JOURNAL OF VIROLOGY, May 2008, p. 4227–4234 Vol. 82, No. 9
0022-538X/08/$08.00+0 doi:10.1128/JVI.02612-07

Without Its N-Finger, the Main Protease of Severe Acute Respiratory Syndrome Coronavirus Can Form a Novel Dimer through Its C-Terminal Domain▽

Nan Zhong,[1,2] Shengnan Zhang,[1,2] Peng Zou,[1,2] Jiaxuan Chen,[1,3] Xue Kang,[1,2] Zhe Li,[1,3] Chao Liang,[1] Changwen Jin,[1,2,3] and Bin Xia[1,2,3]*

Beijing Nuclear Magnetic Resonance Center,[1] College of Chemistry and Molecular Engineering,[2] and College of Life Science,[3] Peking University, Beijing 100871, China

Received 7 December 2007/Accepted 18 February 2008

The main protease (M^{pro}) of severe acute respiratory syndrome coronavirus (SARS-CoV) plays an essential role in the extensive proteolytic processing of the viral polyproteins (pp1a and pp1ab), and it is an important target for anti-SARS drug development. It was found that SARS-CoV M^{pro} exists in solution as an equilibrium of both monomeric and dimeric forms, and the dimeric form is the enzymatically active form. However, the mechanism of SARS-CoV M^{pro} dimerization, especially the roles of its N-terminal seven residues (N-finger) and its unique C-terminal domain in the dimerization, remain unclear. Here we report that the SARS-CoV M^{pro} C-terminal domain alone (residues 187 to 306; M^{pro}-C) is produced in *Escherichia coli* in both monomeric and dimeric forms, and no exchange could be observed between them at room temperature. The M^{pro}-C dimer has a novel dimerization interface. Meanwhile, the N-finger deletion mutant of SARS-CoV M^{pro} also exists as both a stable monomer and a stable dimer, and the dimer is formed through the same C-terminal-domain interaction as that in the M^{pro}-C dimer. However, no C-terminal domain-mediated dimerization form can be detected for wild-type SARS-CoV M^{pro}. Our study results help to clarify previously published controversial claims about the role of the N-finger in SARS-CoV M^{pro} dimerization. Apparently, without the N-finger, SARS-CoV M^{pro} can no longer retain the active dimer structure; instead, it can form a new type of dimer which is inactive. Therefore, the N-finger of SARS-CoV M^{pro} is not only critical for its dimerization but also essential for the enzyme to form the enzymatically active dimer.

A novel coronavirus (CoV) was identified as the etiological agent of the highly epidemic severe acute respiratory syndrome (SARS), which has infected more than 8,400 people, with a high fatality rate of about 10% (3, 14, 16, 25). SARS-CoV is a positive-sense, single-stranded RNA virus. The genome of the virus encodes two overlapping polyproteins, pp1a (486 kDa) and pp1ab (790 kDa), which mediate viral replication and transcription (17, 19, 20). The main protease (M^{pro}) of SARS-CoV, also named 3C-like protease, plays an important role in the extensive proteolytic processing of the viral polyproteins pp1a and pp1ab, which makes it essential for the viral life cycle and represents an attractive target for antiviral agent development (2, 29, 30).

The first crystal structure of SARS-CoV M^{pro} was solved in 2003, and the enzyme is a symmetric homodimer with a fold similar to that of the porcine transmissible gastroenteritis virus (TGEV) M^{pro} (1, 30). The N-terminal domain (residues 1 to 184) of SARS-CoV M^{pro} has a chymotrypsin-like fold, and the C-terminal domain (residues 201 to 303) has a globular fold containing five α-helices (30).

It was reported that SARS-CoV M^{pro} exists in solution as an equilibrium between monomeric and dimeric forms (10), and only the dimeric form of SARS-CoV M^{pro} is active (9). In the crystal structure of SARS-CoV M^{pro}, the N-terminal residues 1 to 7 (N-finger) of each protomer are squeezed in between two protomers and make contacts with both the N-terminal and C-terminal domains of the other protomer, and these contacts are important for dimerization (30). This dimerization pattern is similar to that of the TGEV M^{pro}, in which the role of the N -finger in the dimerization has been analyzed in detail (1). However, previous studies of the N-finger deletion mutants of M^{pro} gave different views of the role of the N-finger in SARS-CoV M^{pro} dimerization. Hsu et al. stated that the N-finger, especially residue R4, is indispensable for the dimerization and enzymatic activity of SARS-CoV M^{pro}, and the monomeric form becomes the predominant form after deletion of the N-terminal four to seven residues (6, 11). On the other hand, Chen et al. concluded that the N-finger is not crucial for the dimerization of SARS-CoV M^{pro} but is fundamental only to the enzymatic activity, and they found that the N-finger deletion mutant and wild-type (WT) SARS-CoV M^{pro} have similar dissociation constants for the dimerization (4). Recently Wei et al. reported that the N-finger deletion mutant of SARS-CoV M^{pro} could not dimerize at all (24). These controversial results turn the role of the N-finger in SARS-CoV M^{pro} dimerization into a mystery.

Meanwhile, Shi et al. proposed that the C-terminal domain plays a critical role in SARS-CoV M^{pro} dimerization based on the observation that the N-terminal domain alone is a monomer and the C-terminal domain alone is only a dimer (22).

* Corresponding author. Mailing address: Beijing Nuclear Magnetic Resonance Center, Peking University, Beijing 100871, P. R. China. Phone: 86-10-6275-8127. Fax: 86-10-6275-3790. E-mail: binxia@pku.edu.cn.

▽ Published ahead of print on 27 February 2008.

Nevertheless, in the crystal structure of SARS-CoV M^{pro}, there is almost no direct contact between the two C-terminal domains of the homodimer. Therefore, it is not obvious how the dimerization of the C-terminal domain is related to the dimerization of SARS-CoV M^{pro}.

