

Foldon unfolding mediates the interconversion between M^{pro}-C monomer and 3D domain-swapped dimer

Xue Kang^{a,b}, Nan Zhong^{a,b}, Peng Zou^{a,b}, Shengnan Zhang^{a,b}, Changwen Jin^{a,b,c}, and Bin Xia^{a,b,c,1}

^aBeijing Nuclear Magnetic Resonance Center, Peking University, Beijing 100871, China; ^bCollege of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China; and ^cSchool of Life Sciences, Peking University, Beijing 100871, China

Edited by S. Walter Englander, The University of Pennsylvania, Philadelphia, PA, and approved August 7, 2012 (received for review March 28, 2012)

The C-terminal domain (M^{pro}-C) of SARS-CoV main protease adopts two different fold topologies, a monomer and a 3D domain-swapped dimer. Here, we report that M^{pro}-C can reversibly interconvert between these two topological states under physiological conditions. Although the swapped α_1 -helix is fully buried inside the protein hydrophobic core, the interconversion of M^{pro}-C is carried out without the hydrophobic core being exposed to solvent. The 3D domain swapping of M^{pro}-C is activated by an order-to-disorder transition of its C-terminal α_5 -helix foldon. Unfolding of this foldon promotes self-association of M^{pro}-C monomers and functions to mediate the 3D domain swapping, without which M^{pro}-C can no longer form the domain-swapped dimer. Taken together, we propose that there exists a special dimeric intermediate enabling the protein core to unpack and the α_1 -helices to swap in a hydrophobic environment, which minimizes the energy cost of the 3D domain-swapping process.

NMR | protein folding

3D domain swapping is a unique mechanism for protein dimerization or oligomerization (1, 2). Through exchanging identical structure elements, two or more molecules of the same protein can form a dimer or higher order oligomer with tight binding interface. In recent years, more and more evidences revealed that 3D domain swapping is involved in protein function regulation (2, 3). Most importantly, 3D domain swapping is indicated to be a mechanism for proteins to form aggregates, fibrils, and amyloids, some of which are associated with neurodegenerative diseases (4, 5).

Most domain-swapped dimers were obtained through nonphysiological treatments (2), such as low pH, high temperature, crystallization, lyophilization, or freezing. In some cases, slow interconversion (days to months) between a monomeric protein and its domain-swapped dimer was observed under physiological or nondenaturing conditions (6–12), while the process could be speed up by changing conditions favoring unfolding (2, 12). For example, the monomer and domain-swapped dimer of cyanovirin-N was reported to be in apparent equilibrium (half conversion time is about 10 h) at 38 °C while its T_m is 60 °C (8). Similar phenomena were observed for GB1 (11), and p13suc1 (12), etc. Although the mechanism for 3D domain swapping remains unclear, it is commonly believed that unfolding of the monomeric protein is necessary for it to release the domains to be swapped (13). For proteins that swap an independently folded domain, it is proposed that the folded monomer is first converted to an “open monomer” in which the swapped domain is detached from the rest of the protein, followed by domain-swapped dimer formation (7, 14, 15). As most proteins swap only a few secondary structural elements that do not fold into an independent domain, it is proposed that these proteins must unfold substantially during the domain swapping process (12, 16–18). For proteins to undergo spontaneous interconversion under physiological or nondenaturing conditions (6–12), either full unfolding of the protein (6, 19) or partial unfolding of the swapped elements (8, 11)

was suggested. However, as there is normally a huge energy barrier between the monomer and domain-swapped dimer, it is still not clear how these proteins can spontaneously unfold and pass the energy barrier (20).

We have previously reported that expression of the C-terminal domain of SARS-CoV main protease (M^{pro}-C, residues 187–306) alone in *E. coli* yields both a monomeric form and a domain-swapped dimeric form due to exchange of α_1 -helices (Fig. 1) (21). We have also demonstrated that the mature protein of the main protease can assemble into a super-active stable octamer due to 3D domain swapping of its C-terminal domain (22). This stable octameric main protease is constantly active since each subunit is locked into the active conformation, unlike the active dimeric enzyme, which is in equilibrium with the inactive monomer. Here, we report that M^{pro}-C can reversibly interconvert between monomeric and domain-swapped dimeric forms under physiological conditions (37 °C, pH 7.0). We demonstrate that the interconversion of M^{pro}-C is controlled by an order-to-disorder transition of its C-terminal α_5 -helix foldon, which mediates the 3D domain swapping process without the protein hydrophobic core being exposed to solvent.

Results

Reversible Interconversion between M^{pro}-C Monomer and Domain-swapped Dimer. Although both the monomer and domain-swapped dimer (DS-dimer) of M^{pro}-C are relatively stable at room temperature with no obvious interconversion in days (21), we found that an equilibrium between the monomer and DS-dimer is established in a few hours when the temperature is raised to 37 °C, starting from either monomer or DS-dimer (Fig. 2*A* and *B*). The observed mass ratios between the monomer and DS-dimer at equilibrium are almost the same (approximately 9:1) in the temperature range of 33–39 °C at a protein concentration of 0.5 mM. The equilibrium dissociation constant (K_d) is determined to be about 8 mM, which can be translated into a Gibbs free energy of dissociation of about 12 kJ/mol at 37 °C based on the Van't Hoff equation.

The rate of interconversion is highly dependent on temperature, and the half conversion time ($t_{1/2}$) is about 2 h at 37 °C and about 20 h at 33 °C (Fig. 2*C* and *D*). Apparently, an increase of temperature by 4 °C results in over 10-fold increase of the conversion rate (Table S1). Accordingly, the interconversion rate at

Author contributions: X.K., C.J., and B.X. designed research; X.K., N.Z., P.Z., and S.Z. performed research; X.K. and B.X. analyzed data; and X.K. and B.X. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates for the NMR ensembles of M^{pro}-C at 25 °C in 2.5 M urea have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2LIZ), and the NMR chemical shifts have been deposited in the BioMagResBank, www.bmr.bwisc.edu (accession no. 17911).

¹To whom correspondence should be addressed. E-mail: binxia@pku.edu.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1205241109/-DCSupplemental.

