

RESEARCH ARTICLE

Three-dimensional domain swapping as a mechanism to lock the active conformation in a super-active octamer of SARS-CoV main protease

Shengnan Zhang^{1,2*}, Nan Zhong^{1,2*}, Fei Xue^{3*}, Xue Kang^{1,2}, Xiaobai Ren^{1,2}, Jiaxuan Chen^{1,4}, Changwen Jin^{1,2,4}, Zhiyong Lou³✉, Bin Xia^{1,2,4}✉

¹ Beijing Nuclear Magnetic Resonance Center, Peking University, Beijing 100871, China

² College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

³ Structural Biology Laboratory, Tsinghua University, Beijing 100084, China

⁴ College of Life Sciences, Peking University, Beijing 100871, China

✉ Correspondence: binxia@pku.edu.cn (B. Xia), louzy@xtal.tsinghua.edu.cn (Z. Lou)

Received February 26, 2010 Accepted March 18, 2010

ABSTRACT

Proteolytic processing of viral polyproteins is indispensable for the lifecycle of coronaviruses. The main protease (M^{Pro}) of SARS-CoV is an attractive target for anti-SARS drug development as it is essential for the polyprotein processing. M^{Pro} is initially produced as part of viral polyproteins and it is matured by autocleavage. Here, we report that, with the addition of an N-terminal extension peptide, M^{Pro} can form a domain-swapped dimer. After complete removal of the extension peptide from the dimer, the mature M^{Pro} self-assembles into a novel super-active octamer (AO- M^{Pro}). The crystal structure of AO- M^{Pro} adopts a novel fold with four domain-swapped dimers packing into four active units with nearly identical conformation to that of the previously reported M^{Pro} active dimer, and 3D domain swapping serves as a mechanism to lock the active conformation due to entanglement of polypeptide chains. Compared with the previously well characterized form of M^{Pro} , in equilibrium between inactive monomer and active dimer, the stable AO- M^{Pro} exhibits much higher proteolytic activity at low concentration. As all eight active sites are bound with inhibitors, the polyvalent nature of the interaction between AO- M^{Pro} and its polyprotein substrates with multiple cleavage sites, would make AO- M^{Pro} functionally much more superior than the M^{Pro} active

dimer for polyprotein processing. Thus, during the initial period of SARS-CoV infection, this novel active form AO- M^{Pro} should play a major role in cleaving polyproteins as the protein level is extremely low. The discovery of AO- M^{Pro} provides new insights about the functional mechanism of M^{Pro} and its maturation process.

KEYWORDS SARS-CoV, main protease, crystal structure, 3D domain swapping, polyprotein processing

INTRODUCTION

SARS coronavirus (SARS-CoV) was identified as the etiological agent of the pandemic transmissible disease severe acute respiratory syndrome (SARS). During the outbreak in 2003, SARS-CoV infected more than 8400 people worldwide with a high fatality rate of about 10% (Chan et al., 2003; Kuiken et al., 2003; Leng and Bentwich, 2003). SARS-CoV is a positive-sense, single-stranded RNA virus, with the 5' two-third genome of the virus encoding two overlapping polyproteins pp1a (486 kDa) and pp1ab (790 kDa). During infection, the two polyproteins are proteolytically processed into 16 nonstructural proteins (nsp1–16) required for viral replication and transcription (Snijder et al., 2003). The proteolytic process is mediated by two virus-encoded proteinases, a papain-like proteinase (nsp3) and a 3C-like proteinase (nsp5) (Snijder et al., 2003). The 3C-like

*These authors contributed equally to the work.

proteinase, also known as the main protease (M^{pro}), is matured from polyproteins through autocleavage and further cleaves all 11 downstream nonstructural proteins. This makes M^{pro} essential for the viral life cycle, and thus an attractive target for anti-SARS drug development (Anand et al., 2003; Yang et al., 2003, 2005).

It has been reported that SARS-CoV M^{pro} exists in solution as an equilibrium between monomeric and dimeric forms (Hsu et al., 2005; Graziano et al., 2006a, b), and only the dimeric form is enzymatically active (Fan et al., 2004; Chen et al., 2005, 2006; Shi and Song, 2006; Chen et al., 2008; Lin et al., 2008; Shi et al., 2008; Hu et al., 2009). Crystal structures of SARS-CoV M^{pro} revealed a symmetric homodimer which has a similar fold to other reported coronavirus M^{pro} structures (Anand et al., 2002, 2003; Yang et al., 2003; Xue et al., 2007). Each protomer in the dimeric structure contains an N-terminal domain (residues 1–184) with a chymotrypsin-like fold and a unique C-terminal domain (residues 201–303) containing five α -helices. The substrate binding pocket is located in a cleft of the M^{pro} N-terminal domain with the catalytic dyad consisting of residues His41 and Cys145 (Yang et al., 2003). The N-terminal residues 1–7 (N-finger) of each protomer are squeezed in between the two protomers and make contacts with both the N-terminal and C-terminal domain of the other protomer (Yang et al., 2003). Therefore, many studies focused on elucidating the roles of N-finger and the C-terminal domain in the dimerization of M^{pro} (Shi et al., 2004; Chen et al., 2005; Wei et al., 2006; Zhong et al., 2008; Hu et al., 2009; Zhong et al., 2009).

We have demonstrated that 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 (Zhong et al., 2008). The M^{pro} C-terminal domain alone (M^{pro} -C) is expressed in *E. coli* as a stable monomer and a stable dimer in solution, with no apparent exchange (Zhong et al., 2009). The stable dimer adopts a novel fold which is characterized by 3D domain swapping with two helices of the two molecules interchange their positions. The N-finger deletion mutant of M^{pro} (M^{pro} - Δ 7) can also form a stable dimer due to 3D domain swapping of its C-terminal domain. However, it is not clear whether 3D domain-swapping plays a role in wild-type M^{pro} .

Three-dimensional domain swapping is a special way of protein association to generate stable multimers, by which identical domains from different molecules exchange their positions (Liu and Eisenberg, 2002; Gronenborn, 2009). In this way, the core tertiary structures of the domain-swapped multimers are similar to those of monomers. Up to date, domain-swapped structures of more than 60 proteins are available in the Protein Data Bank. The biologic roles of 3D domain swapping have been widely discussed and some functions have been elucidated. 3D domain swapping is recognized as a general mechanism for the formation of amyloid fibers which cause conformational diseases

(Yamasaki et al., 2008). HIV-1 capsid dimerization via 3D domain swapping is proposed to be a critical event in immature viral particle assembly (Ivanov et al., 2007). Bovine pancreatic ribonuclease A can oligomerize via 3D domain swapping, and acquires novel biologic functions such as antitumor activity (Libonati et al., 2008).

Here, we report that fusion of an extension peptide to the N terminus of M^{pro} can also result in the formation of a 3D domain-swapped dimer. After proteolytic removal of the extension peptide, we discovered the existence of a previously unknown active form of mature M^{pro} , which is a stable homo-octamer self-assembled by four domain swapped dimers. Each protomer of the active octamer has exactly the same sequence as the previously well characterized matured form of M^{pro} . Crystal structure reveals that, due to 3D domain swapping, the active conformation is locked in the octamer by forming four active units, each adopting the same conformation as the M^{pro} active dimer. At low protein concentration, the proteolytic activity of the M^{pro} active octamer is much higher than that of M^{pro} active dimer which is in equilibrium with an inactive monomer.

RESULTS

M^{pro} can form a 3D domain-swapped dimer with N-terminal extension peptides

It has been pointed out that extra residues to the N terminus of mature M^{pro} result in lower dimer ratio and enzymatic activity (Xue et al., 2007). We constructed an N-terminal extension mutant of M^{pro} (M^{pro} -NE) by fusing a 26-residue peptide to the N terminus of M^{pro} , with an enterokinase cleavage site right in front of the N terminus.

M^{pro} -NE produced in *E. coli* exists as two forms that can be separated by gel-filtration, with retention volumes of 15.5 mL and 13.7 mL (Fig. 1A). The apparent molecular weight calculated based on retention volume is 34.5 kDa for the 15.5-mL fraction, and 71.1 kDa for the 13.7-mL fraction. The two forms should correspond to a monomeric form (15.5-mL fraction, M- M^{pro} -NE) and a dimeric form (13.7-mL fraction, D- M^{pro} -NE) of M^{pro} -NE, as the theoretical molecular weights of the M^{pro} -NE monomer is 36.7 kDa. Both forms migrated on an SDS-PAGE gel at the same position as the M^{pro} -NE monomer, with or without DTT, suggesting that the formation of D- M^{pro} -NE is not through a disulfide bond, even though M^{pro} -NE has 12 free cysteine residues. Similar to the N-finger deletion mutant M^{pro} - Δ 7, no exchange between monomeric and dimeric forms was observed for days.