In order to clarify the controversial issues mentioned above, we have reinvestigated the dimerization of SARS-CoV M^{pro}. We found that the SARS-CoV M^{pro} C-terminal domain alone (M^{pro}-C) exists as a stable monomer and a stable dimer simultaneously. There is no obvious conversion between the two forms. The dimerization interface of the M^{pro}-C dimer is novel and is unrelated to that of SARS-CoV M^{pro} in the crystal structure. Without the N-finger, SARS-CoV M^{pro} can also form a stable dimer due to the dimerization of its C-terminal domain.

MATERIALS AND METHODS

Construction of expression plasmids. For WT SARS-CoV M^{pro}, the DNA fragment encoding residues 1 to 306 was cloned into the pET21a vector and an NdeI restriction site within the coding sequence was removed by changing the codon for H164 from CAT to CAC. A hexahistidine tag (sequence, LEHHHHHH) was engineered at the carboxyl terminus of the protein. For the N-terminal domain of SARS-CoV M^{pro} (M^{pro}-N), the DNA fragment encoding residues 1 to 199 was cloned into the pET21a vector. For the C-terminal domain (M^{pro}-C), the DNA fragment encoding residues 187 to 306 was cloned into the pET21a vector. For the N-finger deletion mutant of SARS-CoV M^{pro} (M^{pro}-Δ7), the DNA fragment encoding residues 8 to 306 was cloned into pET28a with a hexahistidine tag (sequence, LEHHHHHH) attached at the carboxyl terminus.

Protein production and purification. The proteins were all produced in the *Escherichia coli Rosetta* (DE3) *pLysS* strain. The proteins with/without a C-terminal hexahistidine tag were purified using nickel-nitrilotriacetic acid/ion-exchange chromatography and followed by gel filtration (Superdex 75 column) on an ÁKTA fast-protein liquid chromatography system (FPLC) (GE).

Cross-linking experiment. Protein (0.1 mM) was cross-linked with 3 mM ethylene glycolbis(succinimidylsuccinate) (EGS) (Pierce, Rockford, IL) in the reaction buffer (0.1 M potassium phosphate, 0.15 M NaCl, pH 7.2). The reaction mixture was incubated at room temperature for 30 min, and then the reaction was quenched by adding Tris (1 M, pH 7.5) to a final concentration of 50 mM.

Gel filtration analysis. Protein oligomerization was analyzed using a home-packed 16/70 Superdex 75 HR gel filtration column on an ÁKTA FPLC. All protein samples were in 50 mM potassium phosphate buffer (pH 8.0) with 1 mM 1,4-dithiothreitol (DTT). To estimate the apparent molecular mass based on the retention volume, three proteins, myoglobin (17.0 kDa, 88.0 ml), egg albumin (42.7 kDa, 73.4 ml), and bovine albumin V (68.0 kDa, 67.2 ml), were used as the molecular mass standard. A standard calibration curve was obtained by plotting the ratio $(Ve\text{-}V_0)/(Vt\text{-}V_0)$ against the logarithm of molecular mass (Ve is elution volume, V_0 is the void volume, and Vt is the total bed volume).

Enzymatic activity assay. The enzymatic activities of WT SARS-CoV M^{pro} and the mutant M^{pro}-Δ7 were measured using a fluorogenic peptide, MCA-AVLQSGFR-Lys(Dnp)-Lys-NH_2 [more than 95% purity; GL Biochem (Shanghai) Ltd.] as the substrate. The fluorescence intensity was monitored using a Hitachi (Tokyo, Japan) F-4500 fluorescence spectrophotometer with wavelengths of 320 and 405 nm for excitation and emission, respectively. The reaction buffer consisted of 50 mM Tris-HCl (pH 7.3), 1 mM EDTA, and 1 mM DTT. The working concentrations of the protease and the substrate were 1 μM and 40 μM, respectively (29).

NMR spectroscopy. Nuclear magnetic resonance (NMR) samples of uniformly ^{15}N-labeled, $^{15}N/^{13}C$-labeled, and $^{2}H/^{15}N/^{13}C$-labeled M^{pro}-C and the uniformly $^{2}H/^{15}N$-labeled M^{pro}-Δ7 dimer and monomer were prepared. All NMR samples were at a concentration of about 1 mM and were prepared in buffer containing 50 mM potassium phosphate (pH 7.0), 1 mM EDTA, and 0.03% NaN_3 in 90% H_2O–10% D_2O, plus Complete, an EDTA-free protease inhibitor cocktail (Roche, Germany). All NMR experiments were performed at 298 K on a Bruker Avance 500-MHz (with cryoprobe) or 600-MHz NMR spectrometer. Backbone chemical shift assignments were based on a two-dimensional (2D) $^{1}H\text{-}^{15}N$ heteronuclear single-quantum coherence spectrum and three-dimensional HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, HNCO, and HN(CA)CO experiment data (21). All NMR spectra were processed with the NMRPipe software program (7) and analyzed using NMRView software (12). The chemical shift in the ^{1}H dimension was referenced directly to 2,2-dimethyl-2-silapentanesulfonic acid (DSS), whereas the chemical shifts in the ^{13}C and ^{15}N dimensions were indirectly referenced to DSS (26).

Dimer structure modeling. The model of the M^{pro}-C dimer was calculated using the protein–protein docking program HADDOCK (8). The docking was initiated from the C-terminal domain part of the SARS-CoV M^{pro} crystal structure (30). The active residues were defined based on the chemical shift perturbation data, in which residues with the combined NH chemical shift difference between two forms exceeding 0.10 ppm (average + 0.5 × standard deviation) were selected. They were residues R217 to T225, A260, D263, C265 to A267, K269 to L271, and L282. The passive residues were defined as all other surface-accessible residues (residues with more than 55% solvent-accessible surface area, determined using the MOLMOL software program [13]). The ambiguous interaction restraints were defined between the active residues of one protomer and all the active and passive residues of the other protomer. The active residues were set as flexible segments, and the passive residues ± 2 sequential residues were set as semiflexible segments during the calculation.

RESULTS

Two oligomerization states of M^{pro}-C. We have produced the N-terminal domain alone (residues 1 to 199; M^{pro}-N) and the C-terminal domain alone (residues 187 to 306; M^{pro}-C) of SARS-CoV M^{pro} in *E. coli*. In agreement with the previous report, M^{pro}-N behaved as a monomer on the gel filtration column (retention volume, 86.3 ml; apparent molecular mass, 17.7 kDa), and the retention volume did not show significant concentration dependence (data not shown) (22).

