THE PROTEIN SOCIETY

Evidences for the unfolding mechanism of three-dimensional domain swapping

Zhirong Liu,[1,2,3]* and Yongqi Huang[1,2,3]

[1]College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China
[2]Beijing National Laboratory for Molecular Sciences (BNLMS), Peking University, Beijing 100871, China
[3]Center for Quantitative Biology, Peking University, Beijing 100871, China

Received 2 September 2012; Revised 2 December 2012; Accepted 4 December 2012
DOI: 10.1002/pro.2209
Published online 13 December 2012 proteinscience.org

Abstract: The full or partial unfolding of proteins is widely believed to play an essential role in three-dimensional domain swapping. However, there is little research that has rigorously evaluated the association between domain swapping and protein folding/unfolding. Here, we examined a kinetic model in which domain swapping occurred via the denatured state produced by the complete unfolding of proteins. The relationships between swapping kinetics and folding/unfolding thermodynamics were established, which were further adopted as criteria to show that the proposed mechanism dominates in three representative proteins: Cyanovirin-N (CV-N), the C-terminal domain of SARS-CoV main protease (M^{pro}-C), and a single mutant of oxidized thioredoxin (Trx_W28A^{ox}).

Keywords: domain swapping; protein folding; protein unfolding; kinetics; thermodynamics; Cyanovirin-N; SARS-CoV main protease; thioredoxin

Introduction

Three-dimensional domain swapping is a special form of protein oligomerization, where monomers exchange one or more identical structural elements (ranging from secondary structure elements to whole structural domains) to form complexes.[1] Currently, more than 500 domain-swapped structures have been solved.[2] An analysis in the protein structural space suggested that domain swapping is a general property of proteins.[2] Domain swapping possesses many potential biological implications.[1,3] It acts as a mechanism for regulating protein function, and as an evolutionary strategy to create protein complexes. This process is also involved in protein misfolding and aggregation.

In comparison with conventional protein–protein interactions,[4] domain swapping has two distinct kinetic features. First, the interconversion between the monomer and the domain-swapped dimer is generally very slow.[5–7] The equilibrium process may take days or even months. Second, the interconversion rate is very sensitive to temperature.[5,6,8,9] An increase in temperature by 3–4°C is sufficient to produce an overall increase of the conversion rate by 10-fold.[8,9] The extracted enthalpy component of the activation barrier is larger than 100 kcal/mol. These behaviors are closely related to the folding/unfolding process of proteins. In an extreme model (the unfolding mechanism for domain swapping),[6,8,10] it was proposed that domain swapping proceeds via complete unfolding, so the swapping kinetics can be explained in terms of the equilibrium folding/unfolding properties of proteins. Very recently, a study on Cyanovirin-N indeed verified that the energy barrier of domain swapping is very close to the equilibrium unfolding enthalpy of the protein.[8] However, there were also some doubts on the feasibility of the unfolding mechanism of domain swapping.[9,11] The main concern is that the population of the fully unfolded state is too low to account for the observed

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Ministry of Science and Technology of China; Grant number: 2009CB918500; Grant sponsor: National Natural Science Foundation of China; Grant numbers: 20973016 and 11021463.

*Correspondence to: Zhirong Liu, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China. E-mail: liuzhirong@pku.edu.cn

Published by Wiley-Blackwell. © 2012 The Protein Society

swapping rate.[9,11] In addition, careful examination with rigorous formalism is necessary to distinguish domain-swapping via fully unfolded states from domain-swapping via partially unfolded states. For example, the rigorous derivation of the unfolding mechanism for domain swapping predicts that the enthalpy barrier of domain swapping is two times that of the equilibrium unfolding enthalpy of the protein monomer (will be given below), and not one-fold as that defined previously.[8]

Here, we examined the kinetic properties of domain swapping under the unfolding mechanism and analyzed the experimental swapping data in combination with the folding/unfolding data of three proteins: Cyanovirin-N (CV-N), the C-terminal domain of SARS-CoV main protease (M^{pro}-C), and a single mutant of oxidized thioredoxin (Trx_W28A^{ox}). The results showed that domain swapping in all three systems are well described by the unfolding mechanism when the heat capacity difference between the native and denatured states in protein folding/unfolding is appropriately addressed.

