Rᴇᴠɪᴇᴡ

Activation and maturation of SARS-CoV main protease

Bin Xia ✉, **Xue Kang**

Beijing Nuclear Magnetic Resonance Center, College of Chemistry and Molecular Engineering, and School of Life Sciences, Peking University, Beijing 100871, China
✉ Correspondence: binxia@pku.edu.cn

ABSTRACT

The worldwide outbreak of the severe acute respiratory syndrome (SARS) in 2003 was due to the transmission of SARS coronavirus (SARS-CoV). The main protease (M^(pro)) of SARS-CoV is essential for the viral life cycle, and is considered to be an attractive target of anti-SARS drug development. As a key enzyme for proteolytic processing of viral polyproteins to produce functional non-structure proteins, M^(pro) is first auto-cleaved out of polyproteins. The monomeric form of M^(pro) is enzymatically inactive, and it is activated through homo-dimerization which is strongly affected by extra residues to both ends of the mature enzyme. This review provides a summary of the related literatures on the study of the quaternary structure, activation, and self-maturation of M^(pro) over the past years.

KEYWORDS   severe acute respiratory syndrome, M^(pro), structure, dimerization

INTRODUCTION

SARS coronavirus (SARS-CoV) is the etiological agent of the high fatality transmissible disease severe acute respiratory syndrome (SARS) outbreak in 2003 (Chan et al., 2003; Kuiken et al., 2003; Leng and Bentwich, 2003). After infection, two overlapping polyproteins pp1a (486 kDa) and pp1ab (790 kDa) encoded by the SARS-CoV genome are first produced, which are later proteolytically processed into 16 non-structural proteins (nsp1–16) required for viral replication and transcription (Snijder et al., 2003). Two virus-encoded proteinases are involved in this proteolytic process, a papain-like proteinase (nsp3) and a 3C-like proteinase (nsp5) (Snijder et al., 2003). The 3C-like proteinase, also known as

the main protease (M^(pro)), is first auto-cleaved from poly-proteins to yield the mature enzyme and further cleaves all 11 downstream non-structural proteins. Therefore, M^(pro) is an essential viral protein for the viral life cycle, and becomes an attractive target for anti-SARS drug development (Anand et al., 2003; Yang et al., 2003; Yang et al., 2005).

The recombinant M^(pro) expressed from *E. coli* exists in an equilibrium between monomer and dimer (Hsu et al., 2005b; Graziano et al., 2006a, 2006b), and only the dimeric form is enzymatically active for cleaving the Gln-Ser peptide bond of substrates (Fan et al., 2004; Chen et al., 2005; Shi and Song, 2006; Chen et al., 2008a; Lin et al., 2008; Shi et al., 2008; Hu et al., 2009). Therefore, the dimer interface was also suggested to be a potential target for the inhibitor design (Anand et al., 2003; Shi et al., 2004; Hsu et al., 2005b; Yang et al., 2005), and the activation mechanism of M^(pro) has drawn a lot of research attention in order to understand the maturation process of viral polyproteins and the self-maturation mechanism of M^(pro) from viral polyproteins.

In this review, we will summarize the related work on the activation and maturation of M^(pro).

STRUCTURE OF M^(pro) ACTIVE DIMER

The first crystal structure of the SARS-CoV main protease (M^(pro)) with a bound substrate analogue inhibitor was reported in 2003 (Yang et al., 2003). Thereafter, several structures of M^(pro) have been published (Hsu et al., 2005a; Hsu et al., 2005b; Tan et al., 2005; Xu et al., 2005). X-ray crystal structures revealed that M^(pro) forms a homo dimer, and each protomer contains three domains (Fig. 1). Domains I (residues 8–101) and II (residues 102–184) together, also called the N-terminal domain, adopt a fold similar to trypsin-like serine proteases with two anti-parallel β-barrels. Domain III (residues 201–303), also named the C-terminal domain or
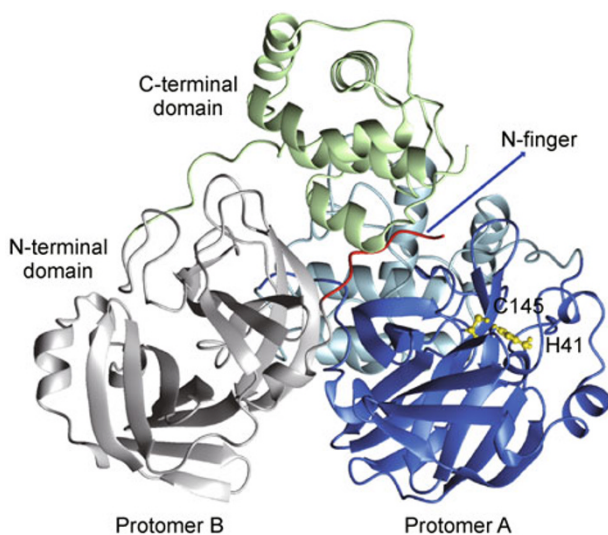
Figure 1. The crystal structure of M^{pro} active dimer (PDB: 2H2Z). The dimer structure is presented as ribbon diagram. The N-terminal and C-terminal of one protomer are colored blue and light blue, respectively; and those of the other protomer are colored grey and light green, respectively, with the N-finger in red. Side-chains of the catalytic dyad Cys145 and His41 are shown in yellow.

the extra domain, contains five α-helices and is connected to the N-terminal domain through a long loop (residues 185–200). The active site of M^{pro} is located at the N-terminal domain in a cleft between domains I and II, and it has a Cys-His catalytic dyad (Cys145 and His41) (Yang et al., 2003) instead of the Ser-His-Asp catalytic triad in common serine proteases. The N-terminal residues 1–7 (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 should be important for the dimerization (Yang et al., 2003). In addition, there are also inter-molecular interactions between the first helices of the two protomers.