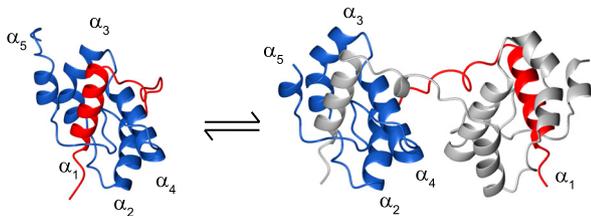


Fig. 1. Illustration of the two forms of $M^{\text{pro-C}}$. Ribbon diagram of (A) solution structure of $M^{\text{pro-C}}$ monomer at 25 °C (PDB code 2K7X), and (B) crystal structure of $M^{\text{pro-C}}$ DS-dimer (PDB code 3EBN) with secondary structure elements annotated. The swapped element α_1 -helix was highlighted in red.

room temperature (25 °C) is predicted to be over 1,000 times slower with $t_{1/2}$ to be months. It is interesting to notice that the mass ratio between the purified monomer and DS-dimer expressed in *E. coli* is usually about 3:2 at different temperatures (18, 25, and 35 °C) (23). It has been suggested that some other factors in the cells may play important roles in facilitating protein oligomerization and stabilizing protein oligomers, such as macromolecular crowding, high local protein concentration, and/or molecular chaperons (1, 2, 24), which remains to be investigated in the future.

Analysis using the transition state theory revealed that the Gibbs free energy of activation is 88 kJ/mol for the 3D domain-swapped dimerization and 99 kJ/mol for the DS-dimer dissociation, while the enthalpy of activation is 374 kJ/mol for dimeriza-

tion and 436 kJ/mol for dimer dissociation, indicating there is a huge energy barrier between the two states (Fig. S1 A–C). The size of this energy barrier is consistent with that of other similar systems, such as stefin A (6) and cyanovirin-N (19). Besides increasing temperature, the interconversion process can also be speed up with the addition of urea (Fig. S1D). In the presence of 2.5 M urea, the $t_{1/2}$ is about 23 h at 25 °C, close to that of 33 °C (Table S1). We tried to investigate the mechanism behind the dramatic acceleration of the interconversion by temperature, as chemical reaction rate is normally increased only by 2–3 times with a 10 °C temperature increase in this temperature range. For the convenience of description, we name 25 °C as “unswappable condition” and 37 °C or 25 °C with 2.5 M urea (25 °C/urea) as “swappable condition.”

Since the swapped α_1 -helix of $M^{\text{pro-C}}$ is buried in the protein hydrophobic core, we wondered whether elevated temperature or urea could cause unfolding of $M^{\text{pro-C}}$. We thus analyzed the thermal stability of $M^{\text{pro-C}}$ using far-UV circular dichroism (CD) spectroscopy (Fig. 2 E and F), and the results showed that the apparent thermal denaturation midpoint (T_m) for $M^{\text{pro-C}}$ monomer and DS-dimer are 59.1 and 59.4 °C, respectively. The denaturation free energy ΔG_U at 25 °C was calculated to be 44.1 kJ/mol for the monomer, noting that the conversion of monomer into DS-dimer could be neglected since the sample concentration is very low (1.4 μM) (Fig. 2F). In contrast, the DS-dimer would firstly dissociate into monomers as the temperature is increased during the thermal unfolding experiment, and the monomer would unfold subsequently. Therefore, the thermal denaturation data of the DS-dimer should mainly reflect unfolding of the monomeric protein, consistent with our observation that the DS-dimer has similar apparent T_m value as the monomer. CD spectrum of the monomer at 37 °C showed a small decrease in the negative ellipticity at 222 nm compared with that at 25 °C, which can be interpreted as a loss of helicity by approximately 4% (Fig. 2E). The stability of monomeric $M^{\text{pro-C}}$ was decreased in the presence of 2.5 M urea with a T_m of 57.1 °C, and the CD spectrum at 25 °C/urea also showed a small decrease in the negative ellipticity at 222 nm. It should be noted that complete unfolding of $M^{\text{pro-C}}$ could not be achieved with urea. Over half of $M^{\text{pro-C}}$ is still folded in the presence of 8 M urea. All these indicate that both forms of $M^{\text{pro-C}}$ are stable and still maintain most of their secondary structures under the swappable conditions.

Consistently, 2D ^1H - ^{15}N HSQC spectra of $M^{\text{pro-C}}$ monomer at 37 °C or 25 °C/urea are overlapped relatively well with that at 25 °C (Fig. 3C), with maximal combined NH chemical shift difference ($\Delta\delta_{\text{comb}}$) to be 0.06 ppm, indicating that the monomer in either swappable condition still maintains the overall fold as that in the unswappable condition. Under both swappable conditions, residues with $\Delta\delta_{\text{comb}}$ over 0.04 ppm are mainly located at L_2 loop, α_3 -helix and α_5 -helix, suggesting that there may be small conformational changes in these regions (Fig. 3 E and F). Both CD and NMR analysis demonstrated that $M^{\text{pro-C}}$ is well folded under swappable conditions.

According to the pre-existing equilibrium theory (25), the native folded state of a protein is in equilibrium with unfolded state, although the population of unfolded state is quite low. Based on ΔG_U (32.1 kJ/mol at 37 °C) of $M^{\text{pro-C}}$ monomer, the amount of the fully unfolded state in 0.5 mM monomer can be estimated to be 2.2×10^{-6} mM at 37 °C (20). Using the diffusion-limited molecular association rate constant of $7 \times 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$ (26), the dimerization rate at 37 °C can be roughly calculated to be 10^{-8} s^{-1} , which results in a conversion $t_{1/2}$ of approximately 2 mo at 37 °C, let alone the typical protein association rate constant is only $10^5 \sim 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$ (27). Therefore, it is thermodynamically impossible that the $M^{\text{pro-C}}$ domain swapping process is undertaken through a fully unfolded intermediate.

