatic activity of H41A and C145A confirmed the catalytic roles of the catalytic dyad (29,30). Similar observations were made for His^{163} and Tyr^{161}. Both of these residues are involved in substrate binding. The main function of Tyr^{161} is to hold His^{163} in place. Mutation of His^{163} led to a relatively small increase in K_m but a large drop in k_{cat} (Table 1). This could result from the enzyme activity assays being performed at pH 7.6, where one protomer of the dimeric enzyme adopts a different conformation from that in acidic environment (7). It has been postulated by Tan et al. (30) and also by Yang et al. (7) that at low pH, protonation of His^{163} induces a series of conformational changes that lead to the collapse of the oxyanion hole and blocks the S1 pocket entry. This would be in full agreement with the observation of a very large decrease in k_{cat} as shown in Table 1. An intriguing question will then be testing the pH-activity profile of H163A mutant to confirm that His^{163} is responsible for the lower activity of 3CLpro at pH 6.0 as predicted by molecular dynamic simulation (30).

The contribution of this study includes our presentation of a clear-cut dissociation-denaturation profile as well as a differential domain unfolding process. Protein folding constitutes one of the most challenging problems in the life science area (31–35). The SARS-CoV main protease has a unique structure with distinct structural domains. Each subunit of the protein has 13 strands and 11 helices, distributed in different structural domains (7). A multiphasic unfolding process may be anticipated. From the present analyses, the differential stability of different structural domains of the enzyme was deduced in terms of their free energy changes during transformation from the folded state to the unfolded state with an intermediate state involvement of folded domain III but unfolded domain (I+II) by the following reasoning.

Bacha et al. (24) had demonstrated the intrinsic instability of the C-terminal domain III by differential scanning calorimetry albeit this technique seemingly does not provide enough resolution to detect the unfolding intermediate. We observed a clear biphasic unfolding curve for the 3CLpro by the protein intrinsic fluorescence. The intermediate form was assigned to a folded C-terminal domain but an unfolded N-terminal domain on the basis of the following observations: *a*), Our unfolding data of 3CLpro clearly demonstrate the biphasic-unfolding process and our data of $3CL_{I+II}$ (N-terminal domain) suggest attributing the first and second phases of the 3CLpro to the N- and C-terminal domains, respectively (Table 2). *b*), Unequivocal assignment of the biphasic unfolding curve to differential domain stability came from results shown in Fig. 7. The W31 and W207/W218 are full-length 3CLpro. The unfolding curves shown in Fig. 7 should be biphasic. However, because the protein was excited with 295 nm UV light, one of the domain unfoldings is invisible and the unfolding curve only registered the structural changes of the tryptophan-containing domain. W31 and W207/W218 thus report conformational changes of the N-terminal and C-terminal domains, respectively. It is then clear that the N-terminal domain is labile than the C-terminal domain. The stabilization of the catalytic N-terminal domain by the C-terminal domain III is thus confirmed. The intrinsic conformational stability of domain (I+II) without domain III has a much higher free energy (7.27 kcal mol^{-1}) than full-length 3CLpro (1.22 kcal mol^{-1}) (Table 2). The significant larger m_{GdnCl} value of $3CL_{I+II}$ (4.94 kcal mol^{-1} M^{-1}) than full-length 3CLpro (1.15 kcal mol^{-1} M^{-1}) clearly indicates the pronounced increase of surface area exposure on denaturation of $3CL_{I+II}$ (26). This actually defines the functional role of domain III in SARS-CoV main protease.

After dissociation, the distribution of three structural forms of 3CLpro in various GdnCl concentrations is very clear. It includes a folded monomer, an intermediate with folded domain III but unfolded domain (I+II), and a completely unfolded form. The enzyme starts to dissociate at very early stage, at GdnCl concentration of <0.1 M. The dissociated monomer then starts to unfold at <0.4 M and reaches the first plateau at 2–3 M. The second unfolding then starts, which reaches another plateau at 4.5 M GdnCl. The complete dissociation and differential unfolding of the structural domains are summarized in Fig. 8. This schematic presentation