Results

Formalism of the unfolding mechanism for domain swapping

In general, domain swapping is described by a dimerization reaction:

$$2\mathrm{M} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \mathrm{D} \quad (1)$$

with the equilibrium dissociation constant $K_d = k_{off}/k_{on}$. The time evolution of monomer and dimer concentrations is given as (see Supporting Information):

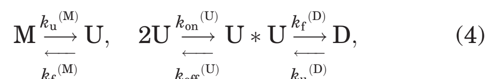
$$\begin{cases} [\mathrm{M}] = [\mathrm{M}]_{\mathrm{eq}} + \left([\mathrm{M}]_{t=0} - [\mathrm{M}]_{\mathrm{eq}}\right) \dfrac{\dfrac{\sqrt{K_d^2 + 8K_d[\mathrm{M}]_{\mathrm{total}}}}{2} \exp\left(-k_{\mathrm{on}} t \sqrt{K_d^2 + 8K_d[\mathrm{M}]_{\mathrm{total}}}\right)}{\dfrac{\sqrt{K_d^2 + 8K_d[\mathrm{M}]_{\mathrm{total}}}}{2} + \left([\mathrm{M}]_{t=0} - [\mathrm{M}]_{\mathrm{eq}}\right)\left[1 - \exp\left(-k_{\mathrm{on}} t \sqrt{K_d^2 + 8K_d[\mathrm{M}]_{\mathrm{total}}}\right)\right]} \\ [\mathrm{D}] = [\mathrm{D}]_{\mathrm{eq}} + \left([\mathrm{D}]_{t=0} - [\mathrm{D}]_{\mathrm{eq}}\right) \dfrac{\dfrac{\sqrt{K_d^2 + 16K_d[\mathrm{D}]_{\mathrm{total}}}}{4} \exp\left(-k_{\mathrm{on}} t \sqrt{K_d^2 + 16K_d[\mathrm{D}]_{\mathrm{total}}}\right)}{\dfrac{\sqrt{K_d^2 + 16K_d[\mathrm{D}]_{\mathrm{total}}}}{4} - \left([\mathrm{D}]_{t=0} - [\mathrm{D}]_{\mathrm{eq}}\right)\left[1 - \exp\left(-k_{\mathrm{on}} t \sqrt{K_d^2 + 16K_d[\mathrm{D}]_{\mathrm{total}}}\right)\right]} \end{cases}, \quad (2)$$

where $[M]_{total}$ and $[D]_{total}$ are the total molar concentrations (irrespective of whether they are present as monomer or dimer) with $[M]_{total} \equiv 2[D]_{total}$, and $[M]_{eq}$ and $[D]_{eq}$ are the corresponding equilibrium concentrations:

$$\begin{cases} [\mathrm{M}]_{\mathrm{eq}} = \dfrac{-K_d + \sqrt{K_d^2 + 8K_d[\mathrm{M}]_{\mathrm{total}}}}{4} \\ [\mathrm{D}]_{\mathrm{eq}} = \dfrac{8[\mathrm{D}]_{\mathrm{total}} + K_d - \sqrt{K_d^2 + 16K_d[\mathrm{D}]_{\mathrm{total}}}}{8} \end{cases}. \quad (3)$$

Equation (2) is generally applicable to various domain swapping processes no matter whether they proceed via complete or partial unfolding. It can be used to fit the experimental kinetic data of domain swapping to extract the constants k_{on}, k_{off}, and K_d. It is noted that when $\left|[\mathrm{M}]_{t=0} - [\mathrm{M}]_{\mathrm{eq}}\right| << \frac{\sqrt{K_d^2 + 8K_d[\mathrm{M}]_{\mathrm{total}}}}{2}$, the denominator on the right side of Eq. (2) is approximately a constant independent on time t, and thus, the kinetics can be described by a single exponential as observed previously.[8,9]

When domain swapping occurs by the unfolding mechanism, i.e., swapping proceeds via complete unfolding, the process is subdivided as:



$$\mathrm{M} \underset{k_f^{(\mathrm{M})}}{\overset{k_u^{(\mathrm{M})}}{\rightleftharpoons}} \mathrm{U}, \quad 2\mathrm{U} \underset{k_{\mathrm{off}}^{(\mathrm{U})}}{\overset{k_{\mathrm{on}}^{(\mathrm{U})}}{\rightleftharpoons}} \mathrm{U}*\mathrm{U} \underset{k_u^{(\mathrm{D})}}{\overset{k_f^{(\mathrm{D})}}{\rightleftharpoons}} \mathrm{D}, \quad (4)$$