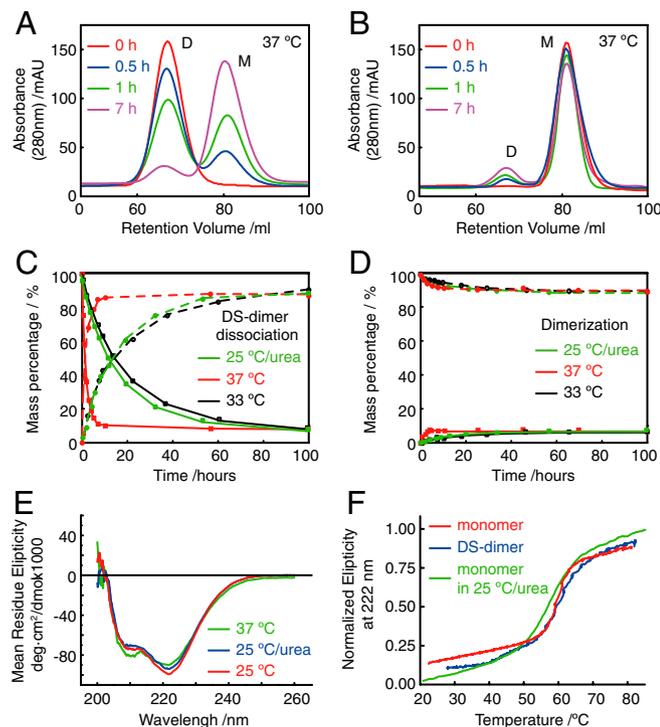


Fig. 2. Interconversion between $M^{\text{pro-C}}$ monomer and DS-dimer and thermal stability of $M^{\text{pro-C}}$. (A) Elution profiles of $M^{\text{pro-C}}$ DS-dimer incubated at 37 °C at 0 h (red), 0.5 h (blue), 1 h (green), and 7 h (purple). (B) Elution profiles of $M^{\text{pro-C}}$ monomer incubated at 37 °C at 0 h (red), 0.5 h (blue), 1 h (green), and 7 h (purple). (C) Mass percentage of the monomer (solid line) and DS-dimer (broken line) as a function of time for the conversion of $M^{\text{pro-C}}$ DS-dimer into monomer at 37 °C (red), 33 °C (black), and 25 °C/urea (green). (D) Mass percentage of the monomer (solid line) and DS-dimer (broken line) as a function of time for the conversion of $M^{\text{pro-C}}$ monomer into DS-dimer at 37 °C (red), 33 °C (black), and 25 °C/urea (green). (E) Comparison of CD spectra of monomeric $M^{\text{pro-C}}$ at 25 °C (red), 37 °C (green), and 25 °C/urea (blue). (F) Thermal denaturation curves of $M^{\text{pro-C}}$ monomer (red), DS-dimer (blue), and monomer in 2.5 M urea (green).

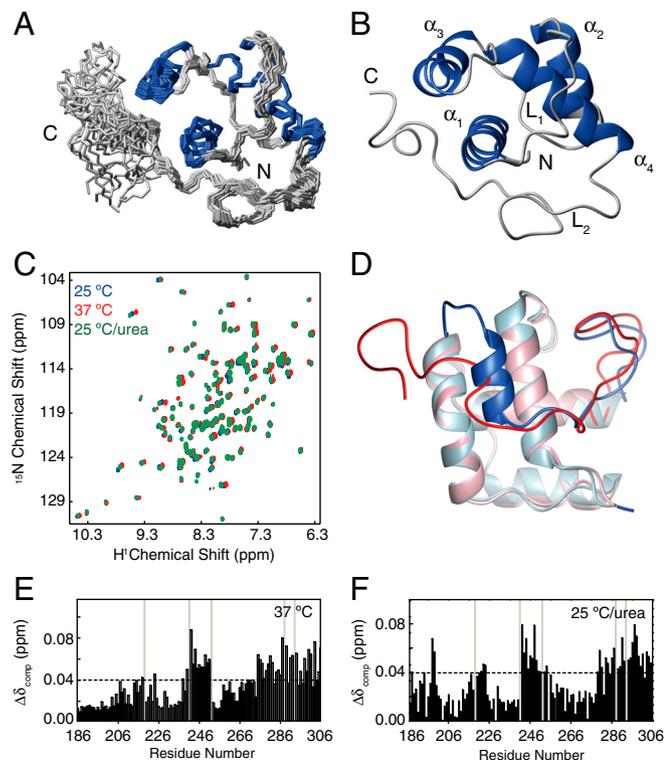


Fig. 3. Structural characterization of $M^{\text{pro-C}}$ monomer under different solution conditions. (A) Backbone traces of 20 representative structures and (B) ribbon diagram of the mean structure of $M^{\text{pro-C}}$ monomer in 25°C/urea (PDB 2LIZ). (C) Overlay of 2D ^1H - ^{15}N HSQC spectra of monomeric $M^{\text{pro-C}}$ at 25°C (blue), 37°C (red), and 25°C/urea (green). (D) Superimposition of $M^{\text{pro-C}}$ monomer structures at 25°C (light blue) and at 25°C/urea (pink) in ribbon diagram. The intact α_5 -helix is in blue, and the disordered “ α_5 -helix” is in red. (E) and (F) Chemical shift differences of monomeric $M^{\text{pro-C}}$ at 25°C with 37°C (E) and 25°C/urea (F), respectively.

Local Conformational Stability of $M^{\text{pro-C}}$ Monomer. Based on structures of the monomer and DS-dimer of $M^{\text{pro-C}}$, it is apparent that the interconversion between the two forms should at least involve unpacking of the protein hydrophobic core and release of the α_1 -helix. We wondered if α_1 -helix could transiently expose to the solvent during the interconversion without fully unfolding the protein. As NMR hydrogen exchange experiment is one of the most effective techniques to monitor protein structure fluctuation, we measured and compared the H/D exchange rates (k_{ex}) for $M^{\text{pro-C}}$ monomer at 25 and 37°C (Fig. S2 and Table S2).

At 25°C, we were able to obtain k_{ex} for 54 out of 120 residues, ranging from 10^{-1} min^{-1} to 10^{-6} min^{-1} , while NH exchange rates for residues L205, A206, L208 and Y209 were too slow to be accurately determined. No amide exchanging too fast to allow the determination of k_{ex} is located on α_1 - α_4 helices. However, k_{ex} was only determined for residue V296, while all other amides on α_5 -helix exchange too fast, indicating that α_5 -helix has relative lower stability than the other helices.