Interestingly, we found that M^{pro}-C was produced in *E. coli* in two forms which could be separated by gel filtration, and the retention volumes of the two forms were 74.6 ml and 86.4 ml, respectively (Fig. 1A). The apparent molecular masses calculated based on their retention volumes are 38.2 kDa for the 74.6-ml fraction and 17.5 kDa for the 86.4-ml fraction. On a sodium dodecyl sulfate (SDS)-polyacrylamide gel, both fractions appeared at the same position under either reducing or nonreducing conditions, with an apparent molecular mass of ~13 kDa (Fig. 1B, lanes M2, M3, D2, and D3). Mass spectrometry analysis also confirmed that the molecular masses of both fractions were the same as the theoretical value (13.4 kDa) for the M^{pro}-C monomer. After treatment with the cross-linking agent EGS, the 86.4-ml fraction appeared at the same position as the untreated sample in SDS-polyacrylamide gel electrophoresis (PAGE), while the 74.6-ml fraction showed two bands. One band was at the same position as the untreated sample, and the other was at the position of ~28 kDa (Fig. 1B, lane M1 and D1). These data suggest that M^{pro}-C is produced in *E. coli* not only in the dimeric form (74.6-ml fraction), as described by Shi et al. (22), but also in a monomeric form (86.4-ml fraction). Meanwhile, the dimerization of M^{pro}-C is noncovalent, and no disulfide bond is involved, even though M^{pro}-C has two free cysteine residues. Surprisingly, we found that the two forms of M^{pro}-C are stable, and there is no obvious conversion between the monomeric and dimeric forms for days at room temperature, which was monitored by gel filtration analysis (Fig. 1A).

Novel interface for the M^{pro}-C dimer. Carefully examining all the crystal structures of SARS-CoV M^{pro} (15, 23, 27, 28, 30), we found that there is almost no direct contact between the C-terminal domains of the two protomers. Only the side

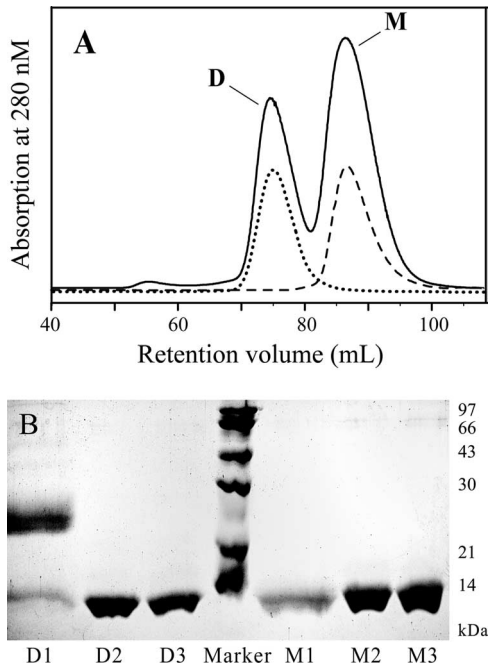

FIG. 1. (A) Elution profile of M^{pro}-C from gel filtration analysis. The solid line is for the purification of the M^{pro}-C protein, and the elution peaks for the monomeric (M) and dimeric (D) forms are indicated. The broken line and dotted line are for the purified M^{pro}-C monomeric and dimeric protein samples after 3 days at room temperature, respectively. (B) SDS-PAGE analysis of M^{pro}-C. Lanes M2 and M3 are the M^{pro}-C monomer with/without 10 mM DTT, respectively; lanes D2 and D3 are the M^{pro}-C dimer with/without 10 mM DTT, respectively; lanes M1 and D1 are the M^{pro}-C monomeric and dimeric forms treated with cross-linking agent EGS; and the center lane is the molecular mass marker.

chains of the C-terminal domain residues T285 and I286 from each protomer are closer than 5 Å in all of these structures. In the structure of TGEV M^{pro}, which has tertiary and quaternary structures similar to those of SARS-CoV M^{pro}, two hydrogen bonds are found between the C-terminal domains of the two protomers. However, it was suggested that the interactions between the C-terminal domains of TGEV M^{pro} appear to be a consequence rather than the cause of the dimerization (1). Thus, the current available structural information cannot elucidate why the SARS-CoV M^{pro} C-terminal domain alone can form a stable dimer. To determine the dimerization interface of the M^{pro}-C dimer, we carried out the backbone NMR resonance assignments for both the monomeric and dimeric forms of M^{pro}-C. For the M^{pro}-C monomer, nearly all backbone NH chemical shift assignments were obtained, with the exception of residues F219 and E288, whose NH signals were missing. For the M^{pro}-C dimer, backbone NH signals for residues F219, R222, F223, and E288 were missing in the 2D ^{1}H-^{15}N HSQC spectrum, while all the other NH signals have been assigned.

The missing NH signals probably resulted from intermediate-time-scale conformational exchange, which causes broadening of the NMR signals beyond detection. Since the chemical shifts of NH signals are sensitive to the local chemical environment change for individual NH group, the dimerization would cause the chemical environment change for the residues at the dimer interface, thus resulting in the NH chemical shift changes for these residues. Therefore, the dimerization interface of the M^{pro}-C dimer can be identified from the comparison of the NH chemical shift differences between the monomeric and dimeric forms of M^{pro}-C.