where k_u and k_f are the unfolding and folding rate constants, respectively, and the superscripts "(M)" and "(D)" denote the monomer and dimer proteins. U is the denatured monomer and U*U is an encounter complex where two monomers come close to each other but do not essentially interact. The equilibrium dissociation constant of the global domain swapping process is thus:

$$K_d = \frac{K_u^{(\mathrm{D})}}{\left[K_u^{(\mathrm{M})}\right]^2} K_d^{(\mathrm{U})}, \quad (5)$$

where $K_u^{(M)} \equiv k_u^{(M)}/k_f^{(M)}$ and $K_u^{(D)} \equiv k_u^{(D)}/k_f^{(D)}$ are the unfolding equilibrium constants for monomer and dimer, respectively, while $K_d^{(U)}$ is the dissociation constant for the step $2U \leftrightarrow U*U$. Under the steady-state assumption and the condition that folding rates are fast (which operates when the denatured states are highly unstable), the time evolution of process (4) can be reduced into process (1) with the

corresponding parameters (see Supporting Information):

$$\begin{cases} k_{\text{on}} = \left[K_{\text{u}}^{(\text{M})}\right]^2 k_{\text{on}}^{(\text{U})} \\ k_{\text{off}} = K_{\text{u}}^{(\text{D})} k_{\text{off}}^{(\text{U})} \end{cases}. \tag{6}$$

This is the main results for the unfolding mechanism of domain swapping. It relates the kinetics of domain swapping to the equilibrium of protein folding/unfolding. By introducing variations (e.g., temperature, denaturant, and mutation) that change the protein stability, Eq. (6) can be critically assessed by examining the quantitative relation between the swapping kinetics and the unfolding thermodynamics. For example, when the temperature is raised to change the domain-swapping rate, Eq. (6) gives that:

$$\begin{cases} \Delta H_{\text{on}}^{\ddagger} = 2\Delta H_{\text{u}}^{(\text{M})} + \Delta H_{\text{on}}^{(\text{U})^{\ddagger}} \\ \Delta H_{\text{off}}^{\ddagger} = \Delta H_{\text{u}}^{(\text{D})} + \Delta H_{\text{off}}^{(\text{U})^{\ddagger}} \end{cases}, \tag{7}$$

where $\Delta H_{\text{on}}^{\ddagger}$ is the enthalpy barrier of k_{on} and $\Delta H_{\text{off}}^{\ddagger}$ is the enthalpy barrier of k_{off}. $\Delta H_{\text{u}}^{(\text{M})}$ and $\Delta H_{\text{u}}^{(\text{D})}$ are the equilibrium unfolding enthalpy of the monomer and the dimer, respectively. $\Delta H_{\text{on}}^{(\text{U})\ddagger}$ and $\Delta H_{\text{off}}^{(\text{U})\ddagger}$ are the enthalpy barrier of the transition between 2U and U*U. Considering that U*U is an encounter complex with no essential interaction between two completely unfolded monomers, $\Delta H_{\text{on}}^{(\text{U})\ddagger}$ and $\Delta H_{\text{off}}^{(\text{U})\ddagger}$ can be ignored. So we have

$$\begin{cases} \Delta H_{\text{on}}^{\ddagger} \approx 2\Delta H_{\text{u}}^{(\text{M})} \\ \Delta H_{\text{off}}^{\ddagger} \approx \Delta H_{\text{u}}^{(\text{D})} \end{cases}. \tag{8}$$

On the other hand, the temperature dependence of K_{d} is usually much smaller than that of k_{on} and k_{off}, so it is derived from $K_{\text{d}} = k_{\text{off}}/k_{\text{on}}$ that:

$$\Delta H_{\text{off}}^{\ddagger} \approx \Delta H_{\text{on}}^{\ddagger}. \tag{9}$$

The relationships can be summarized as:

$$\Delta H_{\text{on}}^{\ddagger} \approx \Delta H_{\text{off}}^{\ddagger} \approx 2\Delta H_{\text{u}}^{(\text{M})} \approx \Delta H_{\text{u}}^{(\text{D})}. \tag{10}$$

Therefore, the kinetic barrier of domain swapping is two times the equilibrium unfolding enthalpy of the monomer and one time the equilibrium unfolding enthalpy change of the dimer. This provides a criterion for the unfolding mechanism of domain swapping. The $k_{\text{on}}^{(\text{U})}$ and $k_{\text{off}}^{(\text{U})}$ can also be extracted from Eq. (6) to see whether they lie in a reasonable range.