Protein local conformational stability may be evaluated by the protection factor (PF) of each residue or by the corresponding apparent free energy (ΔG_{H-D}) (Table S2). Based on PF, the ranking of conformational stability for secondary structure elements of monomeric $M^{\text{pro-C}}$ is: $\alpha_1 > \alpha_4 > \alpha_2 > \alpha_3 > \alpha_5$. As expected, α_1 -helix is the most protected one. Residues with the apparent free energy ΔG_{H-D} (42 ~ 47 kJ/mol) comparable to the denaturation free energy ΔG_U (44.1 kJ/mol) at 25°C are all located on α_1 -helix, suggesting that full unfold is required for these amides to exchange at 25°C.

At 37°C, the swappable condition, k_{ex} of 42 backbone NHs could be measured, while amides of residues L208 and Y209 still

exchange too slowly. Amides of 13 residues located mainly around loops or turns exchange too fast at 37°C, and their k_{ex} can no longer be measured ($>10^{-1} \text{ min}^{-1}$). Comparison of PF reveals that the ranking of conformational stability for secondary structure elements remains the same as that of 25°C. ΔG_{H-D} values of residues on α_1 -helix range from 31 to 34 kJ/mol, which are close to ΔG_U (32.1 kJ/mol) at 37°C and much higher than ΔG_{H-D} values of residues on α_2 - α_4 helices, indicating that the swapped α_1 -helix is still well protected in the hydrophobic core. Therefore, there should be no transient exposure of the protein core to solvent under the swappable condition, which is the same as the case at 25°C. Thus, there must be a special mechanism to enable two monomeric molecules to switch their α_1 -helices without exposing them to solvent.

Unexpectedly, the amide of V296 on α_5 -helix with k_{ex} of $7.6 \times 10^{-5} \text{ min}^{-1}$ at 25°C showed no NH signal even in the first spectrum at 37°C, suggesting it becomes exchanging with solvent very rapidly. This indicates a dramatic decrease in the conformational stability of α_5 -helix. Coincidentally, α_5 -helix also shows significant NH chemical shift changes in swappable conditions, which might suggest that α_5 -helix may play a role in the domain swapping process of $M^{\text{pro-C}}$.

Structure of $M^{\text{pro-C}}$ Monomer Under Swappable Conditions. We determined the solution structure of $M^{\text{pro-C}}$ monomer in 2.5 M urea at 25°C (PDB 2LIZ), which should represent the structure of $M^{\text{pro-C}}$ that is “active” for domain swapping ($M^{\text{pro-C}}_{\text{ds}}$) (Table S3). The previously reported solution structure (PDB 2K7X) of $M^{\text{pro-C}}$ monomer at 25°C without urea can be viewed as an “inactive” structure, which is unswappable ($M^{\text{pro-C}}_{\text{nds}}$).

The overall fold of $M^{\text{pro-C}}_{\text{ds}}$ is still similar to that of $M^{\text{pro-C}}_{\text{nds}}$, except that α_5 -helix (residues P293-C300) in $M^{\text{pro-C}}_{\text{nds}}$ becomes unstructured in $M^{\text{pro-C}}_{\text{ds}}$, consistent with the chemical shift index analysis results (Fig. 3A, B and D, Fig. S3). The other four α -helices (α_1 , residues 201–214; α_2 , residues 227–234; α_3 , residues 244–257; α_4 , residues 261–271) and one loop (L_1 , residues 216–226) of $M^{\text{pro-C}}_{\text{ds}}$, are essentially the same as those in $M^{\text{pro-C}}_{\text{nds}}$. Backbone heavy atom rmsd between $M^{\text{pro-C}}_{\text{ds}}$ and $M^{\text{pro-C}}_{\text{nds}}$ is 0.44 Å for the four α -helices. Residues P293–C300 that form α_5 -helix in $M^{\text{pro-C}}_{\text{nds}}$ become disordered in $M^{\text{pro-C}}_{\text{ds}}$, and form a long flexible tail together with residues S301-Q306.

We also characterized the structure of $M^{\text{pro-C}}$ monomer at 37°C with NMR, and analysis of 3D ^{15}N -NOESY-HSQC spectrum revealed that no NOE characteristic for α -helical structure is observed for residues P293–C300. The number of NOEs observed at 37°C for this region is much less than that in $M^{\text{pro-C}}_{\text{nds}}$ (Fig. S4). These indicate that α_5 -helix in $M^{\text{pro-C}}_{\text{nds}}$ also becomes disordered and flexible at 37°C, similar to the situation in 25°C/urea.

As the disruption of α_5 -helix is the common structural change for both swappable conditions, we speculated that the dramatic acceleration of the interconversion rate may be related to unfolding of α_5 -helix. At 25°C, the unswappable condition, α_5 -helix is well defined and is positioned in parallel to α_1 - and α_3 -helices, making extensive contacts to both. It seems that α_5 -helix in $M^{\text{pro-C}}_{\text{nds}}$ may serve as a “lock” to block the movement of the swapped α_1 -helix. Under swappable conditions, it is possible that the disruption of α_5 -helix in $M^{\text{pro-C}}_{\text{nds}}$ makes it much easier for α_1 -helix to displace the disordered “ α_5 -helix” in $M^{\text{pro-C}}_{\text{ds}}$ and get “unlocked,” which enables and accelerates the domain swapping process.

Disordered “ α_5 -helix” Required for $M^{\text{pro-C}}$ Domain Swapping. If α_5 -helix in $M^{\text{pro-C}}_{\text{nds}}$ does serve as a “lock,” we expected that removal of α_5 -helix would greatly expedite the interconversion process. We thus generated a C-terminal truncation mutant (residues 187–292, $M^{\text{pro-C}}_{\text{C}292}$) of $M^{\text{pro-C}}$ with residues 293–306 (including α_5 -helix) fully removed. To our surprise, $M^{\text{pro-C}}_{\text{C}292}$ expressed in *E. coli* showed only one single peak eluted at 83.9 ml on

gel-filtration column at 25 °C, corresponding to a monomer (Fig. S5A). No dimer could be observed after M^{pro}-C²⁹² was either incubated at 37 °C or treated with 2.5 M urea at 25 °C for hours, indicating that M^{pro}-C²⁹² is a monomer which cannot convert into DS-dimer. This result suggests that the disordered “α₅-helix” of M^{pro}-C is required for the conversion of monomer to DS-dimer.