Most of the NH peaks in 2D 1H - ^{15}N HSQC spectra of M- M^{pro} -NE and D- M^{pro} -NE overlap well, suggesting that most parts of the structures are the same between the two forms. Using the same strategy as for characterizing the domain-swapped dimer of M^{pro} - Δ 7 (Zhong et al., 2008), we compared 2D 1H - ^{15}N HSQC spectra of M- M^{pro} -NE and D- M^{pro} -NE with

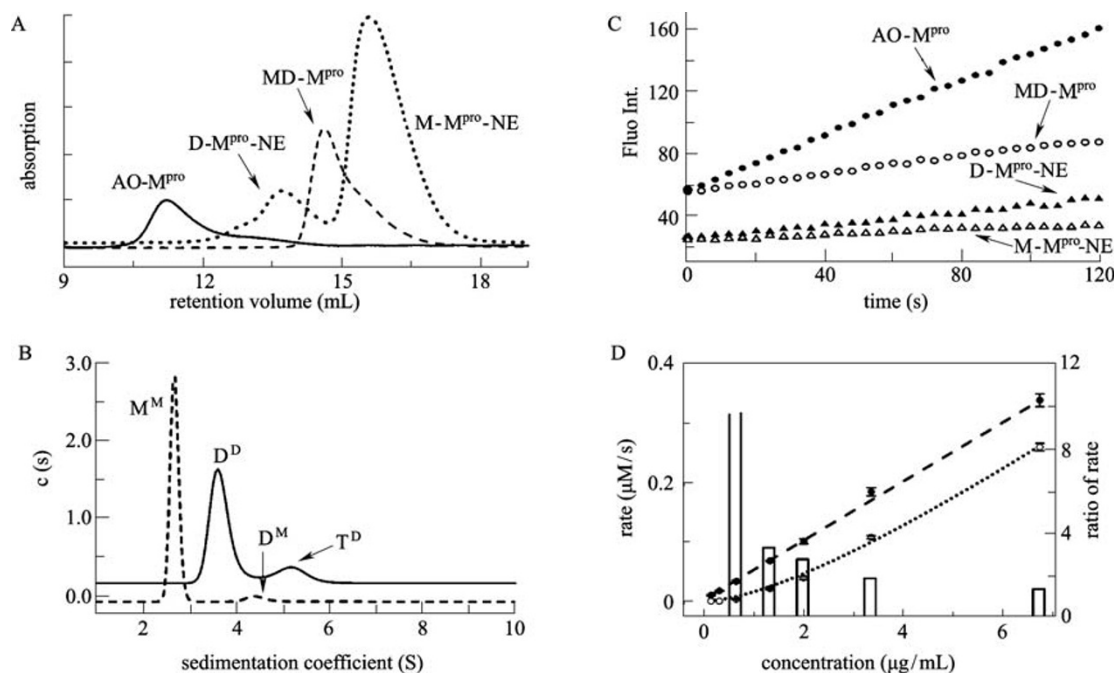


Figure 1. Oligomerization states and enzymatic activity of $M^{\text{pro-NE}}$ and mature M^{pro} . (A) Elution profiles of $M^{\text{pro-NE}}$ (dotted line), $MD-M^{\text{pro}}$ (broken line) and $AO-M^{\text{pro}}$ (solid line) from a gel-filtration column. Elution peaks for $M-M^{\text{pro-NE}}$ and $D-M^{\text{pro-NE}}$ are indicated. The peak heights have been adjusted arbitrarily. (B) Sedimentation coefficient $c(s)$ distribution of $M-M^{\text{pro-NE}}$ (1.2 mg/mL, broken line) and $D-M^{\text{pro-NE}}$ (1.0 mg/mL, solid line). M^{M} and D^{M} represent monomer and dimer of $M-M^{\text{pro-NE}}$, D^{D} and T^{D} represent dimer and tetramer of $D-M^{\text{pro-NE}}$, respectively. (C) Hydrolysis of the fluorogenic substrate by $M-M^{\text{pro-NE}}$ (367 $\mu\text{g/mL}$, open triangle), $D-M^{\text{pro-NE}}$ (367 $\mu\text{g/mL}$, filled triangle), $MD-M^{\text{pro}}$ (1.3 $\mu\text{g/mL}$, open circle) and $AO-M^{\text{pro}}$ (1.3 $\mu\text{g/mL}$, filled circle). (D) Initial hydrolysis rates of $MD-M^{\text{pro}}$ (open circle) and $AO-M^{\text{pro}}$ (filled circles) at different concentrations. The broken line is the linear fitting curve for $AO-M^{\text{pro}}$. The dotted line is the fitting curve for $MD-M^{\text{pro}}$.

those of the monomer and the domain-swapped dimer of $M^{\text{pro-C}}$ (Fig. 2). It was found that the signature backbone NH peaks of the $M^{\text{pro-C}}$ dimer, including those from residues D216, R217, L220, L227, V261, D263, A267, L272, G275, M276, T280 and L282, along with the side-chain NH signal of W218 and the side-chain NH_2 signals of N274 and N277, overlap with those of $D-M^{\text{pro-NE}}$ but not $M-M^{\text{pro-NE}}$ (Fig. 2). Vice versa, the signature peaks of the $M^{\text{pro-C}}$ monomer from residues mentioned above only overlap with those of $M-M^{\text{pro-NE}}$. This implies that the C-terminal domain of $M-M^{\text{pro-NE}}$ has the same fold as $M^{\text{pro-C}}$ monomer, while the C-terminal domain of $D-M^{\text{pro-NE}}$ adopts the same conformation as the domain-swapped dimer of $M^{\text{pro-C}}$, i.e., $D-M^{\text{pro-NE}}$ is also a domain-swapped dimer formed through C-terminal domains of the two molecules.

Both forms of $M^{\text{pro-NE}}$ exhibit some enzymatic activity, though far below that of mature M^{pro} . The initial substrate hydrolysis rates for $M-M^{\text{pro-NE}}$ and $D-M^{\text{pro-NE}}$, both at a total protein concentration of 367 $\mu\text{g/mL}$, are 15.0 nM/s and 6.3 nM/s, respectively (Fig. 1C). Since the monomeric form of mature M^{pro} is inactive, $M-M^{\text{pro-NE}}$ should be inactive being a monomer. It is most likely that $M-M^{\text{pro-NE}}$ need further dimerize in a similar fashion to that of the M^{pro} active dimer to

be active. Meanwhile, as $D-M^{\text{pro-NE}}$ is a domain-swapped dimer unrelated to the M^{pro} active dimer, dimerization of two $D-M^{\text{pro-NE}}$ molecules may also be required for it to adopt an active conformation. Analytical ultracentrifugation analysis confirmed that there is a small fraction of tetramer in $D-M^{\text{pro-NE}}$ ($D-M^{\text{pro-NE}}$ tetramer), and a small fraction of $M-M^{\text{pro-NE}}$ is in a dimeric form ($M-M^{\text{pro-NE}}$ dimer) which should be different from $D-M^{\text{pro-NE}}$ (Fig. 1B). In other words, $M-M^{\text{pro-NE}}$ exists in a monomer-dimer equilibrium, while $D-M^{\text{pro-NE}}$ exists in a dimer-tetramer equilibrium in solution. The dissociation constant can be calculated from the analytical ultracentrifugation analysis data, and the K_d for $D-M^{\text{pro-NE}}$ dimer/tetramer is 0.1 mM, while the K_d for $M-M^{\text{pro-NE}}$ monomer/dimer is 0.8 mM. Therefore, it is estimated that the amount of $D-M^{\text{pro-NE}}$ tetramer is about 2 times over that of the $M-M^{\text{pro-NE}}$ dimer when the total protein concentration is 367 $\mu\text{g/mL}$. This is consistent with the 2-fold difference in the detected enzymatic activity between $D-M^{\text{pro-NE}}$ and $M-M^{\text{pro-NE}}$.