In the first crystal structure of M^{pro}, the two protomers have distinct differences in the substrate-binding sites (Yang et al., 2003). Protomer A resembles a catalytically competent conformation, while the substrate binding pocket collapses in protomer B. The M^{pro} protein used for this structure determination has five extra residues (GPLGS) at the N-terminus (GPLGS-M^{pro}). It was later found that the additional residues to the N-terminus have a negative effect on the enzyme activity (Chou et al., 2004; Chen et al., 2005; Hsu et al., 2005b; Xue et al., 2007), suggesting that the exact N-terminus is important for defining the active site conformation. The crystal structures of M^{pro} with "authentic" N- and C-termini (WT-M^{pro}) were later reported (Xue et al., 2007), and the two protomers in this dimer structure are related by a 2-fold crystallographic symmetry axis, and each has a catalytically competent conformation.

The substrate binding subsite S1 pocket of M^{pro} confers absolute specificity for the Gln-P1 substrate residue on the enzyme (Ziebuhr et al., 2000), and consists of the side-chains of His163, Phe140 and the main chains of Glu166, Asn142, Gly143 and His172. It lies immediately next to the catalytic dyad (Cys145 and His41). In the structure of WT-M^{pro} (Xue et al., 2007), the NH$_3^+$ of Ser1 in one protomer forms a salt bridge with the carboxylate group of Glu166 and donates a hydrogen bond to the main-chain carbonyl group of Phe140 of the other protomer, thus stabilizing the S1 subsite and contributing to the dimerization. In the presence of additional residues at the N-terminus, NH group of Ser1 was unable to form the salt bridge, which induces a series of conformational changes and results in the collapse of the S1 pocket in protomer B of GPLGS-M^{pro} (Yang et al., 2003).

The monomer-dimer equilibrium dissociation constant (K_d) of M^{pro} has been determined by different methods, such as gel filtration chromatography, enzymatic kinetics assay, isothermal titration calorimetry, and analytical ultracentrifugation. However, there is a great discrepancy between K_d values reported by different groups (Chou et al., 2004; Fan et al., 2004; Kuo et al., 2004; Chen et al., 2005; Hsu et al., 2005a, 2005b; Chen et al., 2006; Graziano et al., 2006b; Shi and Song, 2006; Wei et al., 2006; Grum-Tokars et al., 2008). The highest K_d reported is 227 µmol/L from isothermal analysis, and K_d values determined from analytical ultracentrifugation analysis vary from 0.35 nmol/L to 14 µmol/L. Grum-Tokars et al. attributed the K_d value discrepancy to the different constructs used in expressing M^{pro} (Grum-Tokars et al., 2008). They found that M^{pro} with extra residues to the N- or C-termini has a higher K_d value, and the equilibrium is shifted towards the monomer. However, this still cannot explain the reported extremely low K_d of 0.35 nmol/L (Hsu et al., 2005b).

N-FINGER IN M^{pro} DIMERIZATION

The N-terminal domain of M^{pro} adopts a typical chymotrypsin fold, like that of the picornavirus 3C proteases (Seipelt et al., 1999). However, unlike 3C proteases, the M^{pro} is only enzymatically active as a homodimer (Fan et al., 2004) and the extra C-terminal α-helical domain is required for its activity (Shi et al., 2004; Shi and Song, 2006). Therefore, the dimerization mechanism of M^{pro} has been extensively studied.

As mentioned above, the N-finger of each protomer is squeezed between the two protomers and makes inter-molecular contacts, similar to that of the TGEV M^{pro} (Anand et al., 2003). In the crystal structure of WT-M^{pro} (Xue et al., 2007), the NH$_3^+$ group of Ser1 forms a salt bridge with Glu166 of the other protomer, and also forms a hydrogen bond with carbonyl of Phe140 of the other protomer. Arg4 forms an intermolecular salt bridge with Glu290. Ala7 and Val125 form two intermolecular hydrogen bonds. In addition, backbone NH of Ser10 may form a hydrogen bond with the hydroxyl

group of Ser10 on the other protomer, and NH of Gly11 may form intermolecular hydrogen bond with the carboxyl group of Glu14. Meanwhile, side-chain of Met6 makes hydrophobic interaction with Tyr126 and Phe140 of the other protomer.

However, studies of N-finger deletion mutants yielded some controversial results about the role of N-finger in the dimerization of M^{pro}. Chen et al. quantitatively characterized the dimerization features of the N-finger deleted M^{pro} (Chen et al., 2005). Their results showed that the mutant M^{pro} has a similar equilibrium dissociation constant to that of the full-length M^{pro}, even though N-finger deletion results in almost complete loss of enzymatic activity. It was thus concluded that N-finger is not crucial for the dimerization of M^{pro}, but only fundamental to its enzyme activity.

On the other hand, Hsu et al. generated several N-terminal truncated mutants of M^{pro}, and their results showed that M^{pro} still maintains 76% of enzyme activity and is still mainly a dimer after the removal of the N-terminal 3 amino acid residues (Hsu et al., 2005b). However, the removal of N-terminal 4–7 residues resulted in loss of enzyme activity and significantly reduced dimerization ability. They also found that deletion of the last C-terminal helix has a great effect on the dimerization and enzyme activity. As a result, it was concluded that both the C- and the N-terminal regions influence the dimerization and enzyme activity of M^{pro} (Chou et al., 2004; Hsu et al., 2005b). In another study, Wei et al. reported that the N-finger deletion mutant of M^{pro} could not form any dimer, and the N-terminal octapeptide can act as an inhibitor of M^{pro} dimerization (Wei et al., 2006).