We therefore generated six additional truncation mutants, M^{pro}-C²⁹³, M^{pro}-C²⁹⁴, M^{pro}-C²⁹⁵, M^{pro}-C²⁹⁶, M^{pro}-C²⁹⁸, and M^{pro}-C³⁰¹, in which fewer C-terminal residues are removed (Table 1). Among these mutants, M^{pro}-C²⁹³ only exists as a monomer, while all other 5 truncation mutants of M^{pro}-C were produced as both monomer and DS-dimer forms in *E. coli*. The monomer to DS-dimer conversion rates for mutants M^{pro}-C³⁰¹, M^{pro}-C²⁹⁸, M^{pro}-C²⁹⁶, and M^{pro}-C²⁹⁵ are in increasing order. The more C-terminal residues removed, the faster the conversion rate is. However, when residue D295 is further removed from M^{pro}-C²⁹⁵, the resulting mutant M^{pro}-C²⁹⁴ is converting much slower than M^{pro}-C²⁹⁵ (Table 1). Overlay of 2D ¹H-¹⁵N HSQC spectra of monomeric M^{pro}-C²⁹³, M^{pro}-C²⁹⁴, and M^{pro}-C²⁹⁵ shows that most of the NH signals are well overlapped, indicating that there is no major structural difference among these three truncation mutants (Fig. S5D). Moreover, the presence of residue F294 is vital for the conversion of M^{pro}-C²⁹⁴ from monomer into DS-dimer, as further removal of residue F294 from M^{pro}-C²⁹⁴ results in M^{pro}-C²⁹³ no longer possessing the ability of domain swapping. Meanwhile, residue D295 is very important for M^{pro}-C monomer converting into DS-dimer, since including residue D295 renders the conversion rate of M^{pro}-C²⁹⁵ to be approximately 4 times faster than that of M^{pro}-C²⁹⁴. Taken together, these data indicate that the disordered “α₅-helix” of M^{pro}-C^{ds} under swappable conditions should also play a crucial role in mediating the 3D domain swapping.

To rule out the possibility that the change of conversion rate for the truncation mutants is simply due to protein thermal stability change, we also performed thermal denaturation study for all truncation mutants (Table 1). The results showed that all C-terminal truncation mutants are less stable than the wild-type M^{pro}-C, and *T_m* values are lowered by 3 °C to 12 °C. However, although *T_m* values of M^{pro}-C²⁹⁵ (51.6 °C) and M^{pro}-C²⁹⁶ (51.8 °C) are quite similar, the conversion rate of M^{pro}-C²⁹⁵ (629 mM⁻¹ s⁻¹) is over 4 times faster than that of M^{pro}-C²⁹⁶ (151 mM⁻¹ s⁻¹). On the other hand, mutants M^{pro}-C²⁹⁴ and M^{pro}-C²⁹⁶ have similar conversion rates, but *T_m* of M^{pro}-C²⁹⁴ is 3.5 °C lower than that of M^{pro}-C²⁹⁶. Therefore, the difference in the conversion rate for the truncation mutants cannot be attributed to the change of thermal stability.

In the structures of M^{pro}-C monomer and DS-dimer, side-chains of both F294 and D295 are largely exposed to solvent without much contact with other parts of the protein (Fig. S5E). We also studied the effect of point mutation for these two residues on the interconversion between M^{pro}-C monomer and DS-dimer. As shown in Table 1, mutation D295K results in about 20-fold increase of the conversion rate and about 7 °C decrease in *T_m*, while mutations F294A and F294D stabilize the protein and slow down the conversion process by 6 and 8 times, respectively. However, mutation D295A destabilizes the protein, but causes approximately 40% decrease in the conversion rate. Double mutation F294A/D295A decreases *T_m* by 2 °C, but has little effect on the conversion rate. These results confirm that the two residues have a great impact on the conversion rate. If these mutations only affect the stability of α₅-helix or the whole protein, we would expect to see a correlation between *T_m* and the conversion rate. On the contrary, the results show that the change in conversion rate caused by point mutations does not correlate with the change in protein thermal stability, indicating again that residues F294 and D295 should play important roles in mediating the conversion between monomer to DS-dimer.

Taken together, a disordered “α₅-helix” is required and functions to mediate the conversion between monomer to DS-dimer, while the ordered α₅-helix blocks the domain swapping of M^{pro}-C.

Self-association Mediated by Disordered “α₅-helix” of M^{pro}-C. We further characterized M^{pro}-C under the swappable conditions, in an effort to find out how M^{pro}-C behaves differently to enable the domain swapping after α₅-helix becomes disordered.

Interestingly, it was found that some NH peaks of M^{pro}-C at 37 °C display concentration dependent line broadening. Comparison of NH peak intensities of samples with concentrations of 0.1 and 2.0 mM shows that NH signals of residues Q189-D197, T199-A206, T226-N228, E240, T243, M276-G278, and I286, L287, D289-F292, F294-Q306 are getting much weaker at higher concentration (*I*_{0.1 mM}/*I*_{2.0 mM} > 1.5) (Fig. 4C). These residues are mainly located at the N- and C-terminal regions (including α₅-helix), with a few on α₂ and α₃-helices and loop connecting them (Fig. 4D, Fig. S5E). Similar phenomenon was also observed for M^{pro}-C at 25 °C/urea (Fig. 4B), while residues displaying concentration dependent NH signal intensity changes are almost identical to those at 37 °C (Fig. 4D). On the contrary, there is almost no NH signal intensity change for M^{pro}-C at 25 °C between 2 and 0.1 mM samples (Fig. 4A). We also compared 2D ¹H-¹⁵N HSQC spectra of M^{pro}-C²⁹² at different concentrations, and NH signals show little concentration dependent intensity change (Fig. S5C). As the removal of N-terminal ten residues

Table 1. Thermal stability and monomer to DS-dimer conversion ability for wild-type and mutants of M^{pro}-C