M^{pro} can be matured into a super-active octamer

The N-terminal extension peptide of $M^{\text{pro-NE}}$ can be completely removed to produce the exact mature M^{pro}

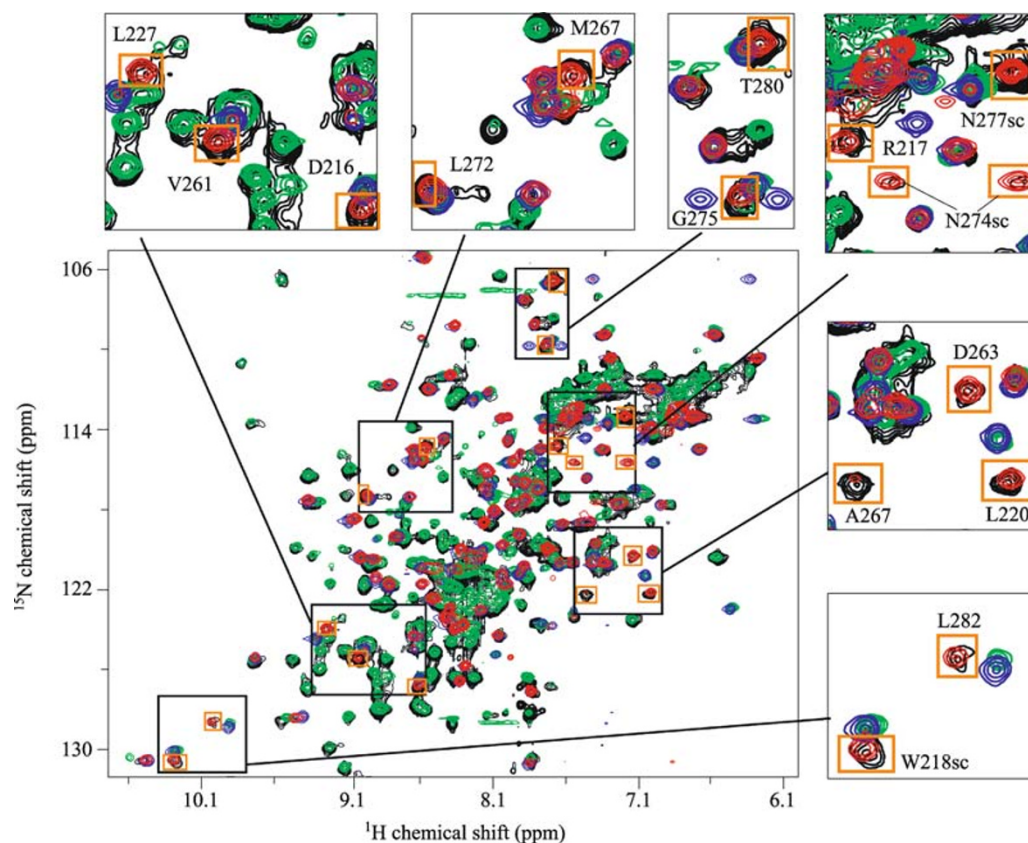


Figure 2. Overlay of 2D ^1H - ^{15}N HSQC spectra of monomeric and dimeric $\text{M}^{\text{pro-NE}}$ and $\text{M}^{\text{pro-C}}$. Black peaks belong to $\text{D-M}^{\text{pro-NE}}$, green peaks belong to $\text{M-M}^{\text{pro-NE}}$, red peaks are from $\text{M}^{\text{pro-C}}$ dimer, and blue peaks are from $\text{M}^{\text{pro-C}}$ monomer. Signature NH peaks of $\text{M}^{\text{pro-C}}$ dimer are indicated by orange squares. Six areas of the spectra are enlarged and displayed for clarity, in which signature NH peaks of $\text{M}^{\text{pro-C}}$ dimer are labeled with one-letter amino acid code and residue number.

through enterokinase digestion. As expected, the retention volume of the mature M^{pro} from $\text{M-M}^{\text{pro-NE}}$ showed an obvious concentration dependence on a gel filtration column, and its enzymatic activity is dependent on both the protein concentration and the dissociation constant (Fig. 1D). Similar methods have been used in previous studies for generating mature M^{pro} (Kuo et al., 2004; Xue et al., 2007; Verschueren et al., 2008).

Surprisingly, a higher order oligomer is observed after removal of the N-terminal extension peptide from $\text{D-M}^{\text{pro-NE}}$, and its retention volume (11.2 mL) is not concentration-dependent, indicating the absence of exchange. The apparent molecular weight calculated based on this retention volume is 271 kDa, which is close to eight times the theoretical molecular weight of M^{pro} (8×33.4 kDa), suggesting that this higher-order oligomer is a homo-octamer.

This newly identified M^{pro} octamer is also enzymatically active. At a total protein concentration of $16.9 \mu\text{g/mL}$, the initial hydrolysis rate of M^{pro} octamer ($0.92 \mu\text{M/s}$) is comparable to that of the previously well-studied form of mature M^{pro} ($0.94 \mu\text{M/s}$) that is in equilibrium between an inactive monomer and an active dimer (MD-M^{pro}). Therefore, we

have discovered a previously unknown form of mature M^{pro} , which is a novel active octamer (AO-M^{pro}). However, in contrast to MD-M^{pro} whose enzymatic activity is dependent on both the enzyme concentration and the dissociation constant, the initial substrate hydrolysis rate of AO-M^{pro} is only linearly dependent on protein concentration. When the total protein concentration is reduced, the activity of MD-M^{pro} decreases much more significantly than that of AO-M^{pro} , and the initial hydrolysis rate of AO-M^{pro} is about 11 times higher than that of MD-M^{pro} at a total protein concentration of $0.68 \mu\text{g/mL}$. Furthermore, below a concentration of $0.34 \mu\text{g/mL}$, the enzymatic activity of MD-M^{pro} is not detectable, while AO-M^{pro} still shows obvious hydrolysis activity (Fig. 1D). Therefore, we discovered a novel super-active octameric form of mature M^{pro} .

Octameric M^{pro} is assembled due to 3D domain swapping

The crystal structure of AO-M^{pro} in complex with the previously reported Michael acceptor inhibitor N3 (Yang et al., 2005) was determined using the molecular replacement (MR) method and refined at 3.2-\AA resolution with a final R_{work}

value of 22.0% ($R_{free} = 27.2\%$) (Fig. 3, PDB code: 3IWM). The crystals belongs to the space group $P4_12_12$, and there are four M^{pro} monomers per asymmetric unit, with a Matthews coefficient of $4.04 \text{ \AA}^3/\text{Da}$, corresponding to 69.6% solvent content (Matthews, 1968). The whole homo-octameric AO- M^{pro} comprises two asymmetric units related via a crystallographic twofold axis (Fig. 4C) and possesses a novel “wheel and axle” shaped structure with approximate dimensions of $135 \text{ \AA} \times 115 \text{ \AA} \times 110 \text{ \AA}$ (Fig. 3A). Together with two other pseudo twofold axes, these features endow the structure with high symmetry.

AO- M^{pro} is a homo-octamer composed of four domain-swapped dimer subunits (DSD), which can be divided into two types according to their roles in structural formation: two identical inner DSDs and two identical outer DSDs (Fig. 3B and 3C). Each DSD adopts a rod-like shape, and is formed by two M^{pro} protomers with the last four helices (residues 227–306) of each C-terminal domain swapped, i.e., the two C-terminal domains within one DSD fold into the same domain-swapped dimer we have reported for M^{pro} C-terminal domain alone (Zhong et al., 2009). The two inner DSDs pack against each other at the middle of the swapped C-terminal