Equations (6) and (10) apply for domain swapping which proceeds via complete unfolding. When swapping proceeds via partial unfolding, there would be no correlation between the swapping kinetics and the global protein stability, and thus Eqs. (6) and (10) are not applicable. Swapping via partial unfolding is still described by Eqs. (2,3), but the thermodynamics of partial unfolding should be instead used to relate to the swapping kinetics of domain swapping. Because the free energy for partial unfolding is usually not as high as that for complete unfolding under native conditions, the interconversion rate for swapping via partial unfolding should be faster and less sensitive to the temperature.

In the following sections, we examined three domain-swapped proteins whose swapping kinetics and folding/unfolding thermodynamics are available in the literature, and showed an agreement between the predictions and the experimental data.

Case study 1: Cyanovirin-N

The protein CV-N is a potent inhibitor of the human immunodeficiency virus and many other viruses.[12] CV-N is composed of 101 amino acids and exists in both a monomer and a domain-swapped dimer.[13] Recently, Liu *et al.*[8] found that the swapping enthalpy barrier is very large and of similar magnitude to the equilibrium unfolding enthalpy of the monomer and dimer, and concluded that domain swapping proceeds via the unfolding mechanism. However, although $\Delta H^{\ddagger} \approx \Delta H_{\text{u}}^{(\text{D})}$ was verified, a relation of $\Delta H^{\ddagger} \approx \Delta H_{\text{u}}^{(\text{M})}$ was observed in their work rather than the predicted $\Delta H^{\ddagger} \approx 2\Delta H_{\text{u}}^{(\text{M})}$ presented in Eq. (10). In addition, they observed that the domain swapping reaction exhibited a single exponential time dependence, which was then used to support the suggestion that the rate-limiting step is M→U. Consequently, we decided to re-examine their data.

In fact, the kinetic data of Liu *et al.*[8] can be well described by Eq. (2). In Figure 1, we refit the conversion data of Liu *et al.*[8] from the wild-type (wt) CV-N domain-swapped dimer to the monomer using Eq. (2) and the enthalpy barrier relation:

$$k_{\text{on}}(T) = k_{\text{on}}(T_0) \exp\left[-\Delta H_{\text{on}}^{\ddagger}\left(\frac{1}{RT} - \frac{1}{RT_0}\right)\right], \tag{11}$$

where T_0 is a reference temperature. It can be seen from Figure 1 that the agreement between the experimental data and Eq. (2) is excellent even if only four parameters were used to globally fit six curves. The discrepancy of Eq. (2) from single exponential behaviors is reflected in the factor

$$\theta = \frac{\dfrac{\sqrt{K_{\text{d}}^2 + 16K_d[\text{D}]_{\text{total}}}}{4}}{\dfrac{\sqrt{K_{\text{d}}^2 + 16K_{\text{d}}[\text{D}]_{\text{total}}}}{4} - \left([\text{D}]_{t=0} - [\text{D}]_{\text{eq}}\right)\left[1 - \exp\left(-k_{\text{on}}t\sqrt{K_{\text{d}}^2 + 16K_{\text{d}}[\text{D}]_{\text{total}}}\right)\right]}. \tag{12}$$

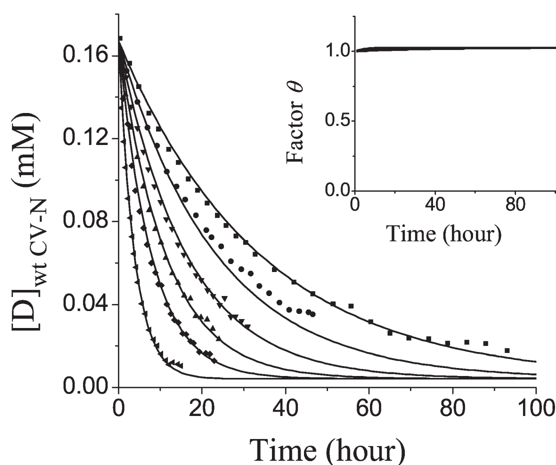


Figure 1. The conversion from the wt CV-N domain-swapped dimer to the monomer at various temperatures: (from top to bottom) 322.5, 323, 323.5, 324, 324.5, and 325.5 K. Scatters are experimental data extracted from Fig. 4(A) of Liu *et al.*[8] Solid lines are the theoretical fits using Eqs. (2,11) with the resulting parameters: $[D]_{total} = 0.167$ m*M*, $K_d = 25.7$ m*M*, k_{on} (322.5 K) $= 6.0 \times 10^{-4}$ M^{-1} s^{-1}, $\Delta H^{\ddagger}_{on} = 151$ kcal/mol. The inset shows the factor θ as given in Eq. (12) which reflects the discrepancy from the single exponential behaviors.