Mutants	Sequence	Domain swapping	<i>T_m</i> (°C)	<i>k_a</i> (mM ⁻¹ ·s ⁻¹) at 37 °C
M ^{pro} -C	187–306	yes	59.1 ± 1.1	11.1 ± 0.3
M ^{pro} -C ³⁰¹	187–301	yes	56.3 ± 0.2	26.2 ± 0.8
M ^{pro} -C ²⁹⁸	187–298	yes	53.8 ± 1.4	52.4 ± 0.1
M ^{pro} -C ²⁹⁶	187–296	yes	51.8 ± 0.4	151 ± 5
M ^{pro} -C ²⁹⁵	187–295	yes	51.6 ± 0.9	629 ± 7
M ^{pro} -C ²⁹⁴	187–294	yes	48.3 ± 0.8	164 ± 8
M ^{pro} -C ²⁹³	187–293	no	48.0 ± 0.5	/
M ^{pro} -C ²⁹²	187–292	no	47.8 ± 0.2	/
M ^{pro} -C ^{F294A}	F294A	yes	60.2 ± 0.9	1.7 ± 0.9
M ^{pro} -C ^{F294K}	F294K	yes	53.6 ± 0.7	37.7 ± 0.4
M ^{pro} -C ^{F294D}	F294D	yes	62.0 ± 0.2	1.3 ± 0.8
M ^{pro} -C ^{F294W}	F294W	yes	59.1 ± 0.2	7.1 ± 0.6
M ^{pro} -C ^{D295A}	D295A	yes	58.3 ± 0.7	7.3 ± 0.9
M ^{pro} -C ^{D295K}	D295K	yes	52.0 ± 0.6	226 ± 3
M ^{pro} -C ^{F294A/D295A}	F294A/D295A	yes	57.0 ± 0.2	10.5 ± 0.8
M ^{pro} -C ^{F294D/D295A}	F294D/D295A	yes	59.0 ± 0.3	5.3 ± 0.9

/ indicates that data is not applicable.

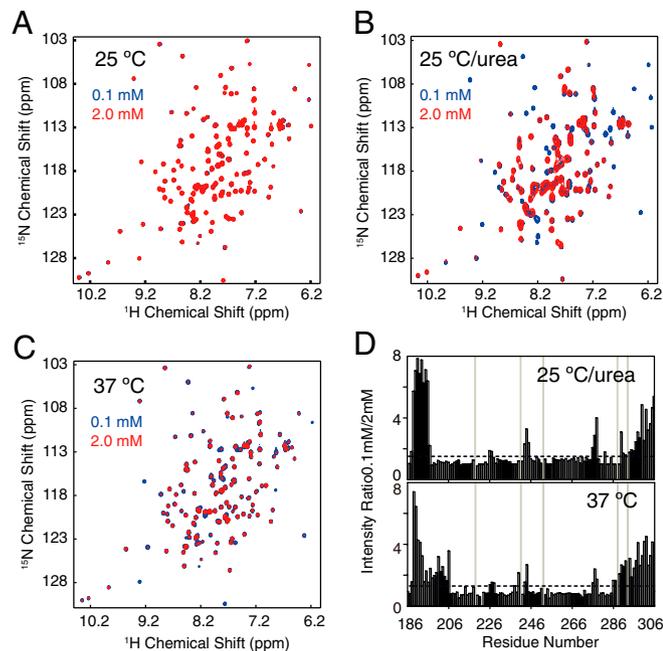


Fig. 4. Self-association of $M^{\text{Pro-C}}$ monomer under swappable conditions. (A–C) Overlay of 2D ^1H - ^{15}N HSQC spectra of $M^{\text{Pro-C}}$ with the concentration of 0.1 mM (blue) and 2.0 mM (red) at 25 °C (A), 25 °C/urea (B), and 37 °C (C). (D) Per residue NH signal intensity ratio of $M^{\text{Pro-C}}$ monomer between samples of 0.1 mM and 2.0 mM at 37 °C (Bottom) and 25 °C/urea (Top).

has almost no effect on the monomer/DS-dimer interconversion (Fig. S5B), this difference observed between the unswappable condition and swappable conditions should be mainly due to the disruption of α_5 -helix in $M^{\text{Pro-C}}_{\text{nds}}$.

Since the concentration dependent NH signal intensity change is an indication of protein self-association, these results suggest that $M^{\text{Pro-C}}_{\text{nds}}$ molecules can self-associate due to the disordered “ α_5 -helix” under swappable conditions. On the contrary, there is no interaction between $M^{\text{Pro-C}}$ monomers when α_5 -helix is intact under the unswappable condition. Thus, it is very likely that the self-association induced by the disordered “ α_5 -helix” is relevant to the domain-swapping process of $M^{\text{Pro-C}}$.

Discussion

3D domain swapping is one of the three proposed models for the amyloid fibril formation (4, 5), and it is important to understand the mechanism for 3D domain swapping, especially the mechanism for the conversion of monomer into 3D domain-swapped dimer/oligomer under physiological conditions.

In the current study, we have demonstrated that $M^{\text{Pro-C}}$ can reversibly interconvert between the monomeric and domain-swapped dimeric states at 37 °C, pH 7.0. As the swapped α_1 -helix of $M^{\text{Pro-C}}$ is completely buried in the protein hydrophobic core, it is inevitable that at least the protein hydrophobic core needs to be unpacked and repacked during the domain swapping process. However, it is thermodynamically impossible for the conversion to take place through the naturally occurred unfolded state according to the pre-existing equilibrium theory (25). Our NH exchange data indicate that the protein core is not exposed to the solvent during the 3D domain swapping process, suggesting that the unpacking and repacking of the hydrophobic core must be undertaken in a hydrophobic environment through a special intermediate state, which is beneficial for the domain swapping since the energy cost can be minimized.