Most of the NH peaks overlap well between the 2D ^{1}H-^{15}N HSQC spectra of the monomeric and dimeric forms of M^{pro}-C (Fig. 2A). As expected, some NH signals exhibit significant chemical shift differences between the two forms. Residues with a large combined NH chemical shift difference ($\Delta\delta_{comb}$) of >0.1 ppm include V212, R217-T225 (F219, R222, and F223 are missing), A260-Q273 (except V261, L262, M264, L268, and L272), and M276. Residues with a $\Delta\delta_{comb}$ value of less than 0.1 ppm but more than 0.05 ppm are the following: I200, A210, N228, F230, Y239, V261, L268, L272, T280, G283, and E290. In addition, the side chain NH_2 signals of residues N214, N221, N274, and N277 also show significant chemical shift differences between the monomeric and dimeric forms (Fig. 2B). Notably, all of the residues mentioned above are located at the loop consisting of residues R217 to T225 (D-loop) and a helix formed by residues A260 to Q273 (D-helix), which is right underneath the D-loop. Thus, the D-loop and D-helix should represent the dimerization interface for the M^{pro}-C dimer. Mapping this dimer interface onto the crystal structure of the SARS-CoV M^{pro} dimer (colored blue and red for different protomers in Fig. 3A), it is obvious that the dimerization interface of the M^{pro}-C dimer is not related to the dimerization interface in the crystal structure of SARS-CoV M^{pro}.

All residues on the D-loop show relatively larger $\Delta\delta_{comb}$ values than those on the D-helix (Fig. 2B). NH signals from residues R222 and F223 are observed for the monomeric form but are not detected for the dimeric form of M^{pro}-C, presumably due to a difference in the conformational exchange rates between the two forms. Also, the D-helix residues which are facing the D-loop have relatively larger $\Delta\delta_{comb}$ values than those on the opposite side (Fig. 2B). Furthermore, the $^{13}C^{\alpha}$ chemical shift differences of most residues on the D-helix are less than 0.2 ppm (within the ^{13}C chemical shift resolution) (Fig. 2C). Since the $^{13}C^{\alpha}$ chemical shift is sensitive to the secondary structure change, this should indicate that this helix does not undergo much conformational change after the dimerization. On the contrary, some of the residues on the D-loop show quite large $^{13}C^{\alpha}$ chemical shift differences (over 2 ppm) between the monomeric and dimeric forms, suggesting that this loop probably changes its conformation upon dimerization (Fig. 2C). This should imply that the NH and $^{13}C^{\alpha}$ chemical shift differences observed for residues on the D-helix are from a secondary effect of the D-loop conformation change due to the dimerization.

Based on the chemical shift perturbation data, a structure model of the M^{pro}-C dimer was calculated using the docking program Haddock. Comparing the final refined models of the dimer with the C-terminal domain structure of SARS-CoV M^{pro}, it seems that the dimerization is mainly due to the hydrophobic interaction between the residue F223 of one M^{pro}-C molecule and the residues W218, F219, and L271 of the other molecule and also possibly a few hydrogen bonds (Fig. 3B).

Dimerization of WT SARS-CoV M^{pro} and its N-finger deletion mutant. Since M^{pro}-C can form a stable dimer, we tried to

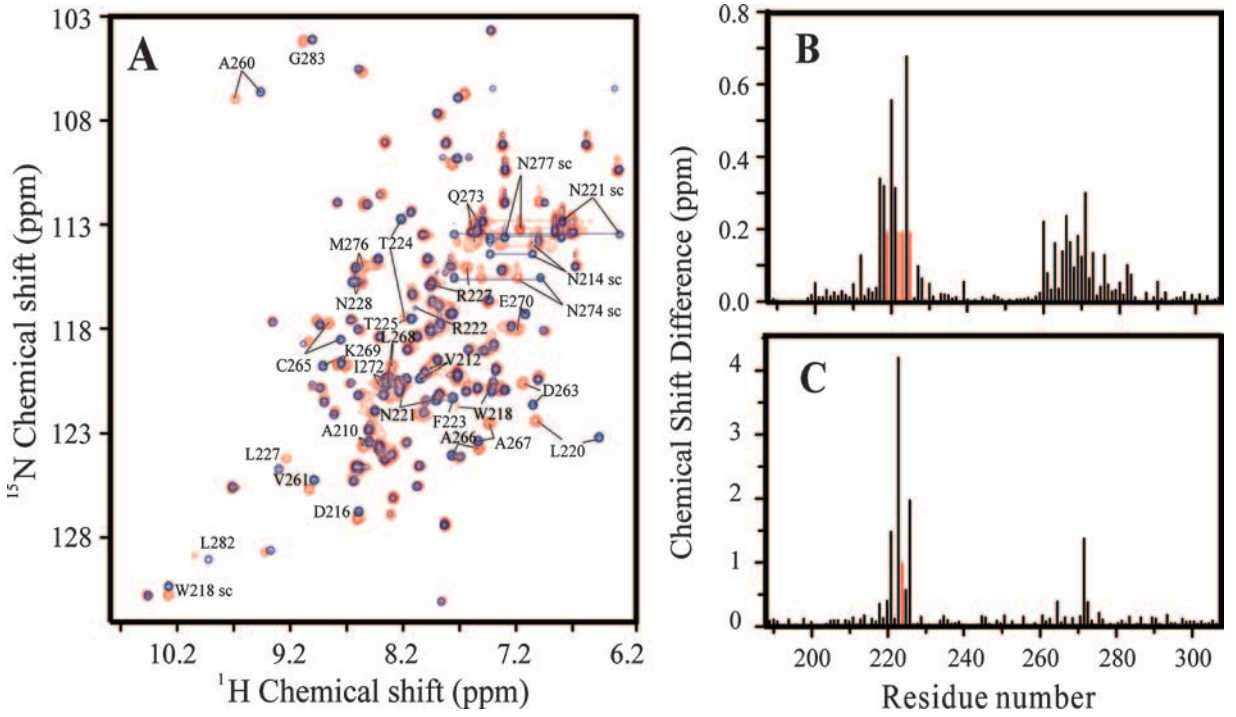

FIG. 2. (A) An overlay of the 2D ^{1}H-^{15}N HSQC spectra of the monomeric (blue) and dimeric (red) forms of M^{pro}-C. The peaks with combined NH chemical shift difference larger than 0.05 ppm are labeled with the one-letter amino acid code and residue number; "sc" is used to indicate the side chain signals. (B) Plot of combined NH chemical shift difference versus residue number. The combined chemical shift difference was calculated using the empirical equation $\Delta\delta_{comb} = [\Delta\delta_{HN}^2 + (\Delta\delta_N/6.5)^2]^{1/2}$, where $\Delta\delta_{HN}$ and $\Delta\delta_N$ represent the chemical shift differences of ^{1}H and ^{15}N, respectively (18). (C) Plot of $^{13}C^{\alpha}$ chemical shift difference versus residue number. Residues without assignment are indicated by short red bars.