The ratio of $\left([D]_{t=0} - [D]_{eq}\right) \Big/ \frac{\sqrt{K_d^2 + 16K_d[D]_{total}}}{4}$ is calculated to be only 0.024, so it is not surprising that the factor θ is very close to 1, as shown in the inset of Figure 1. This explains why the experimental data can also be fitted by a single exponential in Liu *et al.*[8] Such a property is expected to apply also to other proteins when experiment conditions strongly favor the formation of the monomer ($[M]_{total} = 2[D]_{total} \ll K_d$). As a result, the enthalpy barrier $\Delta H^{\ddagger}$ extracted from our fitting (151 kcal/mol) is very close to that derived from the single exponential fitting (145/153 kcal/mol).[8]

On the problem why $\Delta H^{\ddagger} \approx \Delta H_u^{(M)}$ instead of $\Delta H^{\ddagger} \approx 2\Delta H_u^{(M)}$ was observed by Liu *et al.*,[8] we found that the answer lies in the large heat capacity difference (ΔC_p) between the denatured and native proteins.[14] The equilibrium unfolding enthalpy ($\Delta H_u^{(M)}$) was usually measured from thermal melting curves so that its value is applicable near the melting temperature (T_m), while the domain swapping kinetics was measured at lower temperatures, where, according to the equation of:

$$\Delta H_u(T) = \Delta H_u(T_m) + \Delta C_p(T - T_m), \qquad (13)$$

the corresponding $\Delta H_u^{(M)}$ should be lower than that at T_m. By incorporating the effect of ΔC_p, we recovered the relation of $\Delta H^{\ddagger} \approx 2\Delta H_u^{(M)}$ for CV-N as follows:

1 For the CV-N^{P51G} monomer, Table II of Liu *et al.*[8] gave $\Delta H_u(T_m) = 130$ kcal/mol, and Ref. 15 gave $T_m = 71.2$°C. We did not find the corresponding ΔC_p in the literature, so we made use of the experimental fact[16] of ΔG_u (20°C) = 9.8 kcal/mol and the equation:

$$\Delta G_u(T) = \Delta H_u(T_m) + \Delta C_p(T - T_m) - T\left(\frac{\Delta H_u(T_m)}{T_m} + \Delta C_p \ln\frac{T}{T_m}\right) \qquad (14)$$

to estimate ΔC_p to be 2.38 kcal/(mol·K). Therefore, $\Delta H_u^{(M)}$ at the temperature of the swapping measurement (~329 K) is determined to be 93.5 kcal/mol, and thus $2\Delta H_u^{(M)} = 187$ kcal/mol, which is close to the observed $\Delta H^{\ddagger}$ (162 kcal/mol) for swapping kinetics.[8]

2 For the wt CV-N monomer, experiments gave $T_m = 61.3$°C and ΔG_u (20°C) = 4.1 kcal/mol.[15,16] We did not find the experimental results on ΔC_p and ΔH_u (T_m) in the literature. If we assume the wt CV-N has the same ΔC_p value as CV-N^{P51G} above, we can estimate ΔH_u (T_m) = 85 kcal/mol based on Eq. (13). Then, $\Delta H_u^{(M)}$ at the swapping temperature (~325 K) is calculated to be 62.9 kcal/mol, and we have $2\Delta H_u^{(M)} = 126$ kcal/mol. This value is comparable to the observed $\Delta H^{\ddagger}$ value (145/153 kcal/mol).[8]

3 For the dimer, ΔH_u (T_m) of CV-N^{P51G} and CV-N$^{\Delta Q50}$ were reported by Liu *et al.*[8] However, for dimers which are less stable than monomers, their unfolding is usually coupled with a dimer-monomer transition and the extracted unfolding thermodynamics may be problematic. We have ignored the less stable CV-N^{P51G} dimer, and only discussed CV-N$^{\Delta Q50}$ that exists solely as a domain-swapped dimer. The T_m of CV-N$^{\Delta Q50}$ dimer is 50.2°C, which is close to the swapping temperature (~325 K) of wt CV-N. Thus, we directly estimated the $\Delta H_u^{(D)}$ of the wt CV-N dimer at the swapping temperature as the ΔH_u (T_m) value (142 kcal/mol) of CV-N$^{\Delta Q50}$, which is almost identical to the experimental $\Delta H^{\ddagger}$ value (145/153 kcal/mol)[8] of wt CV-N.