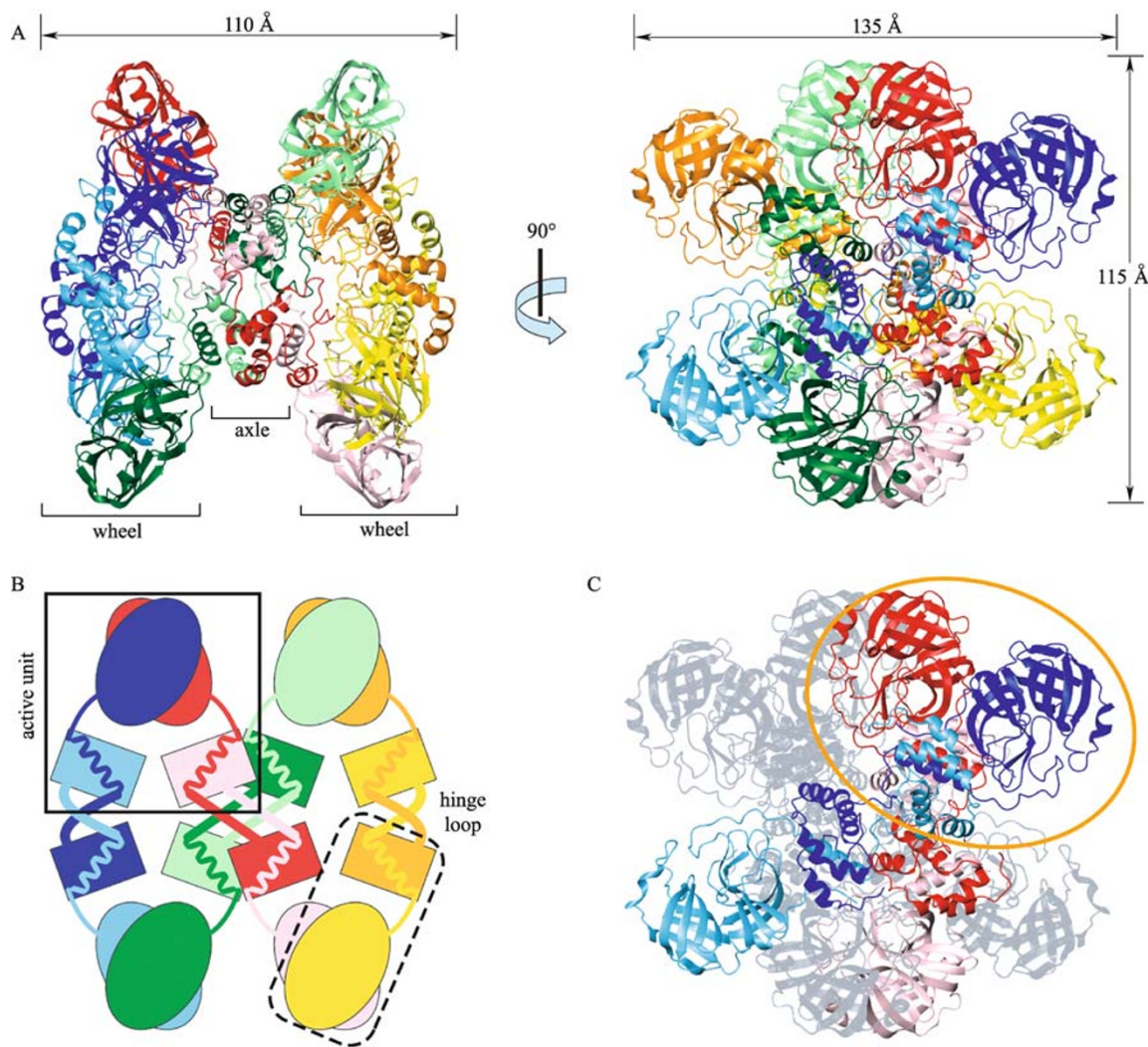


Figure 3. Crystal structure of AO- M^{pro} . (A) Ribbon representations of AO- M^{pro} in two orientations with a 90° rotation. The two protomers of each DSD are colored in red and pink, blue and light blue, dark green and pale green, orange and yellow, respectively. (B) A simplified schematic illustration of the packing of AO- M^{pro} in a similar orientation as the left structure above. The N-terminal domain is represented by an ellipse. The swapped C-terminal domain is represented by a rectangle, while the first helix (remaining helix) of the C-terminal domain is indicated by a spring shape. All eight protomers are colored differently and the color scheme is the same as mentioned in panel A. An active unit is indicated by the black box, and a monomer-like structural unit is marked by the rounded rectangle in broken line. Hinge loops are indicated by thicker lines. (C) Illustration of one inner DSD (red and pink) and one outer DSD (blue and light blue) in ribbon diagram in AO- M^{pro} . The other DSDs are shown in gray dots. An active unit is indicated by the orange ellipse.

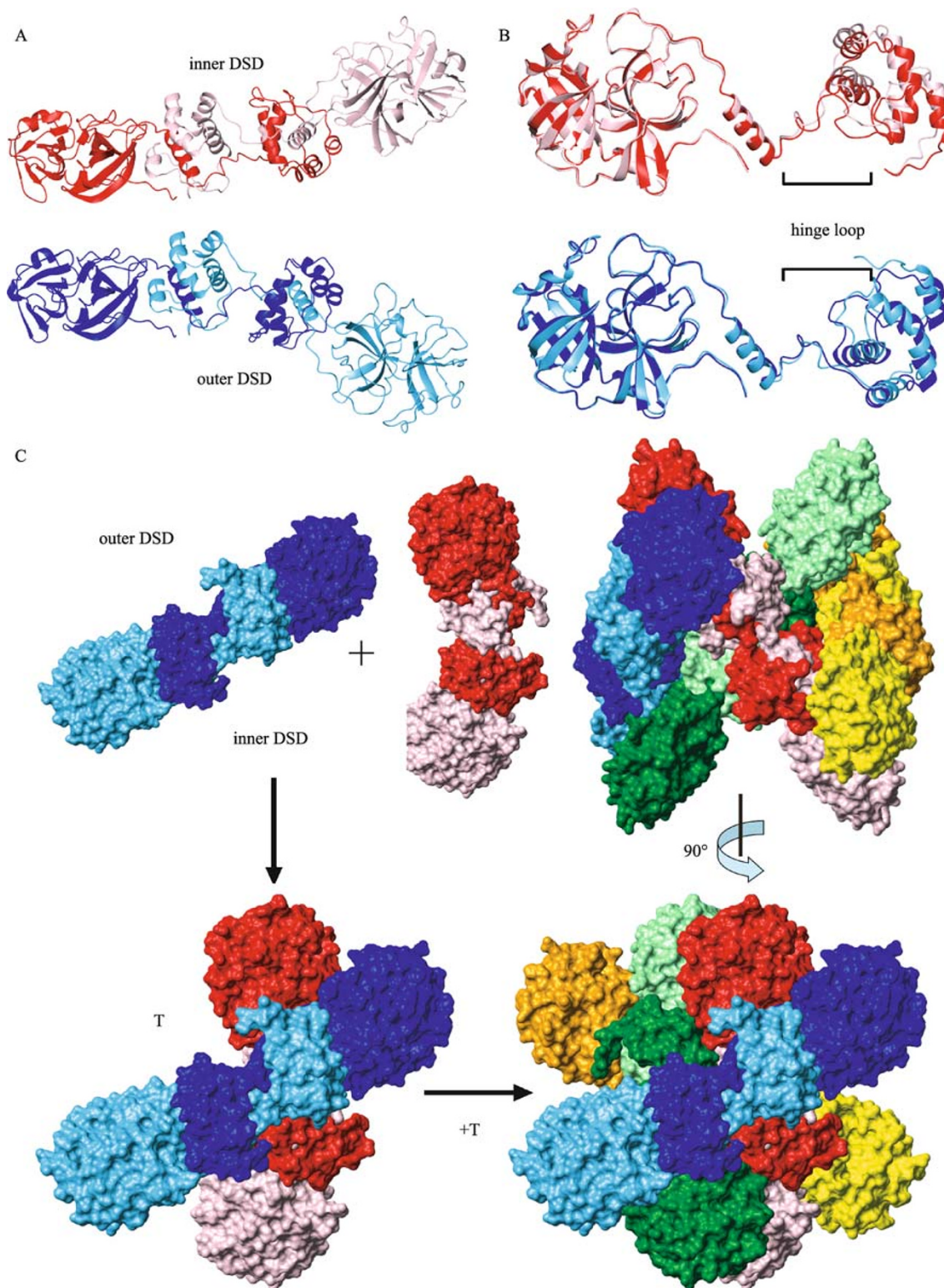


Figure 4. Structural units and assembly of AO-M^{Pro}. (A) Ribbon representations of the two types of DSDs. The upper one is the inner DSD, while the lower one is the outer DSD. The two DSDs are positioned with the two monomer-like structural units on the left side in the same orientation. (B) Ribbon diagrams of the two types of protomers. The upper two protomers are from the inner DSD while the lower two protomers are from the outer DSD. The remaining domains of all protomers are positioned in the same orientation. Hinge loops (residues 214–226) which connect the remaining domains and the swapped domains are indicated. (C) Assembly of AO-M^{Pro} from DSDs. One outer DSD (blue and light blue) and one inner DSD (red and pink) associate into a tetramer, and two tetramers further dimerize through a crystallographic twofold axis to form the octamer.

domains which define the axle region, while two N-terminal domains from both inner DSDs form one wheel together with one outer DSD on either side (Fig. 3A and 3B).