Meanwhile, Shi et al. have reported that the N-terminal domain alone of M^{pro} is a monomer, while the C-terminal domain alone exists as a stable dimer (Shi et al., 2004). Thus, it was proposed that the C-terminal domain plays a role of switching the enzyme from an inactive monomer to an active dimer (Shi et al., 2004; Shi and Song, 2006). However, this is not consistent with the crystal structures of M^{pro}, as the contact area between the two C-terminal domains in the dimer is rather small and appears to be the consequence of dimerization rather than the cause (Yang et al., 2003; Xue et al., 2007).

These controversial issues were clarified in a later study. Zhong et al. re-investigated the dimerization of M^{pro}, and reported that the C-terminal domain alone of M^{pro} (M^{pro}-C) exists in both stable monomeric and stable dimeric forms (Zhong et al., 2008). No conversion was found between the two forms. The stable M^{pro}-C dimer was later demonstrated to be a 3D domain-swapped dimer (Zhong et al., 2009). It was also found that the N-finger deletion mutant of M^{pro} can also form a stable dimer through 3D domain swapping of the C-terminal domains, in addition to be a monomer. Therefore, it was concluded that N-finger is not only critical for the dimerization of M^{pro}, but also essential for it to form the right quaternary structure which is the enzymatically active form (Zhong et al., 2008).

STRUCTURE OF M^{pro} MONOMERIC MUTANTS

Extensive mutagenesis studies have been carried out to probe key residues that have a great impact on the dimerization of M^{pro} (Chou et al., 2004; Barrila et al., 2006; Shi and Song, 2006; Wei et al., 2006; Chang et al., 2007; Chen et al., 2008b; Lin et al., 2008). Although some of the studies were complicated by the use of N-terminal tag for the expressed M^{pro}, most of the results indicated that mutations that significantly reduce dimerization K_d are mainly found for residues at the dimer interface, such as Arg4, Met6, Gly11, Ser139, Glu290, Arg298A, and Gln299 (Fig. 2).

The crystal structures of three monomeric M^{pro} mutants G11A (Chen et al., 2008a, 2008b), S139A (Hu et al., 2009), and R298A (Shi et al., 2008) have been solved, and they are quite similar to each other. Both N-terminal and C-terminal domains of all three monomeric structures still adopt the same overall fold as those in the WT-M^{pro} structure, respectively. However, the inter-domain arrangement of the monomeric mutants has changed dramatically, characterized by an ~40° rotation between the N-terminal domain and the C-terminal domain, compared to the dimeric WT-M^{pro}. Another noticeable change is the displacement of the N-finger. In WT-M^{pro}, Ala7 and Phe8 are positioned near Ser113 of the same protomer, and NH of Phe8 forms a hydrogen bond with the hydroxyl group of Ser113. In the monomeric structures, Ala7 and Phe8 are located close to Val125, and NH of Phe8 forms a hydrogen bond with the carbonyl of Val125 in two of the structures. Also, side-chain of Lys5 is moved closer to Tyr126 and Phe160 and makes hydrophobic interaction in the monomeric mutants. In addition, the conformation of the linker loop between N- and C-terminal domains is also changed.

Examining the active site region of monomeric structures (Fig. 3), detailed structural variations are observed, especially for the loops constituting the catalytic machinery. The conformation of the loop consisting of residues 138–144 is significantly altered, and nearby the β hairpin of residues 111–129 also has large structural dislocation. In the monomeric mutants, a short 3$_{10}$-helix is formed by residues 139–141, which adopt a loop conformation in the WT-M^{pro} structure. The aromatic ring stacking for Phe140 and His163 is critical for stabilizing the substrate-binding pocket of the active M^{pro} dimer. While in the monomeric structure, the ring of Phe140 no longer interacts with that of His163, but is stacking with the ring of Tyr126 instead, leading to the significant conformation twist of residues 142–145 and the collapse of substrate binding pocket.