Case study 2: M^{pro}-C

M^{pro}, the main protease of the SARS coronavirus (SARS-CoV), is a key target for structure-based drug design against SARS.[17] The C-terminal domain of M^{pro} (M^{pro}-C) was found to exist in both monomeric and domain-swapped dimeric forms.[18] Unlike many other domain-swapped proteins, the swapped element of M^{pro}-C is fully buried inside the hydrophobic core rather than at the protein surface, which makes the unfolding mechanism of domain swapping more probable to occur in this system.[19] On the other hand, Kang *et al.*[9] recently measured the swapping kinetics and the folding/unfolding thermodynamics, and concluded that it is thermodynamically impossible for M^{pro}-C to swap through fully unfolded states. A main reason for this suggestion is that $k_{on}^{(U)}$

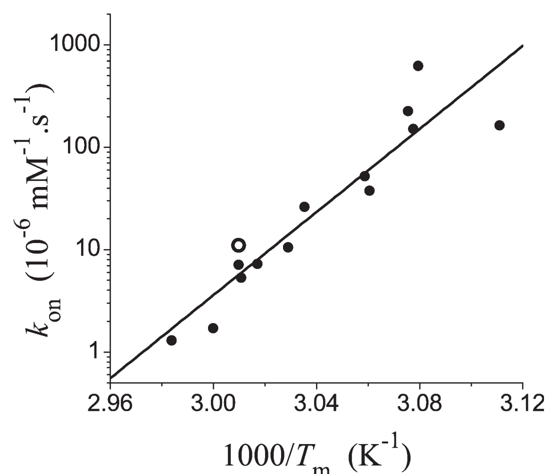


Figure 2. Correlation between the domain-swapping association rate (k_{on}) at 37°C and the melting temperature (T_m) for wt and mutants of M^{pro}-C. The experimental data were taken from Table I of Kang *et al.*[9] The wild-type is high-lighted by using an open circle. The solid line is a linear fit to the data.

calculated from Eq. (6) greatly exceeds the typical protein association rate constants.[9] However, after considering the influence of ΔC_p, we found that the conclusion may change. From Kang *et al.*[9] we have $T_m = 59.1°C$ and ΔH_u (T_m) = 430 kJ/mol for the M^{pro}-C monomer. ($\Delta H_u(T_m)$ was estimated from the data of ΔG_u (25°C) = 44.1 kJ/mol from Kang *et al.*[9] by assuming they did not considered the influence of ΔC_p. A direct refitting to Figure 2(F) of Kang *et al.*[9] results in a similar value of $\Delta H_u(T_m)$.) We are not aware of any experimental results describing the ΔC_p of M^{pro}-C, so we estimated this value from the general scaling law of ΔC_p with respect to the chain length,[14] which predicts that $\Delta C_p = -251 + 0.19$ (ΔASA) $= -251 + 0.19 \times (-907 + 93N)$ (in a unit of cal/(mol·K)) where ΔASA is the change in solvent-accessible surface area upon unfolding and N is the length of the protein. For M^{pro}-C (N = 120), it is estimated that ΔC_p = 7.1 kJ/(mol·K). With these parameters, we re-evaluated the folding/unfolding thermodynamics at the swapping temperature (~37°C) to be: $\Delta H_u^{(M)} = 270$ kJ/mol and $\Delta G_u^{(M)} = 23.3$ kJ/mol. Thus, $2\Delta H_u^{(M)} = 540$ kJ/mol and this value is comparable to the swapping $\Delta H^{\ddagger}$ (374/436 kJ/mol).[9] Based on $K_u^{(M)} = \exp\left[-\frac{\Delta G_u^{(M)}}{RT}\right] = 1.2 \times 10^{-4}$ and the k_{on} value (11.1×10^{-3} M^{-1} s^{-1}) at 37°C from Kang *et al.*,[9] $k_{on}^{(U)}$ is determined from Eq. (6) to be 7.7×10^5 M^{-1} s^{-1}, which is a typical protein association rate limited by diffusion.[4] Consequently, the doubt on the feasibility of the unfolding mechanism may be dismissed.