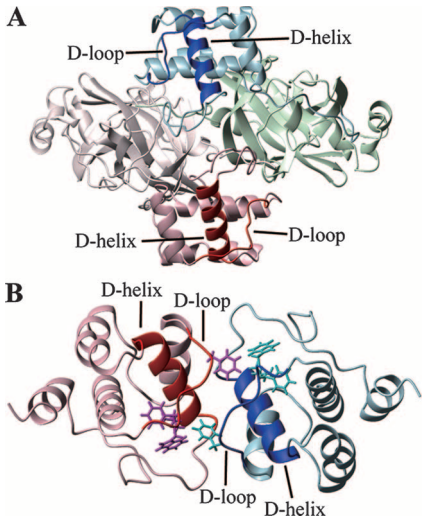

FIG. 3. (A) The dimerization interface of the M^{pro}-C dimer mapped on the crystal structure of SARS-CoV M^{pro}. A ribbon diagram of the crystal structure 1UK3 is shown. The C-terminal domain of one protomer is colored in light blue, and the residues with a $\Delta\delta_{comb}$ value of >0.05 ppm are colored in blue. The C-terminal domain of the other protomer is colored in pink, and the residues with a $\Delta\delta_{comb}$ value of >0.05 ppm are colored in red. (B) Ribbon diagram of a model structure for the M^{pro}-C dimer. The model structure was calculated using the software program Haddock. The dimer interfaces are colored in red and blue in two protomers, respectively. The side chains of residues W218, F219, F223, and L271, which may be important for dimer formation due to hydrophobic interactions at the dimer interface, are shown.

find out whether WT SARS-CoV M^{pro} has a similar C-terminal domain-mediated stable dimeric form which is different from the dimer of the crystal structure. Our study showed that WT SARS-CoV M^{pro} behaves as an equilibrium between the monomeric and dimeric forms on a gel filtration column, and no stable dimeric form could be detected. The retention volumes of WT SARS-CoV M^{pro} were concentration dependent on the gel filtration column: for samples with concentrations of 25, 9, 2, and 0.1 mg/ml, the corresponding retention volumes are 70.2, 70.8, 72.7, and 75.2 ml, respectively (Fig. 4A). At the concentration of 0.1 mg/ml, the estimated apparent molecular mass is 36.7 kDa, very close to the theoretical molecular mass (34.9 kDa) of monomeric WT SARS-CoV M^{pro}. Thus, it seems that WT SARS-CoV M^{pro} could not form a stable dimer through its C-terminal domain.

Interestingly, we found that the N-finger deletion mutant of SARS-CoV M^{pro} (M^{pro}-Δ7) is also produced in *E. coli* in two forms. M^{pro}-Δ7 is eluted from the gel filtration column as two fractions with retention volumes of 63.7 and 74.9 ml, respectively (Fig. 4A). The corresponding estimated apparent molecular masses are 37.4 kDa and 78.5 kDa, respectively. Both fractions run on an SDS-PAGE gel at the same position as monomeric M^{pro}-Δ7, and mass spectrometry analysis showed that both fractions have the same molecular mass as that of monomeric M^{pro}-Δ7 (data not shown). Thus, the 74.9-ml fraction should correspond to a monomeric form of M^{pro}-Δ7, and the 63.7-ml fraction should be a dimeric form. Similar to M^{pro}-C, both the dimeric and monomeric forms of M^{pro}-Δ7

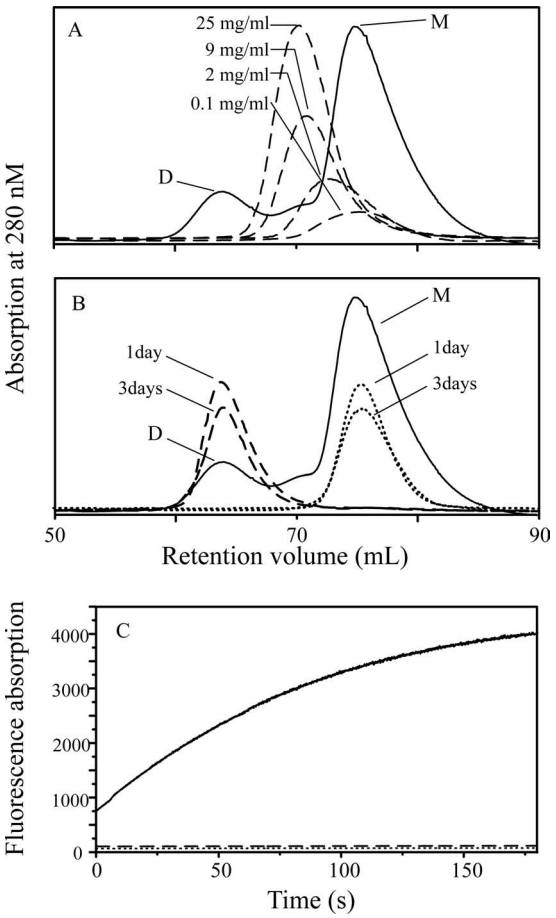

FIG. 4. (A) Gel filtration analysis of WT SARS-CoV M^{pro} and M^{pro}-Δ7. The broken lines represent WT SARS-CoV M^{pro} at the indicated concentrations. The peak heights have been adjusted arbitrarily to make the figure clearer. The solid line is the purification profile for M^{pro}-Δ7 with the elution peaks for the monomeric form (M) and dimeric form (D) marked. (B) Gel filtration analysis of M^{pro}-Δ7 stability. The solid line is the purification profile for M^{pro}-Δ7, with the peaks of the monomeric form (M) and dimeric form (D) indicated. The lines of purified M^{pro}-Δ7 after being placed at room temperature for 1 day and 3 days are indicated. The broken and dotted lines are for the monomeric and dimeric forms, respectively. The peak heights have been adjusted arbitrarily to make the figure clearer. (C) Enzymatic activity of WT SARS-CoV M^{pro} and M^{pro}-Δ7. The solid line is for WT SARS-CoV M^{pro}; the broken line is for the M^{pro}-Δ7 dimer; and the dotted line is for the M^{pro}-Δ7 monomer.

were stable and could not convert into each other for days at room temperature, as monitored by gel filtration analysis (Fig. 4B). The amount of the M^{pro}-Δ7 monomer produced in *E. coli* was normally 10 times higher than that of the M^{pro}-Δ7 dimer (Fig. 4A), and both forms of M^{pro}-Δ7 were almost inactive in the enzymatic assay (Fig. 4C).