The structures of monomeric mutants of M^{pro} provide a reasonable explanation for why dimerization is absolutely essential for catalysis (Chen et al., 2008a; Shi et al., 2008; Hu et al., 2009). In the WT-M^{pro} dimer structure, the loop residues Gly138, Ser139, Phe140, and Leu141, as well as residues Val 125 and Tyr126 of the β hairpin, all make direct contact with
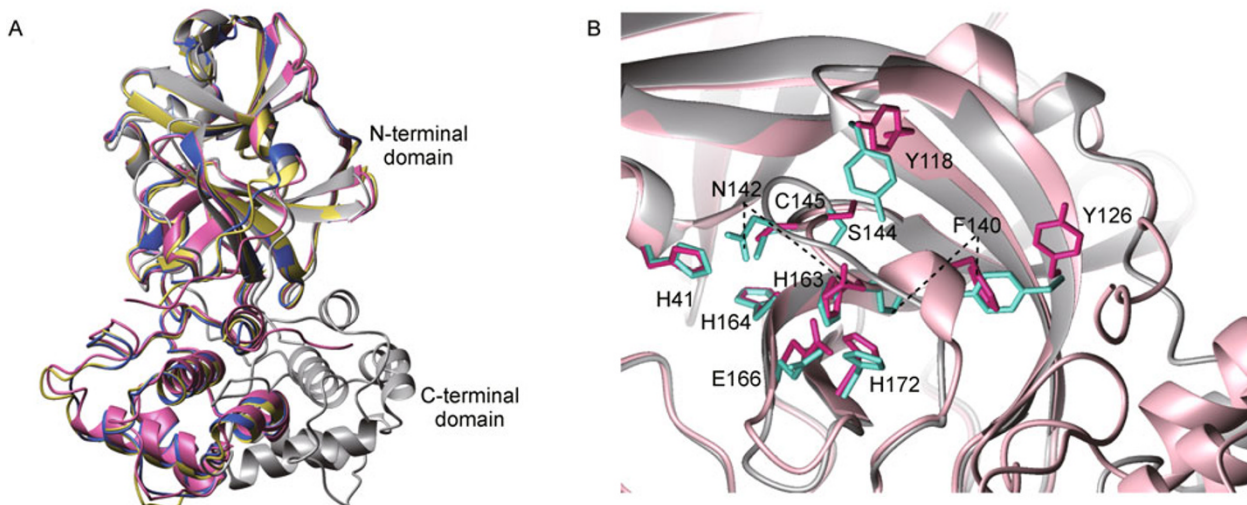
Figure 2. Structure comparison of WT-M^{pro} and three monomeric mutants. (A) Superimposition of WT-M^{pro} (PDB: 2H2Z, grey) and mutants G11A (PDB: 2PWX, blue), R298A (PDB: 2QCY, pink) and S139A (PDB: 3F9E, yellow) in ribbon diagram with N-terminal domains aligned. (B) Comparison of the active site area of WT-M^{pro} (grey) with that of R298A (pink) monomeric mutant. Side-chains of active site residues are shown.

the other protomer (Fig. 3). Therefore, the dimerization is required for these residues to adopt the proper conformations, which further define the right conformation of the active site.

3D DOMAIN SWAPPING OF M^{pro}

The C-terminal domain alone (M^{pro}-C) produced in *E. coli* has two forms, a stable monomeric form and a stable dimeric form due to 3D domain swapping (Zhong et al., 2008). The solution structure of the M^{pro}-C monomer is almost identical to that of the C-terminal domain in the crystal structure of full length M^{pro}, suggesting that the deletion of N-terminal domain does not change the fold of the C-terminal domain. Crystal structure of the M^{pro}-C dimer structure revealed that the first helices of the two molecules exchange their positions, and each is enwrapped by four other helices from the other molecule (Zhong et al., 2009). Each folding subunit of M^{pro}-C domain-swapped dimer still has the same general architecture as that of the monomer. The two folding subunits of the M^{pro}-C dimer are linked by two hinge loops consisting of residues 216–226, which is the loop following the first helix in the M^{pro}-C monomer. This elucidates the structure basis for the exceeding stability of the M^{pro}-C dimer and the lack of exchange between the monomer and dimer. Also, the structure of M^{pro}-C dimer indicates that it is unlikely that the dimerization of the C-terminal domain plays a role in switching the enzyme from the inactive monomeric form to the active dimeric form, as it has been suggested (Shi et al., 2004; Hu et al., 2009).
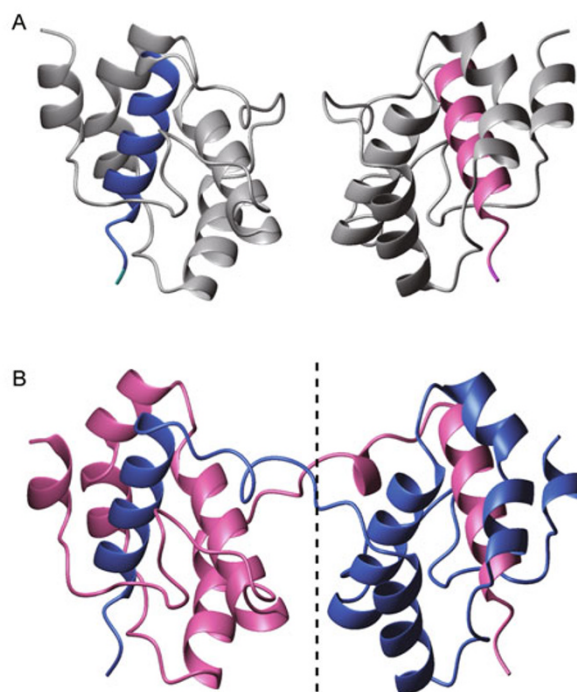

Figure 3. Structures of M^{pro}-C monomer and dimer. (A) Solution structure of M^{pro}-C (PDB: 2K7X) monomer in ribbon diagram is shown in two orientations with α_1-helix colored in pink and blue, respectively. (B) Ribbon diagrams of crystal structure of the 3D domain swapped dimer M^{pro}-C (PDB: 3EBN) with the two protomers colored in pink and blue, respectively.

Surprisingly, the 3D domain swapping of the M^{pro}-C can serve as a mechanism to enable the enzyme to be locked in the active conformation and to be constantly active. It was found that a mutant M^{pro} (M^{pro}-NE) with a 26-residue N-terminal fusion peptide including an enterokinase cleavage site right in front of the N-terminus also exists as a stable monomeric form and a stable dimeric form (Zhang et al., 2010). NMR study revealed that the M^{pro}-NE dimer is also formed through 3D domain swapping of their C-terminal domains. When the N-terminal extension peptide of M^{pro}-NE dimer is completely removed to produce the exact mature M^{pro} through enterokinase digestion, a stable octamer is automatically assembled. The M^{pro} octamer is also enzymatically active, and its activity is comparable to that of the previously well-studied form of mature M^{pro} that is in equilibrium between an inactive monomer and an active dimer, at a protein concentration of 16.9 μg/mL. However, when the total protein concentration is reduced to 0.68 μg/mL, the relative activity of the M^{pro} octamer is over 10 times higher.