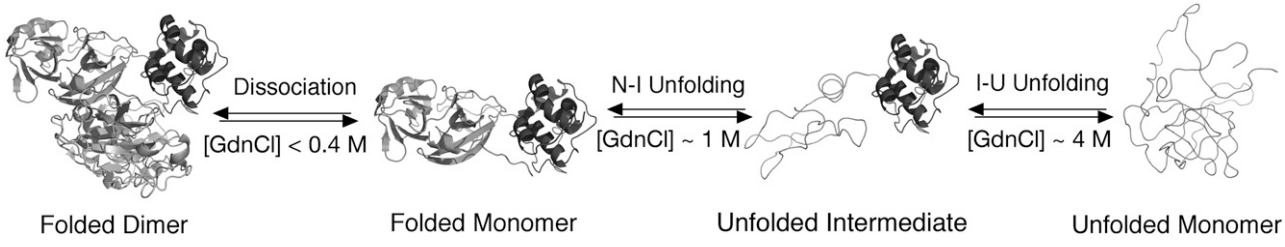

FIGURE 8 A schematic model for the dissociation and unfolding of SARS-CoV main protease in GdnCl. The stepwise dissociation and unfolding of the dimeric 3CLpro are shown schematically with N-terminal domain in red and C-terminal domain in blue. Unfolding portions of the polypeptide are shown in red thread.

only lists the major enzyme species of the complex folding process. The sharp unfolding curves for $3CL_{I+II}$ (Fig. 6) and the completely reversible of the unfolding process (Fig. 2 *B*) consistent with that $3CL_{I+II}$ without domain III can fold independently into an intact chymotrypsin-like fold. Domain III thus must have extra functional role.

Shi and Song (36) have observed a nanochannel between domain III and the active center and proposed a regulatory effect of the extra domain III in relaying the catalytic machinery. Chen et al. (21) suggest that the enzyme follows an association-activation-catalysis-dissociation catalytic cycle. This will provide further insight into the functional role of domain III in the dimerization and involvement in the enzyme activity regulation beyond simple stabilization of the catalytic N-terminal domain.

In summary, both $3CL_{I+II}$ and 3CLpro show a completely reversible unfolding/refolding process in GdnCl. We demonstrate here a clear-cut case of the differential stability of quaternary and domain structures. The unfolding studies of SARS-CoV main protease may shed light on the inherent structural stability of this important viral enzyme.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

The authors thank Ms. Wei-Hsin Chang and Mr. Meng-Ying Tsai for technical assistance.

This work was supported by National Science Council, Republic of China (NSC 94-2320-B-010-060).