For each DSD, the remaining domain consisting 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, form a monomer-like structural unit (Fig. 3B). Therefore, a DSD can also be viewed as two monomer-like structural units connected by two hinge loops (residues 214–226) that connect the remaining domain and the swapped domain within one protomer (Supplemental Fig. 1). The structures of all eight monomer-like structural units are almost identical, and each adopts essentially the same conformation as that of one protomer in the crystal structure of the M^{Pro} active dimer (Yang et al., 2003; Xue et al., 2007). Although the two protomers within one DSD have very similar conformations, structures of the inner DSD protomers differ greatly from those of the outer DSD protomers. While all eight protomers have almost identical structures for the swapped domain and the remaining domain, the difference lies on the relative orientation of the swapped domain toward the remaining domain in each protomer. Specifically, superimposing the remaining domains, the two swapped domains nearly rotate 180° to each other between the two types of protomers, which is resulted from the conformation difference in the hinge loop region (Fig. 4B). This difference in protomers leads to the conformational difference between inner and outer DSDs, which is characterized by the difference in the relative orientation of the two monomer-like structural units within one DSD (Fig. 4A).

One inner DSD and one outer DSD can associate into a tetramer, in which two monomer-like structural units from both DSDs interact to form an M^{Pro} active dimer-like conformation (active unit) (Fig. 3B, 3C and 4C), and there is no direct contact between the other two “free” monomer-like structural units. This may imply that the tetrameric form of D-M^{Pro}-NE could adopt a similar structure in solution. Two tetramers form an octamer via a crystallographic two-fold axis, in which the four “free” monomer-like structural units form two active units, each consisting of one monomer-like structural unit from an inner DSD of one tetramer and one from an outer DSD of the other tetramer. The four active units in an AO-M^{Pro} have almost identical overall folds with a backbone r.m.s deviation of ~ 0.9 Å for regions of regular secondary structure. Moreover, the conformation of the active unit in AO-M^{Pro} is essentially the same as that of M^{Pro} active dimer (PDB ID: 2HOB) with a backbone r.m.s deviation of ~1.0 Å for secondary structure regions (Fig. 5A).

Consistent with the architecture of the octamer, there are eight active sites in AO-M^{Pro}, and each one is bound with a Michael acceptor inhibitor N3 in similar conformation and orientation. Comparison of the N3-bound active site conformations between AO-M^{Pro} and M^{Pro} active dimer (PDB ID: 2HOB) shows that the active site residues of both active

forms have very similar conformation, especially those directly interacting with the inhibitor N3 (Fig. 5B). In addition, inhibitor N3 also binds active sites of AO-M^{Pro} and M^{Pro} active dimer in similar fashions. All eight active sites are located at the wheel regions of AO-M^{Pro}, and each wheel has four active sites. Two of them are located at the axle side of a wheel, while the other two are located at the exterior side of a wheel (Fig. 5C). The two active sites in one active unit are facing opposite directions. The eight active sites are dispersedly distributed in AO-M^{Pro}, and the shortest distance between two active sites is over 40 Å.

DISCUSSION

Our discovery of AO-M^{Pro} indicates that mature M^{Pro} actually has at least three different quaternary structure forms: an inactive monomeric form in equilibrium with an active dimeric form (Fan et al., 2004), and a super-active octameric form (Fig. 6). All three forms have exactly the same primary sequence. However, different from the M^{Pro} active dimer, AO-M^{Pro} is assembled in a very unique way characterized by 3D domain swapping.

In AO-M^{Pro}, the two protomers of each DSD swap their whole C-terminal domains except the first helices, which results in a large binding interface (~ 5200 Å² compared to ~27,000 Å² of overall molecular surface of one DSD) to stabilize the DSD. As a result, two monomer-like structural units connected by hinge loops are formed in each DSD. Meanwhile, two monomer-like structural units from different DSDs form an active unit through the same interactions as those monomer-like structural unit of the M^{Pro} active dimer. However, different from M^{Pro} active dimer, these interactions occur among four different protomers instead of two for each active unit (Fig. 3B and 3C). Therefore, each active unit is connected to two other active units by hinge loops in AO-M^{Pro}. All these features stabilize the octamer structure and enable the active conformation to be locked. In addition, flexibility of hinge loops is characteristic of 3D domain swapping, which enables the formation of four active units in AO-M^{Pro}. 3D domain swapping is generally regarded as a mechanism for protein oligomerization and aggregation. Here, in SARS-CoV M^{Pro}, 3D domain swapping serves as a mechanism to enhance the dimerization affinity for the active conformation formation and enables the enzyme to be locked in the active conformation and to be constantly active, representing a new role of 3D domain swapping in protein functions.

It has been established that the monomeric form of M^{Pro} is inactive, and homo dimerization is required for the enzyme to adopt an enzymatically active conformation. However, activity of the M^{Pro} active dimer, which is in equilibrium with the inactive monomer, is dependent on both the enzyme concentration and the dissociation constant. Evidently, a critical concentration threshold must be reached before the active dimer is constantly present. In contrast, the activity of

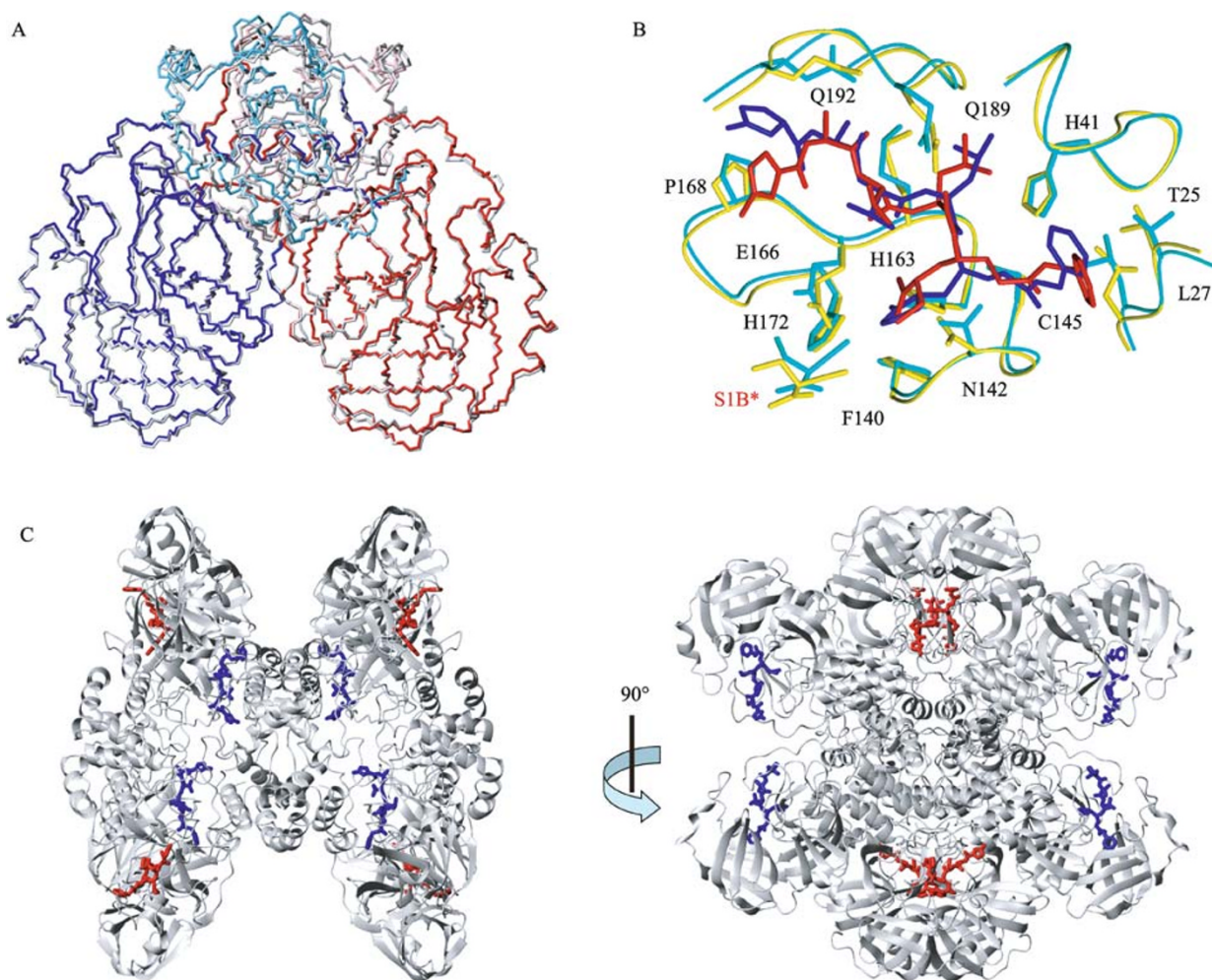


Figure 5. Active conformation and active sites. (A) Superposition of one active unit of AO-M^{PRO} (the same color scheme in Fig. 4C) with M^{PRO} active dimer (2HOB, gray). (B) Comparison of one inhibitor N3 bound active pocket in AO-M^{PRO} (cyan) with that in M^{PRO} active dimer (yellow, PDB ID: 2HOB). The corresponding inhibitors are in red and blue, respectively. (C) Ribbon diagrams of AO-M^{PRO} with eight inhibitors shown in stick display. The four inhibitors located at inside wheel are colored blue, while the other four at outside wheel are colored red.