Crystal structure revealed that the M^{pro} octamer is composed of four domain-swapped dimer subunits, and each is formed by two M^{pro} protomers with the last four helices (residues 227–306) of each C-terminal domain swapped (Zhang et al., 2010) (Fig. 4). The two C-terminal domains within one dimer subunit adopt the same fold as the 3D domain-swapped dimer of M^{pro}-C. In each dimer subunit, the remaining domain consists of the N-terminal domain of one protomer and the first helix of its C-terminal domain, together with the four swapped helices (swapped domain) from the other protomer, adopting an architecture just like a molecule in the structure of M^{pro} active dimer. Thus, the dimer subunit can also be viewed as two monomer-like structural units connected by two hinge loops. Two monomer-like structural units from different dimer subunits can pack into an active structural unit through the same interactions as those in the M^{pro} active dimer, enabling the formation of four active structural units with totally eight active sites. However, these interactions occur among four different protomers for each active structural unit in the M^{pro} octamer, and each active unit is connected to two other active units by hinge loops. All these features stabilize the octamer structure and enable the active conformation to be locked.

This unique structure perfectly explains why the relative enzyme activity of the M^{pro} octamer is much higher at low concentration. As the M^{pro} active dimer is in equilibrium with the inactive monomer, the active dimer fraction will be greatly reduced as the protein concentration goes down, while the M^{pro} active octamer is stable and constantly active. It was speculated that M^{pro} active octamer may play an important role in polyprotein maturation at an early stage of SARS-CoV infection when concentrations of pp1a, pp1b, and mature M^{pro} are extremely low, if it does exist in the initial period of SARS-CoV infection (Zhang et al., 2010).

POLYPROTEIN MATURATION MECHANISM

SARS-CoV M^{pro} is initially produced as the Nsp5 domain of the viral polyproteins pp1a and pp1ab before they are proteolytically processed into a total of 15 or 16 non-structural proteins. M^{pro} is first self-matured from polyproteins through auto-cleavage, and becomes a mature M^{pro} to further process other non-structural proteins out of pp1a and pp1ab. As most studies were focused on the mature M^{pro}, the auto-cleavage process is less well understood even though it is the first essential step for the viral polyprotein maturation.

Lin et al. first developed a Vero cell-based assay to demonstrate the auto-cleavage activity of M^{pro} (Lin et al., 2004). Shan et al. have introduced a 31-mer peptide with an auto-cleavage site to the N-terminus of M^{pro} and found the peptide could be auto-cleaved efficiently when expressing in *E. coli* (Shan et al., 2004). Interestingly, they reported that the C145S mutant of M^{pro} still showed a weak auto-cleavage activity, while the trans-cleavage activity was undetectable. This may suggest that there could be different mechanisms for M^{pro} to perform auto-cleavage and trans-cleavage.

Hsu et al. expressed C145A mutants of M^{pro} with a 10-residue extension in native polyprotein sequences to the N-terminus and C-terminus respectively, and found both mutants show significantly increased dimerization K_d compared to the mature M^{pro} (Hsu et al., 2005a). They also performed proteolysis assay using an N-terminal Trx-tagged and C-terminal GST-tagged C145A mutant with 10-residue extensions to both N- and C- termini of M^{pro} as substrate, and found that the mature M^{pro} cleaved the N-terminal site more efficiently than the C-terminal site. From the crystal structure of C145A mutant, they found that the active site of one protomer in the C145A dimer binds the C-terminal six amino acids of another molecule from a nearby asymmetric unit, which may mimic the product-bound form in the auto-cleavage process. They proposed that M^{pro} in polyproteins may still form some active dimer and thus process another molecule, as the K_d (0.35 nmol/L) they determined from AUC analysis for mature M^{pro} is a few order of magnitude smaller than those from other studies (10^{-6} to 10^{-5} mol/L) (Chou et al., 2004; Fan et al., 2004; Kuo et al., 2004; Chen et al., 2005; Hsu et al., 2005a; Hsu et al., 2005b; Chen et al., 2006; Shi and Song, 2006; Wei et al., 2006).