Since M^{pro}-Δ7 and M^{pro}-C behave similarly in terms of oligomerization states and stability, we supposed that the dimerization of M^{pro}-Δ7 was also due to its C-terminal domain. To prove this, the 2D ^{1}H-^{15}N HSQC spectra of the uniformly $^{2}H/^{15}N$-labeled M^{pro}-Δ7 monomer and dimer were collected and compared. Most of the NH peaks in the spectra of these two forms overlap well, suggesting that most parts of the structures are the same between the monomeric and dimeric forms of M^{pro}-Δ7. Since the NH signals for residues on the D-loop and D-helix of M^{pro}-C have different chemical shifts between the monomeric and dimeric forms, the characteristic chemical shifts of these NH signals in each form can be used to distinguish the conformation of the D-loop and the D-helix between the two forms. Therefore, if the M^{pro}-Δ7 dimer is dimerized through its C-terminal domain in the same fashion as the M^{pro}-C dimer, we would expect to see the dimer interface residues in the M^{pro}-Δ7 dimer have the same NH chemical shifts as those in the M^{pro}-C dimer. Superimposing the 2D ^{1}H-^{15}N HSQC spectra of both the monomeric and dimeric forms of M^{pro}-Δ7 and M^{pro}-C, it was found that the NH signals of the D-loop and the D-helix residues R217, L220, L227, N228, V261, D263, A267, M276, L282, and G283, along with the side chain NH signal of W218 and the side chain NH_2 signals of N277, could be identified unambiguously, while the other unique NH signals from the D-loop and the D-helix could not be distinguished clearly due to signal overlap with the NH peaks from the N-terminal domain of M^{pro}-Δ7. It is clearly demonstrated in Fig. 5 that the above-mentioned NH peaks of dimeric M^{pro}-Δ7 (black peaks) overlap and only overlap with those of the dimeric M^{pro}-C (red peaks). Vice versa, those signature NH signals from monomeric M^{pro}-Δ7 (green peaks) coincide and only coincide with those of monomeric M^{pro}-C (blue peaks). These suggest that the D-loop and the D-helix in monomeric M^{pro}-Δ7 retain the same conformation as those in monomeric M^{pro}-C and the conformations of the D-loop and the D-helix in dimeric M^{pro}-Δ7 and dimeric M^{pro}-C are the same. Thus, the M^{pro}-Δ7 dimer should have the same dimerization interface as the M^{pro}-C dimer. Therefore, the N-finger deletion mutant of SARS-CoV M^{pro} also exists as a stable monomer and a stable dimer, and the dimer is formed through its C-terminal domain dimerization, which is different from the dimerization of WT SARS-CoV M^{pro} (Fig. 6).

DISCUSSION

Three research articles have been published concerning the dimerization of the N-finger deletion mutant of SARS-CoV M^{pro}, and inconsistent conclusions about the role of the N-finger in the dimerization have been given. None of them has reported the existence of a novel stable dimeric form of M^{pro}-Δ7. Hsu et al. (11) and Chen et al. (4) both reported that M^{pro}-Δ7 is in an equilibrium of monomeric and dimeric forms. From our point of view, since the mutant protein samples they studied were purified through a one-step nickel-nitrilotriacetic acid affinity column, the stable dimeric form was therefore not separated from the stable monomeric form in their samples, and the sample could be mistakenly treated as an equilibrium between the monomer and the dimer. Our results agree with those of Hsu et al. better in that the monomeric form is the major form for M^{pro}-Δ7, since the monomeric and dimeric M^{pro}-Δ7 proteins were normally produced at about a 10:1 ratio from *E. coli* in our current study (Fig. 4A). However, Chen et al. reported that the monomer/dimer ratio for M^{pro}-Δ7 is similar to that for WT SARS-CoV M^{pro}. From their paper, we found that both the WT and mutant proteins they studied have a 14-residue fusion tag with a sequence of "MRGSHHHHH HGSTM" at the N terminus of the protein sequences (4). It is possible that this tag may have an effect on the dimerization