AO-M^{PRO} only depends on concentration, and it would remain active once formed. As a result, AO-M^{PRO} exhibits significantly higher activity than MD-M^{PRO} at low concentrations.

Moreover, as there are four active units and thus eight active sites in AO-M^{PRO} and the substrates of M^{PRO} have multiple M^{PRO} cleavage sites (7 for pp1a and 11 for pp1ab), the interaction between AO-M^{PRO} and pp1a/pp1ab could very well be a unique case of polyvalent interaction between enzyme and substrate. Polyvalent interactions are characterized by the simultaneous binding of multiple ligands on one biologic entity (a molecule, a surface, etc) to multiple receptors on another (Mammen et al., 1998), and have been implicated to play pivotal roles in some important biologic processes, such as cancer metastasis (Cattaruzza and Perris, 2005), virus and bacterial infections (Mammen et al., 1998), wound healing

(Minor and Peterson, 2002), chromatin modifications (Ruthenburg et al., 2007), and transcription regulations (Mammen et al., 1998). In a polyvalent interaction, the binding affinity between a multivalent ligand and a multivalent receptor increases with the increasing number of ligands and receptors. As the eight active sites are dispersedly distributed and are all accessible, it is quite possible for AO-M^{PRO} to simultaneously recognize and digest multiple cleavage sites of one polyprotein substrate. Thus, compared to M^{PRO} active dimer which has only two active sites, the binding affinity between AO-M^{PRO} and its polyprotein substrates could be much higher, with the potential to increase the proteolytic activity of AO-M^{PRO} toward pp1a/pp1ab. Furthermore, it was reported that M^{PRO} exhibits quite different affinities and activities toward the 11 different M^{PRO} cleavage sequences

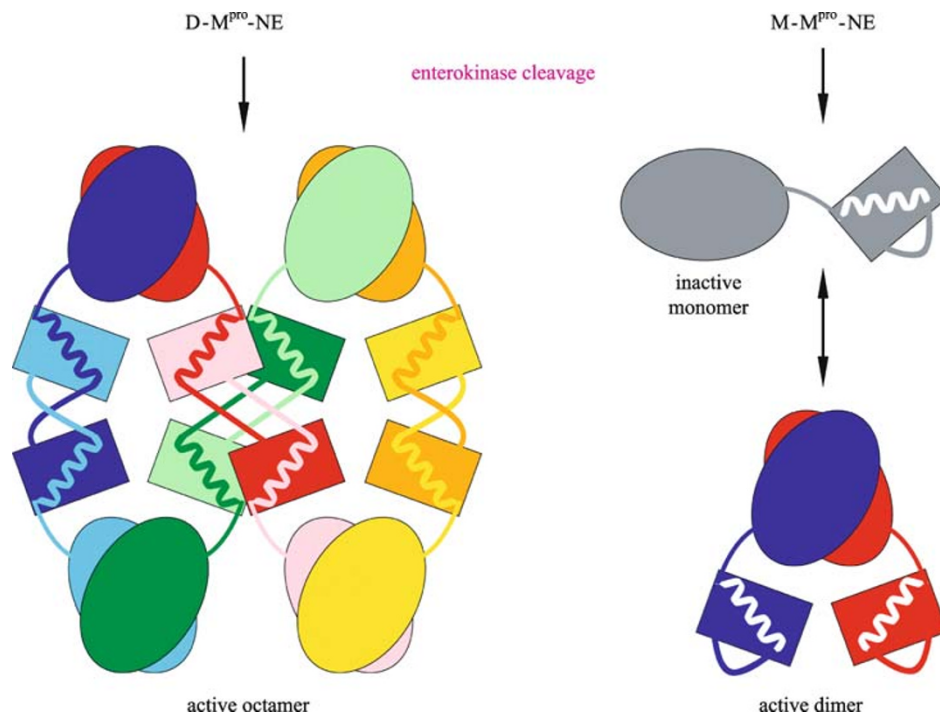


Figure 6. A schematic diagram illustrating the generation of the three different quaternary structure forms of mature M^{pro} .

in the polyprotein substrates (Fan et al., 2004). It is expected that the polyvalent interaction nature and the simultaneous cleavage ability of $AO-M^{\text{pro}}$ should significantly increase the affinity and activity of M^{pro} toward low efficiency cleavage sites by decreasing the dissociation rate (k_{off}) of the enzyme/substrate complex. In addition, with 8 active sites, it is also possible for $AO-M^{\text{pro}}$ to hydrolyze more than one polyprotein substrate at the same time. Therefore, although we have only shown that $AO-M^{\text{pro}}$ has higher activity for a single site substrate than M^{pro} active dimer at low protein concentrations, it is theoretically conceivable that the actual activity of $AO-M^{\text{pro}}$ toward polyprotein substrates could be even much higher than that of the well-known M^{pro} active dimer with only two active sites.

In all crystal structures of SARS-CoV M^{pro} , the C terminus is exposed while the N-finger is buried in the dimer interface with several important interactions (Yang et al., 2003; Lee et al., 2005; Tan et al., 2005; Xu et al., 2005; Xue et al., 2007). Xue et al. (2007) have reported that extra residues to the N terminus of M^{pro} , but not the C terminus, can significantly reduce the activity of SARS-CoV M^{pro} . This is consistent with the low enzymatic activity of $M^{\text{pro}}\text{-NE}$ we observed, which may imply that mature M^{pro} is quite different from M^{pro} in polyproteins prior to self-cleavage. It has been proposed that M^{pro} has to be activated by auto-cleavage from this inactive polyprotein precursor before it can cleave the polyprotein at other cleavage sites (Anand et al., 2005). In a sense, the process for generating mature M^{pro} from $M^{\text{pro}}\text{-NE}$ is somewhat analogous to that of the M^{pro} maturation process, and

$AO-M^{\text{pro}}$ matured from $D-M^{\text{pro}}\text{-NE}$ should be as biologically relevant as the commonly studied forms of the mature M^{pro} .

Although we don't have evidence that $AO-M^{\text{pro}}$ exists in SARS-CoV infected cells, it can be predicted that $AO-M^{\text{pro}}$ should be functionally superior to $MD-M^{\text{pro}}$ for SARS-CoV if it does exist, especially during the initial period of infection when the production level of mature M^{pro} is extremely low. Therefore, it is tempting to speculate that $AO-M^{\text{pro}}$ 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.

SARS-CoV M^{pro} has been extensively investigated via structural and biochemical studies (Anand et al., 2005; Bartlam et al., 2005; Po-Huang, 2006; Yang et al., 2006). Almost all these studies are carried out with mature M^{pro} , and little is clearly known about M^{pro} in the viral polyproteins and its maturation process. Recently, Chen et al. reported that two monomeric mutants of M^{pro} can still perform N-terminal autocleavage, while the dimerization of mature protease and trans-cleavage activity following auto-processing are completely lost (Chen et al., 2010). They proposed that the auto-processing of M^{pro} from polyproteins is through an "intermediate dimer" structure which does not strictly depend on the active conformation existing in mature M^{pro} , and the possible "intermediate dimer" may be formed through C-terminal domain dimerization. It is of interest to find out whether this "intermediate dimer" is related to the $M^{\text{pro}}\text{-NE}$ domain-swapped dimer we observed. The finding by Chen et al. and our discovery of the super-active octameric form of

SARS-CoV main protease indicate that the functional mechanism of this important enzyme may be far more complicated than we thought, and that more studies are necessary to reveal its genuine biologic function mechanism.