We have showed that the C-terminal α_5 -helix of $M^{\text{Pro-C}}$ monomer under unswappable condition is disrupted and becomes disordered under the swappable conditions. Analysis of $M^{\text{Pro-C}}$ monomer and DS-dimer structures indicate that this order-

to-disorder conformational transition should favor the domain swapping process. An intact α_5 -helix locks the structure into an “inactive” state for domain swapping, while disruption of α_5 -helix under swappable conditions should serve as a key to unlock the protein into an “active” state and initiate the swapping of α_1 -helix. Furthermore, we found that residues within α_5 -helix should also play a role to mediate the domain swapping process, in addition to the role of “lock/unlock.” Studies of C-terminal truncation mutants reveal that fully removal of α_5 -helix results in loss of $M^{\text{Pro-C}}$ ’s ability to form domain-swapped dimers, while inclusion of residue 294 at the C-terminus is a “must” for $M^{\text{Pro-C}}$ truncation mutants to convert into domain-swapped dimers. Site-specific mutagenesis results also indicate that residues 294 and 295 have a great impact on the rate of the monomer to DS-dimer conversion. Meanwhile, we observed that $M^{\text{Pro-C}}$ monomers can self-associate upon disruption of α_5 -helix but not when α_5 -helix is intact or absent, which lead us to speculate that the disordered “ α_5 -helices” mediated protein self-association may enable $M^{\text{Pro-C}}$ to form a special dimeric intermediate that facilitates the swapping of α_1 -helices without exposing them to solvent.

Taken all together, we propose a mechanism for the interconversion between $M^{\text{Pro-C}}$ monomer and domain-swapped dimer (Fig. 5). The C-terminal α_5 -helix of $M^{\text{Pro-C}}$ monomer must first undergo an order-to-disorder transition, which converts $M^{\text{Pro-C}}$ into an active conformation for domain swapping. The disordered “ α_5 -helix” then mediates protein self-association and unpacking of the protein’s hydrophobic core to form a special intermediate dimer, which allows the exchange of α_1 -helices in a hydrophobic environment. Next, each swapped α_1 -helix is repacked with three other α -helices from the partner monomer to form hydrophobic cores of the domain-swapped dimer. Lastly, the disordered “ α_5 -helix” becomes ordered and locks the protein into the DS-dimer conformation. Although our efforts were mainly focused on the study of the formation of DS-dimer by $M^{\text{Pro-C}}$ monomer, we have also characterized the DS-dimer with NMR under swappable condition, while the domain swapped dimer can dissociate into the monomer. Comparison of 2D ^1H - ^{15}N HSQC spectra of $M^{\text{Pro-C}}$ monomer and DS-dimer at 37 °C reveals that NH signals of “ α_5 -helix” residues are exactly overlapped (Fig. S6D), indicating that “ α_5 -helices” of DS-dimer are also disordered under swappable condition. Although the two α_5 -helices of $M^{\text{Pro-C}}$ DS-dimer are positioned on far-most opposite sides in the crystal structure (Fig. S6C) (21), we found that the two α_5 -helices can be close to each other in solution using NMR spin label experiments. A paramagnetic spin label agent MTSL attached to C300 on α_5 -helix of one molecule of the DS-dimer can result in almost all NH signals of residues on α_5 -helix of the other molecule vanished (Fig. S6A–C). This suggests that the hinge loop of the DS-dimer is flexible and there is no defined relative orientation for the two monomer-like subunits in solution. Therefore, it is possible for the conversion of DS-dimer into monomer to take place through a reversed process of the dimerization (Fig. 5).

It was demonstrated that many proteins act as accretions of cooperative unfolding/refolding foldon units under native

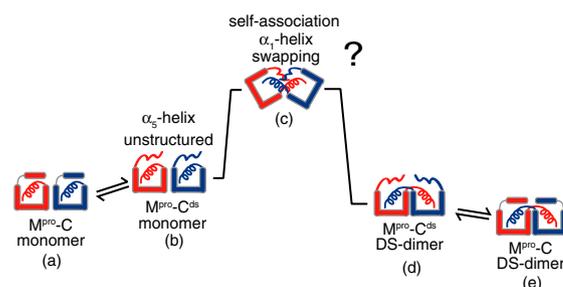


Fig. 5. Proposed mechanism for $M^{\text{Pro-C}}$ 3D domain swapping process.

conditions, while the less stable foldons are suggested to be related to protein functional behaviors (28). A foldon-dependent molecular switching mechanism has been previously proposed for the conversion of metamorphic protein Mad2 from the check-point inactive open form (O-Mad2) to an active close form (C-Mad2), which is catalyzed by Mad1-C-Mad2 complex. The catalytic C-Mad2 is suggested to selectively bind and stabilize a partially unfolded intermediate I-Mad2, which also lowers the energy level of the transition state in favor of the conversion from O-Mad2 to C-Mad2 (29). While in the case of M^{Pro}-C, since the folding/unfolding of α_5 -helix does not affect the conformation and packing of the other four α -helices, α_5 -helix should be the least stable foldon that unfolds first in M^{Pro}-C. As we have showed that the unfolding of α_5 -helix is necessary for M^{Pro}-C monomers to self-associate and form the special intermediate dimer facilitating the exchange of α_1 -helices, the conversion rate must be strongly affected by the α_5 -helix foldon stability. It is also possible that the self-association of M^{Pro}-C mediated by the disordered " α_5 -helices" can further stabilize the special dimeric intermediate state in favor of the formation of domain-swapped dimer, as in the case for Mad2 (29). The domain swapping process of M^{Pro}-C is therefore dependent on the dynamic foldon unfolding and refolding under native condition.

Meanwhile, our mutagenesis study results indicate that the conversion rate does not correlate directly with the protein thermal stability, and the unfolding of α_5 -helix serves to generate the disordered " α_5 -helix" necessary for mediating the domain swapping process. This provides a special "gain-of-function" example for the order-to-disorder transition of a protein region. A similar phenomenon has been reported for the redox-regulated chaperone Hsp33, of which the C-terminal redox-switch domain becomes unfolded to be active, while it is well folded in the inactive state (30). Nowadays, it's well accepted that the intrinsically dis-

ordered protein domain/region plays important biological roles in regulation of cellular processes which often associate with a disorder-to-order conformational transition, as demonstrated by many examples (31, 32). Our results suggest that an order-to-disorder conformational transition may also play important roles in regulating protein functions, as its opposite process does.