In 2010, two groups reported that the dimerization of the mature M^{pro} is enhanced by substrates. Cheng et al. reported that the initial velocity curves of R298A and R298L monomeric mutants show a cooperative effect (Cheng et al., 2010). AUC analyses indicate that the K_d values of both mutants are decreased by over 10 fold in the presence of small peptide substrates, even though the substrate is partially cleaved during the centrifugation process. The substrate-induced dimerization is reversible after removal of the substrate. Although they found that the K_d values with/without the substrate are comparable for WT-M^{pro}, it could be that most of
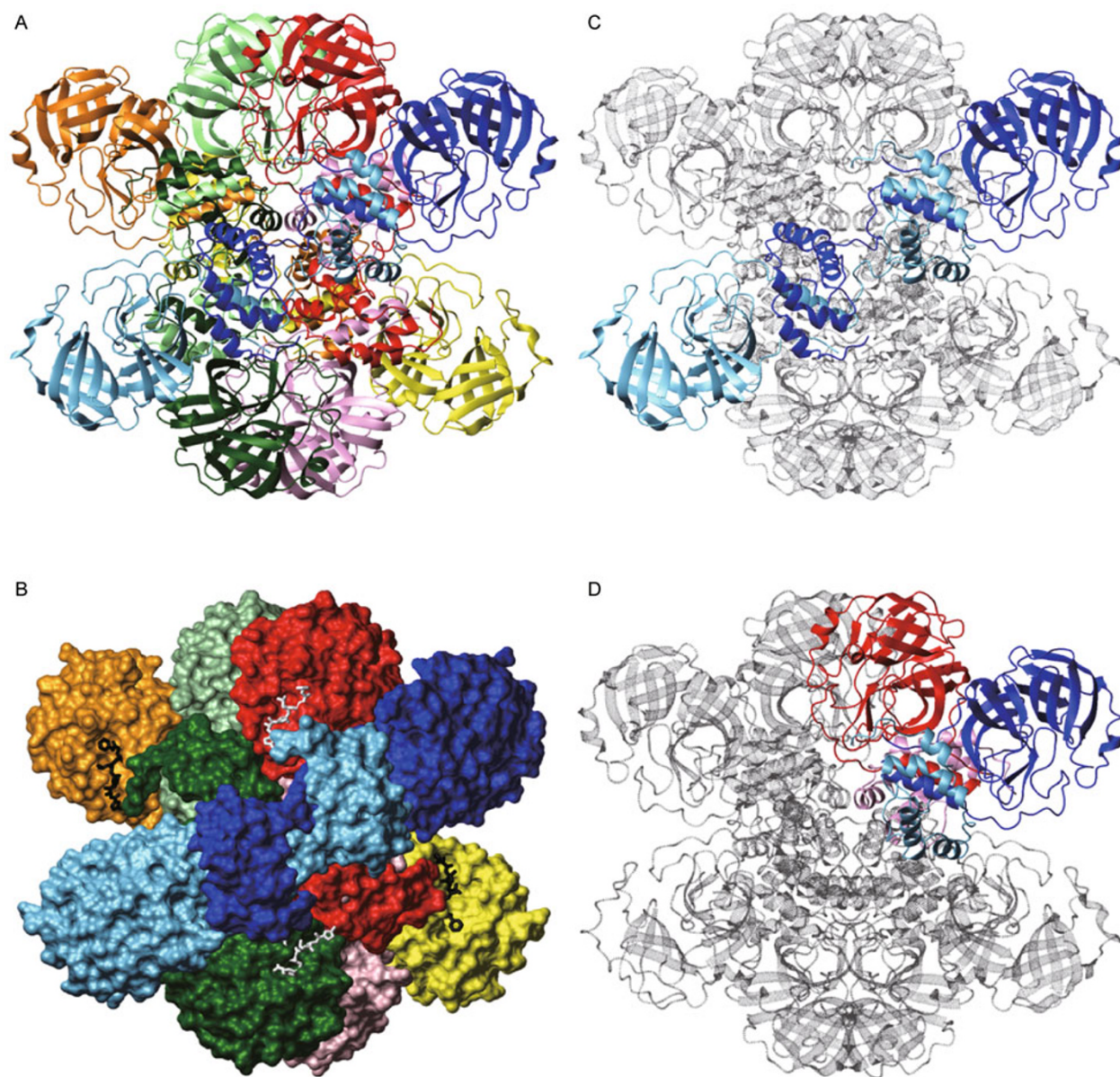
Figure 4.   Crystal structure of M^pro octamer. (A) M^pro octamer structure in ribbon diagram; (B) M^pro octamer structure in space fill mode; (C) Illustration of one domain-swapped dimer subunit (blue and light blue) in M^pro octamer; (D) Illustration of one active unit (blue and light blue, red and pink) in M^pro octamer.

the substrate have been cleaved early in the centrifugation process. They also reported that the dimerization of the E166A/R298A mutant is not enhanced by the substrate, and proposed that Glu166 plays a pivotal role in connecting the substrate binding site with the dimer interface.

Wei et al. first reported that an isatin inhibitor 5f of M^pro, which is also a substrate analog, can strongly induce M^pro dimer formation (Wei et al., 2010). The half maximal effective concentration (EC50) of dimer induction is ~ 1 μmol/L at an

M^pro concentration of 3 μmol/L, implying that only one substrate bound M^pro molecule can easily form a dimer with a free M^pro molecule. This may be the reason for the 15-fold difference in K_d values determined from the enzymatic assay and from the analytical ultra-centrifugation, suggesting that substrate of M^pro may allosterically regulate the enzyme activity.

The same group went further to address the effect of substrate on M^pro in polyprotein (Li et al., 2010). They

constructed two artificial polyproteins, His-C-XX(Q/E)-3CLP-Y and His-C-XX-C145A-Y, containing N-terminal CFP, M^{pro} (3CLP) with an N-terminal flanking 6-residue substrate peptide (XX), and C-terminal YFP. The substrate site Q to E mutation and the C145A mutation resulted in the two polyproteins not self-cleavable. It was found that both polyproteins are monomers in solution, while dimer can be observed in the presence of substrate analog 5f. His-C-XX(Q/E)-3CLP-Y is active and it can digest both the peptide substrate and the inactive polyprotein His-C-XX-C145A-Y. It was thus concluded that M^{pro} in the polyprotein is monomeric and the substrate-induced dimerization is essential for its enzyme activity. A maturation mechanism was also proposed: two M^{pro} in polyprotein can form transient dimer stabilized by binding the N-terminal substrate site of another M^{pro} in polyprotein, and further cleave to free its N terminus.