MATERIALS AND METHODS

Cloning, protein production, and purification

To construct the expression plasmid of M^{PRO}-NE, the coding sequence for the mature M^{PRO} was amplified by PCR with a forward primer (5'-GGCCGCCATATGGACGACGACGACAAAAGTGGTTTTAG-GAAAATG-3') and a reverse primer (5'-CGCACGATCTCGAGT-TATTGGAAGGTAACACCAG-3'). The PCR product was cloned into the pET28a vector between NdeI and XhoI sites. The resulting plasmid encodes M^{PRO}-NE which has an extra 26-residue fusion peptide (MGSSHHHHHSSGLVPRGSHMDDDDK) including an enterokinase cleavage site right in front of the N terminus of mature M^{PRO}. The plasmid was transformed into the *E. coli* BL21 (DE3) strain for protein production. Cells were grown in LB medium containing 100 µg/mL of kanamycin at 35°C until A₆₀₀ reached 0.8, and were induced with 0.5 mM IPTG. After 8 h, the cells were harvested by centrifugation at 5000 g for 10 min. The pelleted cells were suspended in 50 mM potassium phosphate buffer with 300 mM NaCl at pH 8.0 (buffer A), and then stored at -80°C. After cell lysis by sonication and removal of cell debris by centrifugation at 24,000 g for 20 min, the supernatant was applied onto a Ni-NTA column equilibrated with buffer A, and then washed by 50 mM imidazole in buffer A and further purified with gel-filtration (Superdex 200 10/300 GL) on an ÄKTA fast protein liquid chromatography system (FPLC) (GE, USA) with 50 mM Tris buffer (pH 8.0, 1 mM DTT). Thus, M-M^{PRO}-NE and D-M^{PRO}-NE were completely separated. At a concentration of 1 mg/mL, purified M-M^{PRO}-NE or D-M^{PRO}-NE was digested by enterokinase in a reaction buffer of 50 mM Tris-HCl (pH 8.0, 1 mM CaCl₂, 0.1% Tween-20, 1 mM DTT) at room temperature for 12 h, and then the mature M^{PRO} was purified using Superdex 200 column with 50 mM Tris buffer (pH 7.4, 1 mM EDTA, 1 mM DTT).

Analytical ultracentrifugation

The sedimentation velocity experiments were carried out with a Beckman Coulter ProteomeLab™ XL-I at the Institute of Biophysics, Chinese Academy of Sciences (Beijing, China). All AUC runs were carried out at a speed of 50,000 rpm at 16°C. The sample volume was 400 µL with a protein concentration of ~ 1 mg/mL in 50 mM Tris buffer (pH 7.4, 1 mM EDTA, 0.5 mM DTT). A wavelength of 280 nm was used to record the UV absorption of the cells which were scanned every 5 min for 5 h. The data were analyzed with the SedFit program (version 11.71 from 07/2008).

The dissociation constant of M-M^{PRO}-NE was calculated from the peak area integrals of the monomer (A^M) and dimer (A^D) peaks in AUC analysis data. As the total protein concentration (c) is known, the monomer concentration can be calculated using equation $M = c/(1 + A^D/A^M)$, and the dimer concentration can be calculated as $D = c/2(1 + A^M/A^D)$. Therefore, the dissociation constant can be calculated as $K_d = M^2/D$. The dissociation constant between D-M^{PRO}-NE dimer and tetramer was calculated in the same way.

Enzymatic activity assay

The enzymatic activity was determined by the peptide cleavage assay using a peptide substrate with a sequence of MCA-AVLQSGFR-Lys (Dnp)-Lys-NH₂ (more than 95% purity; GL Biochem (Shanghai) Ltd.). The reaction was monitored with a Hitachi (Tokyo, Japan) F-4500 fluorescence spectrophotometer at 25°C using wavelengths of 320 nm and 405 nm for excitation and emission, respectively. The reaction was initiated by adding varied concentrations of enzyme to a 600 µL system containing 60 µM substrates with the reaction buffer of 50 mM Tris-HCl (pH 7.4) with 1 mM EDTA and 1 mM DTT. When the enzyme concentration is less than 0.34 µg/mL, the reaction was recorded for 10 min. Initial hydrolysis rates were calculated by linearly fitting the linear portion of the data using Microsoft Office Excel 2003. For mature M^{PRO}, enzyme activities were measured for concentrations range from 0.17 to 3.4 µg/mL, and the apparent dimer-monomer dissociation constant K_d values were obtained by fitting the plot of reaction rate versus enzyme concentration to the following equation, assuming the dimer is active and the monomer is inactive:

$$v = A[K_d + 4c - \sqrt{K_d^2 + 8K_d * c}]/8 \quad (1)$$

In the equation, v is the observed hydrolysis rate, A is the hydrolysis rate of the active dimer, and c is the total protein concentration.

NMR sample preparation and NMR spectroscopy

E. coli cells were allowed to grow in M9 minimal medium prepared with D₂O, containing ¹⁵N-labeled ammonium chloride (Marley et al., 2001). The uniformly ¹⁵N, ²H-labeled D-M^{PRO}-NE and M-M^{PRO}-NE proteins were purified by the abovementioned methods. NMR samples of D-M^{PRO}-NE and M-M^{PRO}-NE were at a concentration of about 1 mM, with a buffer containing 50 mM potassium phosphate (pH 7.0), 1 mM EDTA, 0.03% NaN₃, in 90% H₂O/10% D₂O, plus Complete, an EDTA-free protease inhibitor cocktail (Roche, Germany).

All 2D NMR experiments were performed at 298 K on a Bruker Avance 800 MHz spectrometer. NMR spectra were processed with NMRPipe (Delaglio et al., 1995), and analyzed using NMRView (Johnson and Blevins, 1994). The chemical shift in the ¹H dimension was referenced directly to 2,2-dimethyl-2-silapentanesulfonic acid (DSS), whereas the chemical shifts in the ¹⁵N dimensions were indirectly referenced to DSS (Wishart et al., 1995).

Crystallization

The freshly prepared protein was incubated with inhibitor N3 with a molar ratio of 1:1 at 4°C and then concentrated to 30 mg/mL in 20 mM Tris buffer (pH 7.0). Crystallization was performed by the sitting-drop vapor-diffusion method at 16°C in 48-well plates. The initial crystallization conditions were screened using Hampton Research Crystal Screen Kits (Hampton Research Corporation). One microliter of the protein solution was mixed with 1 µL reservoir solution and equilibrated against 100 µL reservoir solution. Small crystals could be found in several conditions within three days. A series of crystallization grids was used to optimize the crystallization condition at 18°C using the hanging-drop vapor-diffusion method by mixing 1 µL AO-M^{PRO} protein in the storage buffer with an equal volume of the

reservoir solution containing 0.2 M sodium chloride, 0.1 M Bis-Tris (pH 5.5) and 25% (w/v) PEG3350. Fine shaped and good quality crystals (size ~30×50×200 μm) suitable for data collection appeared in 14 d and were soaked in a cryo-protectant solution consisting of 4 M sodium formate. Crystals were flash-frozen in liquid nitrogen and then transferred into a dry nitrogen stream at 100K for X-ray diffraction data collection.

Data collection and structure determination

A 3.2 Å resolution diffraction data set was collected at 100 K from a single AO-M^{pro} crystal using an ADSC Q270 detector on the BL17A beamline at Photon Factory, KEK (Japan). A total of 360 frames of data were collected with 0.5° oscillation width. Processing of diffraction images and scaling of the integrated intensities were performed using the HKL2000 software package (Otwinowski and Minor, 1997). The crystal belongs to space group $P4_12_12$, with unit cell parameters $a = b = 161.9 \text{ \AA}$, $c = 166.4 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$. We assumed that there are four M^{pro} molecules in the asymmetric unit, corresponding to a Matthews coefficient of 1.92 and solvent content of 37% (Matthews, 1968). Initial phases were obtained by molecular replacement with PHASER (McCoy et al., 2007) with the crystal

structure of SARS-CoV main protease (PDB code: 2H2Z) as the search model. The final manual structure rebuilding and refinement were performed in COOT (Emsley and Cowtan, 2004), Refmac (Murshudov et al., 1997) and Phenix (Adams et al., 2002) with the guidance of the $2Fo-Fc$ and $Fo-Fc$ density maps. During the later stages of positional refinement, restraints were relaxed and a bulk solvent correction was applied under the guidance of R_{free} . Model geometry was verified using the program PROCHECK (Laskowski et al., 1993). Final refinement statistics are summarized in Table 1. Figures were created by using MOLMOL (Koradi et al., 1996).