REFERENCES

1. Chan, H. L., S. K. Tsui, and J. J. Sung. 2003. Coronavirus in severe acute respiratory syndrome (SARS). *Trends Mol. Med.* 9:323–325.
2. Matsuyama, S., M. Ujike, S. Morikawa, M. Tashiro, and F. Taguchi. 2005. Protease-mediated enhancement of severe acute respiratory syndrome coronavirus infection. *Proc. Natl. Acad. Sci. USA.* 102:12543–12547.
3. Yang, H., W. Xie, X. Xue, K. Yang, J. Ma, W. Liang, Q. Zhao, Z. Zhou, D. Pei, J. Ziebuhr, R. Hilgenfeld, K. Y. Yuen, et al. 2005. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol.* 3:1742–1752.
4. Wei, P., K. Fan, H. Chen, L. Ma, C. Huang, L. Tan, D. Xi, C. Li, Y. Liu, A. Cao, and L. Lai. 2006. The N-terminal octapeptide acts as a dimerization inhibitor of SARS coronavirus 3C-like proteinase. *Biochem. Biophys. Res. Commun.* 339:865–872.
5. Liang, P. H. 2006. Characterization and inhibition of SARS-coronavirus main protease. *Curr. Top. Med. Chem.* 6:361–376.
6. Tong, L. 2002. Viral proteases. *Chem. Rev.* 102:4609–4626.
7. Yang, H., M. Yang, Y. Ding, Y. Liu, Z. Lou, Z. Zhou, L. Sun, L. Mo, S. Ye, H. Pang, G. F. Gao, K. Anand, et al. 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.
8. Bartlam, M., H. Yang, and Z. Rao. 2005. Structural insights into SARS coronavirus proteins. *Curr. Opin. Struct. Biol.* 15:664–672.
9. Anand, K., G. J. Palm, J. R. Mesters, S. G. Siddell, J. Ziebuhr, and R. Hilgenfeld. 2002. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J.* 21:3213–3224.
10. Anand, K., J. Ziebuhr, P. Wadhwani, J. R. Mesters, and R. Hilgenfeld. 2003. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science.* 300:1763–1767.
11. Chou, C. Y., H. C. Chang, W. C. Hsu, T. Z. Lin, C. H. Lin, and G. G. Chang. 2004. Quaternary structure of the severe acute respiratory syndrome (SARS) coronavirus main protease. *Biochemistry.* 43:14958–14970.
12. Hsu, W. C., H. C. Chang, C. Y. Chou, P. J. Tsai, P. I. Lin, and G. G. Chang. 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.
13. Hsu, M. F., C. J. Kuo, K. T. Chang, H. C. Chang, C. C. Chou, T. P. Ko, H. L. Shr, G. G. Chang, A. H. Wang, and P. H. Liang. 2005. Mechanism of the maturation process of SARS-CoV 3CL protease. *J. Biol. Chem.* 280:31257–31266.
14. Braman, J., C. Papworth, and A. Greener. 1996. Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods Mol. Biol.* 57:31–44.
15. Royer, C. A., C. J. Mann, and C. R. Matthews. 1993. Resolution of the fluorescence equilibrium unfolding profile of Trp aporepressor using single tryptophan mutants. *Protein Sci.* 2:1844–1852.
16. Schuck, P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* 78:1606–1619.
17. Brown, P. H., and P. Schuck. 2006. Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys. J.* 90:4651–4661.
18. Lamm, O. 1929. Die differentialgleichung der ultrazentrifugierung. *Ark. Mat. Astr. Fys.* 21B:1–4. [in German].
19. Pace, C. N. 1990. Measuring and increasing protein stability. *Trends Biotechnol.* 8:93–98.
20. Morjana, N. A., B. J. McKeone, and H. F. Gilbert. 1993. Guanidine hydrochloride stabilization of a partially unfolded intermediate during the reversible denaturation of protein disulfide isomerase. *Proc. Natl. Acad. Sci. USA.* 90:2107–2111.
21. Chen, H., P. Wei, C. Huang, L. Tan, Y. Liu, and L. Lai. 2006. Only one protomer is active in the dimer of SARS 3C-like proteinase. *J. Biol. Chem.* 281:13894–13898.
22. Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.
23. Lee, T. W., M. M. Cherney, C. Huitema, J. Liu, K. E. James, J. C. Powers, L. D. Eltis, and M. N. James. 2005. Crystal structures of the main peptidase from the SARS coronavirus inhibited by a substrate-like aza-peptide epoxide. *J. Mol. Biol.* 353:1137–1151.
24. Bacha, U., J. Barrila, A. Velazquez-Campoy, S. A. Leavitt, and E. Freire. 2004. Identification of novel inhibitors of the SARS coronavirus main protease 3CLpro. *Biochemistry.* 43:4906–4912.
25. Shi, J., Z. Wei, and J. Song. 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.
26. Shortle, D. 1995. Staphylococcal nuclease: a showcase of m-value effects. *Adv. Protein Chem.* 46:217–247.
27. Fink, A. L. 1995. Molten globules. *Methods Mol. Biol.* 40:343–360.
28. Hegyi, A., A. Friebe, A. E. Gorbalenya, and J. Ziebuhr. 2003. Mutational analysis of the active centre of coronavirus 3C-like proteases. *J. Gen. Virol.* 83:581–593.
29. Huang, C., P. Wei, K. Fan, Y. Liu, and L. Lai. 2004. 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry.* 43:4568–4574.

30. Tan, J., K. H. Verschueren, K. Anand, J. Shen, M. Yang, Y. Xu, Z. Rao, J. Bigalke, B. Heisen, J. R. Mesters, K. Chen, X. Shen, et al. 2005. pH-dependent conformational flexibility of the SARS-CoV main proteinase (M(pro)) dimer: molecular dynamics simulations and multiple X-ray structure analyses. *J. Mol. Biol.* 354:25–40.

31. Dill, K. A., and H. S. Chan. 1997. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4:10–19.

32. Jahn, T. R., and S. E. Radford. 2005. The Yin and Yang of protein folding. *FEBS J.* 272:5962–5970.

33. Gruebele, M. 2005. Downhill protein folding: evolution meets physics. *C. R. Biol.* 328:701–712.

34. Lindorff-Larsen, K., P. Rogen, E. Paci, M. Vendruscolo, and C. M. Dobson. 2005. Protein folding and the organization of the protein topology universe. *Trends Biochem. Sci.* 30:13–19.

35. Oliveberg, M., and E. I. Shakhnovich. 2006. Folding and binding: the conformational repertoire of proteins: folding, aggregation and structural recognition. *Curr. Opin. Struct. Biol.* 16:68–70.

36. Shi, J., and J. Song. 2006. The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain. *FEBS J.* 273: 1035–1045.

37. DeLano, W. L. 2002. The Pymol Molecular Graphic System. DeLano Scientific, San Carlos, CA.