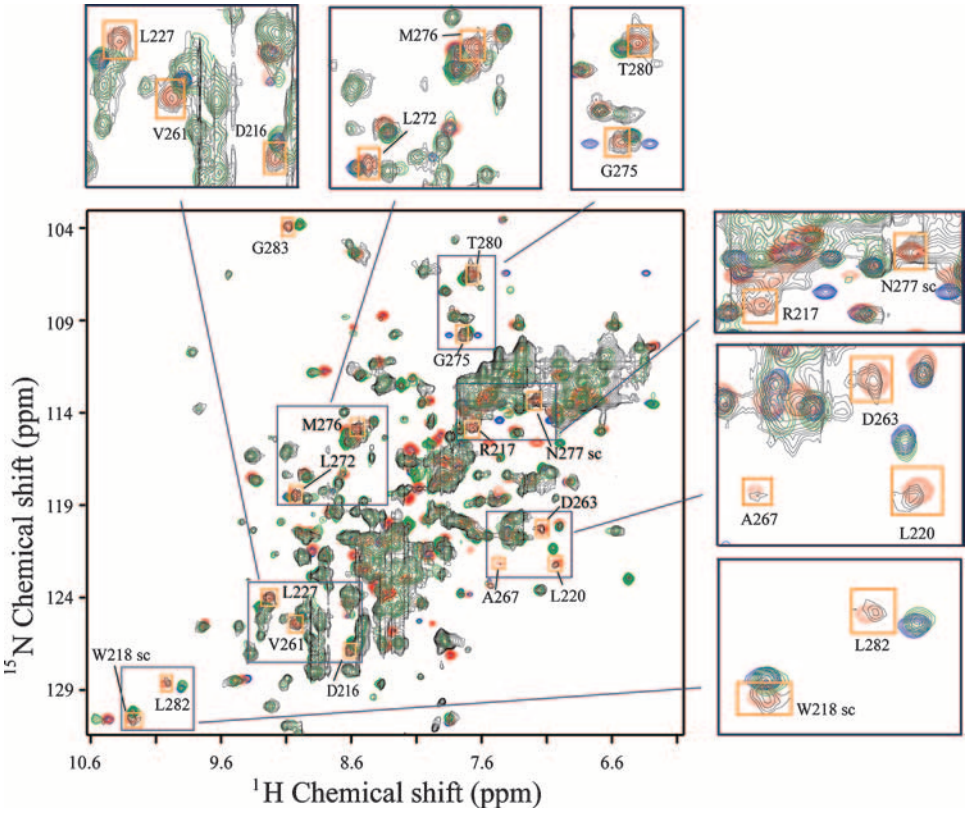

FIG. 5. An overlay of 2D ^{1}H-^{15}N HSQC spectra of monomeric and dimeric M^{pro}-Δ7 and M^{pro}-C. The black peaks belong to the M^{pro}-Δ7 dimer, the green peaks belong to the M^{pro}-Δ7 monomer, the red peaks are from the M^{pro}-C dimer, and the blue peaks are from the M^{pro}-C monomer. The signature NH peaks of the M^{pro}-C dimer are indicated by orange squares, which are labeled with a one-letter amino acid code and a residue number. Six areas of the spectra are enlarged and displayed for clarity.

and thus interfered with their study on the oligomerization of M^{pro}-Δ7. Chen et al. also presented a model structure for the M^{pro}-Δ7 dimer based on molecular dynamics simulation. Although there is difference in dimerization between this model structure and the crystal structure of WT SARS-CoV M^{pro}, the difference is not so significant and there is no sign of the C-terminal domain dimerization in their model structure (Fig. 6). The M^{pro}-Δ7 sample Wei et al. used was purified by an affinity column followed by gel filtration, and it is possible that the stable dimeric form was ignored during the later purification step (24). If that is the case, the sample they studied should be the same as our M^{pro}-Δ7 monomer sample, which cannot form a dimer from equilibrium.

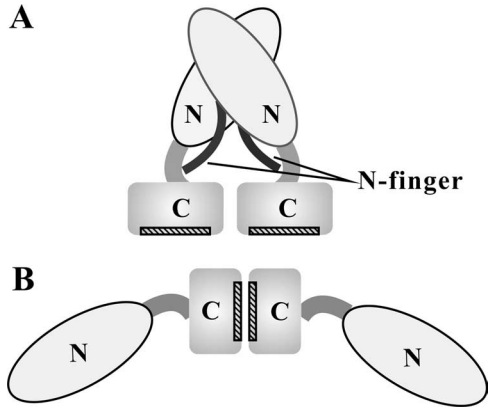

FIG. 6. Cartoon diagrams illustrating the dimerization pattern of WT SARS-CoV M^{pro} (A) or M^{pro}-Δ7 (B). The N-terminal and C-terminal domains are labeled "N" and "C", respectively. The N-finger is illustrated as a thick black line and is indicated in the figure. The novel dimer interface of the C-terminal domain is represented by hatched bars.

Shi et al. have suggested that the C-terminal domain of SARS-CoV M^{pro} plays a role in switching the enzyme from the inactive form (monomer) to the active form (dimer), solely based on the observation that the C-terminal domain alone of SARS-CoV M^{pro} forms a stable dimer (22). Meanwhile, based on the report by Shi et al., along with the crystal structure of a monomeric SARS-CoV M^{pro} G11A mutant, Chen et al. have proposed that the dimerization of SARS-CoV M^{pro} is initiated by the dimerization of its C-terminal domain in a recent study (5). However, our results indicate that M^{pro}-C not only exists in a stable dimeric form but also in a stable monomeric form. In addition, the dimerization interface of the M^{pro}-C dimer is unrelated to that of WT SARS-CoV M^{pro}. Therefore, it is not reasonable to assume a role of the C-terminal domain in the dimerization of SARS-CoV M^{pro} solely based on the existence of the M^{pro}-C dimer. Although the dimerization of M^{pro}-C may be related to the dimerization and function of SARS-CoV M^{pro}, the currently available evi-

dences are not enough to propose it is biologically relevant. More investigations need to be carried out to explore the biological relevance of the C-terminal domain dimerization before a conclusion can be reached.

In summary, we have reinvestigated the dimerization of SARS-CoV M^{pro}. We found that its C-terminal domain alone (M^{pro}-C) exists in both stable monomeric and stable dimeric forms, and there is no conversion between the two forms at room temperature. The stable M^{pro}-C dimer has a novel dimerization interface which has no direct correlation with the WT SARS-CoV M^{pro} dimerization interface, and no C-terminal domain dimerization form could be found in WT SARS-CoV M^{pro}. However, once the N-terminal seven residues (N-finger) are deleted, the truncated SARS-CoV M^{pro} mutant (M^{pro}-Δ7) can form a new type of dimer through its C-terminal domain dimerization, and the native active dimer form, which is in equilibrium with the monomer, no longer exists. Therefore, the N-finger is not only critical for the dimerization of SARS-CoV M^{pro} but also essential for it to form the right quaternary structure which is the enzymatically active form. Since the dimerization of SARS-CoV M^{pro} is very important for its function and the currently available data are not very consistent, our study helps to clarify the previous controversial statements about the roles of N-finger and the C-terminal domain in SARS-CoV M^{pro} dimerization and should contribute to the elucidation of the dimerization mechanism of SARS-CoV M^{pro}.