ACKNOWLEDGMENTS

This work was supported by Grant No. 2003CB514104 from the National Basic Research Program (973 Program) and Grant No. 30125009 from National Natural Science Foundation of China to Bin Xia, Grant No. 2006AA02A323 from the National Programs for High Technology Research and Development Program (863 Program) to Changwen Jin, and Grant No. 2009ZX09311-001 to Zhiyong Lou from the National Major Projects of China. We gratefully thank Dr. Zengyi Chang for his assistance on gel filtration analysis, and thank Ms. Xiaoxia Yu for assisting the analytical ultracentrifugation experiment.

Table 1 Data collection and refinement statistics

data collection statistics

cell parameters	$a = b = 161.9 \text{ \AA}$, $c = 166.4 \text{ \AA}$ $\alpha = \beta = \gamma = 90^\circ$
space group	$P4_12_12$
wavelength used (Å)	1.000
resolution (Å)	50.0(3.3) ^c – 3.2
No. of all reflections	839,197
No. of unique reflections	37,177
completeness (%)	89.9 (64.6)
average $I/\sigma(I)$	8.4 (2.2)
R_{merge}^a (%)	15.5 (65.2)
refinement statistics	
No. of reflections used ($\sigma(F) > 0$)	33,573
R_{work}^b (%)	22.0
R_{free}^b (%)	27.2
r.m.s.d. bond distance (Å)	0.009
r.m.s.d. bond angle (°)	1.526
average B-value (Å ²)	97.1
ramachandran plot (excluding Pro & Gly)	
Res. in most favored regions	727 (69.8%)
Res. in additionally allowed regions	249 (23.9%)
Res. in generously allowed regions	40 (3.8%)

^a $R_{merge} = \sum_h \sum_l |I_{ih} - \langle I_h \rangle| / \sum_h \sum_l \langle I_h \rangle$, where $\langle I_h \rangle$ is the mean of the observations I_{ih} of reflection h .

^b $R_{work} = \sum (|F_p(obs) - F_p(calc)|) / \sum |F_p(obs)|$; R_{free} = R factor for a selected subset (5%) of the reflections that was not included in prior refinement calculations.

^c Numbers in parentheses are corresponding values for the highest resolution shell.

ABBREVIATIONS

DSD, domain-swapped dimer subunits; DSS, 2,2-dimethyl-2-silapentanesulfonic acid; DTT, 1,4-dithiothreitol; EDTA, ethylene diamine tetraacetic acid; FPLC, fast protein liquid chromatography system; HSQC, heteronuclear single quantum coherence; MR, molecular replacement; M^{pro}, main protease; NMR, nuclear magnetic resonance; RMSD, root mean square deviation; SARS, severe acute respiratory syndrome; SARS-CoV, SARS coronavirus; WT, wild-type

REFERENCES

- Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.-W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D* 58, 1948–1954.
- Anand, K., Palm, G.J., Mesters, J.R., Siddell, S.G., Ziebuhr, J., and Hilgenfeld, R. (2002). Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra α -helical domain. *EMBO J* 21, 3213–3224.
- Anand, K., Yang, H., Bartlam, M., Rao, Z. & Hilgenfeld, R. (2005). Coronavirus main proteinase: target for antiviral drug therapy. In: *Coronaviruses with special emphasis on first insights concerning SARS*, A. Schmidt, M.H. Wolff, and O.F. Weber, ed. (Switzerland, Basel; Birkhauser Verlag). pp. 173–199.
- Anand, K., Ziebuhr, J., Wadhvani, P., Mesters, J.R., and Hilgenfeld, R. (2003). Coronavirus main proteinase (3CL^{pro}) structure: basis for design of anti-SARS drugs. *Science* 300, 1763–1767.
- Bartlam, M., Yang, H., and Rao, Z. (2005). Structural insights into SARS coronavirus proteins. *Curr Opin Struct Biol* 15, 664–672.
- Cattaruzza, S., and Perris, R. (2005). Proteoglycan control of cell movement during wound healing and cancer spreading. *Matrix Biol* 24, 400–417.
- Chan, H.L., Tsui, S.K., and Sung, J.J. (2003). Coronavirus in severe acute respiratory syndrome (SARS). *Trends Mol Med* 9, 323–325.
- 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, 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., Hu, T., Zhang, J., Chen, J., Chen, K., Ding, J., Jiang, H., and Shen, X. (2008). Mutation of Gly11 on the dimer interface results in the complete crystallographic dimer dissociation of SARS-CoV 3CL^{pro}: Crystal structure with molecular dynamics simulations. *J Biol Chem* 283, 554–564.
- Chen, S., Jonas, F., Chen, C., and Higenfiled, 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.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6, 277–293.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126–2132.
- 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.
- Gronenborn, A.M. (2009). Protein acrobatics in pairs—dimerization via domain swapping. *Curr Opin Struct Biol* 19, 39–49.
- Hsu, W.C., Chang, H.C., Chou, C.Y., Tsai, P.J., Lin, P.I., and Chang, G.G. (2005). Critical assessment of important regions in the subunit association and catalytic action of the severe acute respiratory syndrome coronavirus main protease. *J Biol Chem* 280, 22741–22748.
- 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.
- Ivanov, D., Tsodikov, O.V., Kasanov, J., Ellenberger, T., Wagner, G., and Collins, T. (2007). Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc Natl Acad Sci U S A* 104, 4353–4358.
- Johnson, B.A., and Blevins, R.A. (1994). NMR View: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4, 603–614.
- Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics* 14, 51–55.
- 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.-A., and Liang, P.-H. (2004). Characterization of SARS main protease and inhibitor assay using a fluorogenic substrate. *Biochem Biophys Res Commun* 318, 862–867.
- Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26, 283–291.
- Lee, T.W., Cherney, M.M., Huitema, C., Liu, J., James, K.E., Powers, J.C., Eltis, L.D., and James, M.N. (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.
- Leng, Q., and Bentwich, Z. (2003). A novel coronavirus and SARS. *N Engl J Med* 349, 709.
- Libonati, M., Gotte, G., and Vottariello, F. (2008). A novel biological actions acquired by ribonuclease through oligomerization. *Curr Pharm Biotechnol* 9, 200–209.
- 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.

- Liu, Y., and Eisenberg, D. (2002). 3D domain swapping: as domains continue to swap. *Protein Sci* 11, 1285–1299.
- Mammen, M., Choi, S.-K., and Whitesides, G.M. (1998). Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew Chem Int Ed* 37, 2754–2794.
- Marley, J., Lu, M., and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* 20, 71–75.
- Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491–497.
- McCoy, A., Grosse-Kunstleve, R., Adams, P., Winn, M., Storoni, L., and Read, R. (2007). Phaser crystallographic software. *J Appl Cryst* 40, 658–674.
- Minor, K.H., and Peterson, C.B. (2002). Plasminogen activator inhibitor type 1 promotes the self-association of vitronectin into complexes exhibiting altered incorporation into the extracellular matrix. *J Biol Chem* 277, 10337–10345.
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53, 240–255.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Macromolecular Crystallography, part A*, C.W. Carter Jr., and R.M. Sweet, eds. (Academic Press), pp. 307–326.
- Po-Huang, L. (2006). Characterization and inhibition of SARS-coronavirus main protease. *Curr Top Med Chem* 6, 361–176.
- Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* 8, 983–994.
- 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.
- 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.
- 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., Verschuere, 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.
- Verschuere, K.H.G., Pumpor, K., Anemüller, S., Chen, S., Mesters, J.R., Hilgenfeld, R., (2008). A structural view of the inactivation of the SARS coronavirus main proteinase by benzotriazole esters. *Chem Biol* 15, 597–606.
- 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.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L., and Sykes, B.D. (1995). ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J Biomol NMR* 6, 135–140.
- Xu, T., Ooi, A., Lee, H.C., Wilmoth, 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.
- Yamasaki, M., Li, W., Johnson, D.J., and Huntington, J.A. (2008). Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* 455, 1255–1258.
- 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.
- 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., Bartlam, M., and Rao, Z. (2006). Drug design targeting the main protease, the Achilles' heel of coronaviruses. *Curr Pharm Des* 12, 4573–4590.
- Zhong, N., Zhang, S., Zou, P., Chen, J., Kang, X., Li, Z., Liang, C., Jin, C., and Xia, B. (2008). Without its N-finger, SARS-CoV main protease can form a novel dimer through its C-terminal domain. *J Virol* 82, 4227–4234.
- 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.