In another study, Chen et al. reported an interesting observation that the expression of GST or Trx tagged M^{pro} mutants E290R or R298E with an N-terminal M^{pro} cleavage site only results in the production of mature proteins as WT-M^{pro} does, even though no trans-cleavage activity could be detected for the two mutants (Chen et al., 2010). *In vitro* proteolysis assay revealed that M^{pro} mutant E290R still possesses obvious enzymatic activity towards the inactive GST tagged M^{pro} mutant C145A/E290R with an N-terminal M^{pro} cleavage site, but not for the effector domain of influenza A virus non-structural protein 1 with the same N-terminal cleavage site, while WT-M^{pro} can cleave both. These are consistent with the previously mentioned observation by Shan et al. (Shan et al., 2004), suggesting that the active dimer conformation of M^{pro} is unnecessary for the N-terminal auto-cleavage of M^{pro} since the E290R mutant is unable to form dimer. It was thus proposed that the N-terminal auto-cleavage might only need two "immature" M^{pro} in monomeric polyproteins to form an "intermediate" dimer that is not related to the active dimer of the mature M^{pro} (Chen et al., 2010).

SUMMARY

The main protease of SARS-CoV is a very interesting protein to study. It has three quaternary structure forms: an inactive monomer in equilibrium with an active dimer, and a stably active octamer. The dimerization of M^{pro} and thus its enzyme activity are very sensitive to extra residues to the N- or C-termini, and can be enhanced in the presence of substrate. The folding of its C-terminal domain is regulated by the N-finger. However, the more we know about M^{pro} from *in vitro* studies, the more questions we need to answer about its behavior *in vivo*. Can M^{pro} in polyprotein form the active dimer, or the 3D domain-swapped dimer? How exactly is it matured from the polyprotein? Is the auto-cleavage mechanism really different from that of trans-cleavage? Does the octamer exist *in vivo*? Is the self-association under regulation by other environmental factors, such as the lipid membrane, as it was suggested that M^{pro} in the polyprotein is flanked by transmembrane domains (Nsp4 and Nsp6) anchored to the double-membrane vesicles in the SARS-CoV infected cells (Knoops et al., 2008).

It has been pointed out that "perhaps the most important insight made over the past several years is that coronaviruses have and will likely continue to cross species and cause disease in unrelated hosts." (Perlman and Netland, 2009) The transmission may still result in severe disease as the SARS epidemic in the future. As an important drug target, more future studies are required to answer those questions for the main protease of SARS-CoV.

ACKNOWLEDGEMENTS

REFERENCES

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.

Barrila, J., Bacha, U., and Freire, E. (2006). Long-range cooperative interactions modulate dimerization in SARS 3CLpro. Biochemistry 45, 14908–14916.

Chan, H.L., Tsui, S.K., and Sung, J.J. (2003). Coronavirus in severe acute respiratory syndrome (SARS). Trends Mol Med 9, 323–325.

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.

Chen, H., Wei, P., Huang, C., Tan, L., Liu, Y., and Lai, L. (2006). Only one protomer is active in the dimer of SARS 3C-like proteinase. J Biol Chem 281, 13894–13898.

Chen, S., Chen, L., Tan, J., Chen, J., Du, L., Sun, T., Shen, J., Chen, K., Jiang, H., and Shen, X. (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.

Chen, S., Hu, T., Zhang, J., Chen, J., Chen, K., Ding, J., Jiang, H., and Shen, X. (2008a). Mutation of Gly-11 on the dimer interface results in the complete crystallographic dimer dissociation of severe acute respiratory syndrome coronavirus 3C-like protease: crystal structure with molecular dynamics simulations. J Biol Chem 283, 554–564.

Chen, S., Jonas, F., Shen, C., and Higenfeld, R. (2010). Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. Protein Cell 1, 59–74.

Chen, S., Zhang, J., Hu, T.C., Chen, K.X., Jiang, H.L., and Shen, X. (2008b). Residues on the dimer interface of SARS coronavirus 3C-like protease: dimer stability characterization and enzyme catalytic activity analysis. J Biochem 143, 525–536.

Cheng, S.C., Chang, G.G., and Chou, C.Y. (2010). Mutation of Glu-166 blocks the substrate-induced dimerization of SARS corona-

virus main protease. Biophys J 98, 1327–1336.

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.

Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma, L., Lai, B., Pei, J., Liu, Y., Chen, J., et al. (2004). Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. J Biol Chem 279, 1637–1642.

Graziano, V., McGrath, W.J., DeGruccio, A.M., Dunn, J.J., and Mangel, W.F. (2006a). Enzymatic activity of the SARS coronavirus main proteinase dimer. FEBS Lett 580, 2577–2583.

Graziano, V., McGrath, W.J., Yang, L., and Mangel, W.F. (2006b). SARS CoV main proteinase: The monomer-dimer equilibrium dissociation constant. Biochemistry 45, 14632–14641.

Grum-Tokars, V., Ratia, K., Begaye, A., Baker, S.C., and Mesecar, A. D. (2008). Evaluating the 3C-like protease activity of SARS-Coronavirus: Recommendations for standardized assays for drug discovery. Virus Res 133, 63–73

Hsu, M.F., Kuo, C.J., Chang, K.T., Chang, H.C., Chou, C.C., Ko, T.P., Shr, H.L., Chang, G.G., Wang, A.H., and Liang, P.H. (2005a). Mechanism of the maturation process of SARS-CoV 3CL protease. J Biol Chem 280, 31257–31266.

Hsu, W.C., Chang, H.C., Chou, C.Y., Tsai, P.J., Lin, P.I., and Chang, G.G. (2005b). 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.