Traditionally, studies of the mechanism for 3D domain swapping mainly focused on the swapped elements and the hinge loop regions (2, 8, 33). It is distinctive that an unswapped element of a protein plays a key role in mediating the 3D domain swapping process under native condition. Our results on M^{Pro}-C shed new light on the mechanism of protein 3D domain swapping, and provide a sound mechanism on how two well-folded protein molecules can spontaneously cross a huge energy barrier to swap structure elements under physiological condition. It is still not clear how exactly the disordered " α_5 -helix" mediates the 3D domain swapping of M^{Pro}-C, and further studies are needed to reveal its detailed molecular mechanism.

Materials and Methods

Wide-type M^{Pro}-C and all truncation as well as point mutants of M^{Pro}-C were expressed and purified as described previously (21). All NMR experimental data were collected on Bruker Avance 500-MHz or 800-MHz spectrometers equipped with triple-resonance cryoprobes. Relevant data collection and refinement statistics for structure determination are provided in Table S3. A complete description of the materials and experimental procedures is included in *SI Text*.

ACKNOWLEDGMENTS. We thank Chao Liang, Dr. Si Wu, and Prof. Luhua Lai for their assistance on experiments. We also thank Prof. Xianming Pan, Prof. Zhirong Liu, and Prof. Lin Wang for their helpful advices and discussions. This work was supported by Grant No. 2012CB910700, Grant No. 2011IM030300, and Grant No. 2003CB514104 from the 973 Program of China, and Grant No. 31170682 from NSFC of China. All NMR experiments were carried out at Beijing NMR Center.

- Gronenborn AM (2009) Protein acrobatics in pairs—Dimerization via domain swapping. *Curr Opin Struct Biol* 19:39–49.
- Liu Y, Eisenberg D (2002) 3D domain swapping: As domains continue to swap. *Protein Sci* 11:1285–1299.
- Rousseau F, Schymkowitz JW, Itzhaki LS (2003) The unfolding story of three-dimensional domain swapping. *Structure* 11:243–251.
- Bennett MJ, Sawaya MR, Eisenberg D (2006) Deposition diseases and 3D domain swapping. *Structure* 14:811–824.
- van der Wel PC (2012) Domain swapping and amyloid fibril conformation. *Prión* 6:211–216.
- Jerala R, Zerovnik E (1999) Accessing the global minimum conformation of stefin A dimer by annealing under partially denaturing conditions. *J Mol Biol* 291:1079–1089.
- Miller KH, Marqusee S (2011) Propensity for C-terminal domain swapping correlates with increased regional flexibility in the C-terminus of RNase A. *Protein Sci* 20:1735–1744.
- Gronenborn AM, et al. (2002) The domain-swapped dimer of cyanovirin-N is in a metastable folded state: Reconciliation of X-ray and NMR structures. *Structure* 10:673–686.
- Ivanov D, et al. (2007) Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc Natl Acad Sci USA* 104:4353–4358.
- Penin F, et al. (2001) Evidence for a dimerisation state of the Bacillus subtilis catabolite repression HPr-like proteins, Crh. *J Mol Microbiol Biotechnol* 3:429–432.
- Byeon IJ, Louis JM, Gronenborn AM (2004) A captured folding intermediate involved in dimerization and domain-swapping of GB1. *J Mol Biol* 340:615–625.
- Rousseau F, Schymkowitz JW, Wilkinson HR, Itzhaki LS (2001) Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc Natl Acad Sci USA* 98:5596–5601.
- Rousseau F (2008) *Protein Dimerization and Oligomerization in Biology*, ed J Matthews (Landes Bioscience, Austin), pp 1–13.
- Hirota S, et al. (2010) Cytochrome c polymerization by successive domain swapping at the C-terminal helix. *Proc Natl Acad Sci USA* 107:12854–12859.
- Liu Y, Gotte G, Libonati M, Eisenberg D (2001) A domain-swapped RNase A dimer with implications for amyloid formation. *Nat Struct Biol* 8:211–214.
- Lopez-Alonso JP, et al. (2010) NMR spectroscopy reveals that RNase A is chiefly denatured in 40% acetic acid: Implications for oligomer formation by 3D domain swapping. *J Am Chem Soc* 132:1621–1630.
- Yang S, et al. (2004) Domain swapping is a consequence of minimal frustration. *Proc Natl Acad Sci USA* 101:13786–13791.
- Hayes MV, Sessions RB, Brady RL, Clarke AR (1999) Engineered assembly of intertwined oligomers of an immunoglobulin chain. *J Mol Biol* 285:1857–1867.
- Liu L, Byeon IJ, Bahar I, Gronenborn AM (2012) Domain swapping proceeds via complete unfolding: A 19F- and 1H-NMR study of the Cyanovirin-N protein. *J Am Chem Soc* 134:4229–4235.
- Malevanets A, Sirota FL, Wodak SJ (2008) Mechanism and energy landscape of domain swapping in the B1 domain of protein G. *J Mol Biol* 382:223–235.
- Zhong N, et al. (2009) C-terminal domain of SARS-CoV main protease can form a 3D domain-swapped dimer. *Protein Sci* 18:839–844.
- Zhang S, et al. (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, et al. (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.
- Hartl FU, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature* 475:324–332.
- Tsai CJ, Kumar S, Ma B, Nussinov R (1999) Folding funnels, binding funnels, and protein function. *Protein Sci* 8:1181–1190.
- von Smoluchowski M (1916) Three presentations on diffusion, molecular movement according to Brown and coagulation of colloid particles. *Phys Z* 17:557–571.
- Northrup SH, Erickson HP (1992) Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc Natl Acad Sci USA* 89:3338–3342.
- Englander SW, Mayne L, Krishna MM (2007) Protein folding and misfolding: Mechanism and principles. *Q Rev Biophys* 40:287–326.
- Skinner JJ, Wood S, Shorter J, Englander SW, Black BE (2008) The Mad2 partial unfolding model: Regulating mitosis through Mad2 conformational switching. *J Cell Biol* 183:761–768.
- Ilbert M, et al. (2007) The redox-switch domain of Hsp33 functions as dual stress sensor. *Nat Struct Mol Biol* 14:556–563.
- Wright PE, Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* 19:31–38.
- Uversky VN (2011) Intrinsically disordered proteins from A to Z. *Int J Biochem Cell Biol* 43:1090–1103.
- Huang Y, Cao H, Liu Z (2012) Three-dimensional domain swapping in the protein structure space. *Proteins* 80:1610–1619.