ACKNOWLEDGMENTS

All NMR experiments were carried out at the Beijing NMR Center. We thank Xuemin Zhang for kindly providing the cDNA of SARS-CoV, and we thank Zihe Rao and Haitao Yang for kindly providing the substrate for the enzyme activity assay and other assistance.

This research was supported by grant 2003CB514104 (to B.X. and C.J.) from the 973 Program, grant 30125009 from NSFC to B.X., and grant 2006AA02A323 from the 863 Program to C.J.

REFERENCES

1. **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.
2. **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.
3. **Chan, H. L., S. K. Tsui, and J. J. Sung.** 2003. Coronavirus in severe acute respiratory syndrome (SARS). Trends Mol. Med. **9:**323–325.
4. **Chen, S., L. Chen, J. Tan, J. Chen, L. Du, T. Sun, J. Shen, K. Chen, H. Jiang, and X. Shen.** 2005. Severe acute respiratory syndrome coronavirus 3C-like proteinase N terminus is indispensable for proteolytic activity but not for enzyme dimerization. Biochemical and thermodynamic investigation in conjunction with molecular dynamics simulations. J. Biol. Chem. **280:**164–173.
5. **Chen, S., T. Hu, J. Zhang, J. Chen, K. Chen, J. Ding, H. Jiang, and X. Shen.** 2008. Mutation of Gly11 on the dimer interface results in the complete crystallographic dimer dissociation of SARS-CoV 3CLpro: crystal structure with molecular dynamics simulations. J. Biol. Chem. **283:** 554–564.
6. **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.
7. **Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax.** 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR **6:**277–293.
8. **Dominguez, C., R. Boelens, and A. M. Bonvin.** 2003. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc. **125:**1731–1737.
9. **Fan, K., P. Wei, Q. Feng, S. Chen, C. Huang, L. Ma, B. Lai, J. Pei, Y. Liu, J. Chen, and L. Lai.** 2004. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. J. Biol. Chem. **279:**1637–1642.
10. **Graziano, V., W. J. McGrath, L. Yang, and W. F. Mangel.** 2006. SARS CoV main proteinase: the monomer-dimer equilibrium dissociation constant. Biochemistry **45:**14632–14641.
11. **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.
12. **Johnson, B. A., and R. A. Blevins.** 1994. NMR View: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR **4:**603–614.
13. **Koradi, R., M. Billeter, and K. Wuthrich.** 1996. MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. **14:**51–55, 29–32.
14. **Kuiken, T., R. A. Fouchier, M. Schutten, G. F. Rimmelzwaan, G. van Amerongen, D. van Riel, J. D. Laman, T. de Jong, G. van Doornum, W. Lim, A. E. Ling, P. K. Chan, J. S. Tam, M. C. Zambon, R. Gopal, C. Drosten, S. van der Werf, N. Escriou, J. C. Manuguerra, K. Stohr, J. S. Peiris, and A. D. Osterhaus.** 2003. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet **362:**263–270.
15. **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.
16. **Leng, Q., and Z. Bentwich.** 2003. A novel coronavirus and SARS. New Engl. J. Med. **349:**709.
17. **Marra, M. A., S. J. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, J. Khattra, J. K. Asano, S. A. Barber, S. Y. Chan, A. Cloutier, S. M. Coughlin, D. Freeman, N. Girn, O. L. Griffith, S. R. Leach, M. Mayo, H. McDonald, S. B. Montgomery, P. K. Pandoh, A. S. Petrescu, A. G. Robertson, J. E. Schein, A. Siddiqui, D. E. Smailus, J. M. Stott, G. S. Yang, F. Plummer, A. Andonov, H. Artsob, N. Bastien, K. Bernard, T. F. Booth, D. Bowness, M. Czub, M. Drebot, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Meyers, A. Kabani, Y. Li, S. Normand, U. Stroher, G. A. Tipples, S. Tyler, R. Vogrig, D. Ward, B. Watson, R. C. Brunham, M. Krajden, M. Petric, D. M. Skowronski, C. Upton, and R. L. Roper.** 2003. The genome sequence of the SARS-associated coronavirus. Science **300:**1399–1404.
18. **Mulder, F. A., D. Schipper, R. Bott, and R. Boelens.** 1999. Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins. J. Mol. Biol. **292:**111–123.
19. **Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini.** 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science **300:**1394–1399.
20. **Ruan, Y. J., C. L. Wei, A. L. Ee, V. B. Vega, H. Thoreau, S. T. Su, J. M. Chia, P. Ng, K. P. Chiu, L. Lim, T. Zhang, C. K. Peng, E. O. Lin, N. M. Lee, S. L. Yee, L. F. Ng, R. E. Chee, L. W. Stanton, P. M. Long, and E. T. Liu.** 2003. Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection. Lancet **361:**1779–1785.
21. **Sattler, M., J. Schleucher, and C. Griesinger.** 1999. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. Nucl. Magn. Reson. Spectrosc. **34:**93–158.
22. **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.
23. **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, H. Jiang, and R. Hilgenfeld.** 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.
24. **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.
25. **World Health Organization.** 2003. SARS outbreak contained worldwide. World Health Organization, Geneva, Switzerland. http://www.who.int/mediacentre/news/releases/2003/pr56/en/.
26. **Wishart, D. S., C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes.** 1995. 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J. Biomol. NMR **6:**135–140.
27. **Xu, T., A. Ooi, H. C. Lee, R. Wilmouth, D. X. Liu, and J. Lescar.** 2005.

Structure of the SARS coronavirus main proteinase as an active C2 crystallographic dimer. Acta Crystallogr. F **61:**964–966.
28. **Xue, X., H. Yang, W. Shen, Q. Zhao, J. Li, K. Yang, C. Chen, Y. Jin, M. Bartlam, and Z. Rao.** 2007. Production of authentic SARS-CoV M(pro) with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. J. Mol. Biol. **366:**965–975.
29. **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, L. Wong, G. Gao, S. Chen, Z. Chen, D. Ma, M. Bartlam, and Z. Rao.** 2005. Design of wide-spectrum inhibitors targeting coronavirus main proteases. PLoS Biol. **3:**e324.
30. **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, M. Bartlam, R. Hilgenfeld, and Z. Rao.** 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.