Hu, T., Zhang, Y., Li, L., Wang, K., Chen, S., Chen, J., Ding, J., Jiang, H., and Shen, X. (2009). Two adjacent mutations on the dimer interface of SARS coronavirus 3C-like protease cause different conformational changes in crystal structure. Virology 388, 324–334.

Seipelt, J., Guarne, A., Bergmann, E., James, M., Sommergruber, W., Fita, I., and Skern, T., (1999). The structures of picornaviral proteinases. Virus Res 62, 159–168.

Knoops, K., Kikkert, M., Worm, S.H., Zevenhoven-Dobbe, J.C., van der Meer, Y., Koster, A.J., Mommaas, A.M., and Snijder, E.J. (2008). SARS-coronavirus replication is supported by a reticulo-vesicular network of modified endoplasmic reticulum. PLoS Biol 6, e226.

Kuiken, T., Fouchier, R.A., Schutten, M., Rimmelzwaan, G.F., van Amerongen, G., van Riel, D., Laman, J.D., de Jong, T., van Doornum, G., Lim, W., et al. (2003). Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet 362, 263–270.

Kuo, C.J., Chi, Y.H., Hsu, J.T., and Liang, P.H. (2004). Characterization of SARS main protease and inhibitor assay using a fluorogenic substrate. Biochem Biophys Res Commun 318, 862–867.

Leng, Q., and Bentwich, Z. (2003). A novel coronavirus and SARS. N Engl J Med 349, 709.

Li, C., Qi, Y., Teng, X., Yang, Z., Wei, P., Zhang, C., Tan, L., Zhou, L., Liu, Y., and Lai, L. (2010). Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like proteinase. J Biol Chem 285, 28134–28140.

Lin, C.W., Tsai, C.H., Tsai, F.J., Chen, P.J., Lai, C.C., Wan, L., Chiu, H. H., and Lin, K.H. (2004). Characterization of trans- and cis-cleavage activity of the SARS coronavirus 3CLpro protease: basis

for the in vitro screening of anti-SARS drugs. FEBS Lett 574, 131–137.

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.

Perlman, S., and Netland, J. (2009). Coronaviruses post-SARS: update on replication and pathogenesis. Nat Rev Microbiol 7, 439–450.

Seipelt, J., Guarné, A., Bergmann, E., James, M., Sommergruber, W., Fita, I., and Skern, T. (1999). The structures of picornaviral proteinases. Virus Res 62, 159–168.

Shan, Y.F., Li, S.F., and Xu, G.J. (2004). A novel auto-cleavage assay for studying mutational effects on the active site of severe acute respiratory syndrome coronavirus 3C-like protease. Biochem Biophys Res Commun 324, 579–583.

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.

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

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.

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L., Guan, Y., Rozanov, M., Spaan, W.J., and Gorbalenya, A.E. (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol 331, 991–1004.

Tan, J., Verschueren, K.H., Anand, K., Shen, J., Yang, M., Xu, Y., Rao, Z., Bigalke, J., Heisen, B., Mesters, J.R., 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.

Wei, P., Fan, K., Chen, H., Ma, L., Huang, C., Tan, L., Xi, D., Li, C., Liu, Y., Cao, A., et al. (2006). The N-terminal octapeptide acts as a dimerization inhibitor of SARS coronavirus 3C-like proteinase. Biochem Biophys Res Commun 339, 865–872.

Wei, P., Li, C.M., Zhou, L., Liu, Y., and Lai, L.H. (2010). Substrate Binding and Homo Dimerization of SARS 3CL Proteinase are Mutual Allosteric Effectors. Acta Phys Chim Sin 26, 5.

Xu, T., Ooi, A., Lee, H.C., Wilmouth, R., Liu, D.X., and Lescar, J. (2005). Structure of the SARS coronavirus main proteinase as an active C2 crystallographic dimer. Acta Crystallogr Sect F Struct Biol Cryst Commun 61, 964–966.

Xue, X., Yang, H., Shen, W., Zhao, Q., Li, J., Yang, K., Chen, C., Jin, Y., Bartlam, M., and Rao, Z. (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.

Yang, H., Xie, W., Xue, X., Yang, K., Ma, J., Liang, W., Zhao, Q., Zhou, Z., Pei, D., Ziebuhr, J., et al. (2005). Design of wide-spectrum inhibitors targeting coronavirus main proteases. PLoS Biol 3, e324.

Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., et al. (2003). The crystal structures of severe

acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc Natl Acad Sci U S A 100, 13190–13195.

Zhang, S.N., Zhong, N., Xue, F., Kang, X., Ren, X.B., Jin, C.W., Lou, Z.Y., and Xia, B. (2010). Three-dimensional domain swapping as a mechanism to lock the active conformation in a super-active octamer of SARS-CoV main protease. Protein Cell 1, 371–383.

Zhong, N., Zhang, S., Xue, F., Kang, X., Zou, P., Chen, J., Liang, C., Rao, Z., Jin, C., Lou, Z., *et al.* (2009). C-terminal domain of

SARS-CoV main protease can form a 3D domain-swapped dimer. Protein Sci 18, 839–844.

Zhong, N., Zhang, S., Zou, P., Chen, J., Kang, X., Li, Z., Liang, C., Jin, C., and Xia, B. (2008). Without its N-finger, the main protease of severe acute respiratory syndrome coronavirus can form a novel dimer through its C-terminal domain. J Virol 82, 4227–4234.

Ziebuhr, J., Snijder, E.J., and Gorbalenya, A.E. (2000). Virus-encoded proteinases and proteolytic processing in the Nidovirales. J Gen Virol 81, 853–879.