Kang *et al.* have also constructed various mutants of M^{pro}-C and measured their thermal stability and domain swapping kinetics.[9] By redrawing their data in Figure 2, it is clearly demonstrated that, despite some fluctuations, there is a tight correlation between k_{on} and T_m. This result strongly suggests that the change in the swapping kinetics of the mutants is simply because of a change in the protein thermal stability. A linear fitting to $\ln k_{on} \sim 1/T_m$ gives an effective enthalpy difference of 388 kJ/mol, which is very similar to the observed swapping $\Delta H^{\ddagger}$ value (373/433 kJ/mol).[9] The calculated $k_{on}^{(U)}$ for the mutants falls in the range of 1.0–16.9 $\times$ 10^5 M^{-1} s^{-1} (Supporting Information Table S1), the majority of which are close to that for the wide-type. The consistence among $k_{on}^{(U)}$ for the wide-type and various mutants of M^{pro}-C lends support to the unfolding mechanism, i.e., swapping proceeds via complete unfolding.

It is noted that the state U*U in Eq. (4) is not well established, so the physicochemical meaning of the parameters such as $k_{on}^{(U)}$ may depend on the system. In the unfolding mechanism we discussed, U*U is assumed to be made of two completely unfolded monomers. If swapping proceeds via partial unfolding, then U*U should be composed of partially unfolded monomers and the derived swapping kinetics would relate to the thermodynamics of partial unfolding. Thus, there does not exist a tight correlation between the swapping kinetics and the global protein stability for swapping proceeds via partial unfolding.

Case study 3: Trx_W28A^{ox}

Thioredoxin (Trx) plays an essential role in many biological processes, including cellular redox balance, promotion of cell growth, and inhibition of apoptosis.[20] Garcia-Pino *et al.* showed that a single active-site mutation on the oxidized form (Trx_W28A^{ox}) converts the protein into a biologically inactive domain-swapped dimer.[21] The swapped dimer of Trx_W28A^{ox} is a kinetically trapped species. Its unfolding is not reversible, i.e., it spontaneously refolds to the monomer after thermal unfolding. (In Garcia-Pino *et al.*,[21] the transition was written as $S_2 \rightarrow 2I$, which is synonymous to our notation of D→2M here. For the unfolding of monomer, the van't Hoff enthalpy instead of the calorimetric enthalpy was assigned to $\Delta H_u^{(M)}$ since van't Hoff enthalpy better reflects the properties of denatured population.) In other words, the K_d is very large in this system and the measured swapping kinetics is mainly determined by k_{off}. Based on Table III of Garcia-Pino *et al.*,[21] we have calculated that $\Delta H_u^{(D)} = \Delta H_{D \rightarrow 2M} + 2\Delta H_u^{(M)} = 14.4 + 2 \times 47.1 = 109$ kcal/mol. (The swapping kinetics was extracted from the temperature dependence of the apparent excess heat capacity in Garcia-Pino *et al.*,[21] so the swapping temperature was close to the melting temperature and we need not consider ΔC_p in the calculation). This $\Delta H_u^{(D)}$ value is very similar to the measured swapping kinetic barrier $\Delta H^{\ddagger}$ (120 kcal/mol),[21] suggesting the unfolding mechanism to be responsible for the domain swapping of Trx_W28A^{ox}.

Discussion

Although there is currently no unifying molecular mechanism describing domain swapping, it is generally believed that the monomer should be fully or partially unfolded in swapping. However, criteria should be developed to rigorously test any proposed mechanism. In this article, we established the formalism of the (fully) unfolding mechanism for domain swapping and used the obtained criteria to analyze the properties of three representative proteins. We are not suggesting that the unfolding mechanism is universal for all swapped proteins since each protein may behave in a distinct manner; however, the criteria presented were met in the examined proteins. Therefore, the unfolding mechanism probably dominates in these example systems.

The developed formalism can be extended to describe other experiments of domain swapping. For example, when denaturants are used to increase the rate of the swapping kinetics, it is predicted from the unfolding mechanism that the slope of the logarithmic swapping rate as a function of the denaturant concentration is two times as that of the unfolding equilibrium constants for the monomer. This remains to be verified in future work.

It should also be interesting to explore the mechanism where proteins are partially unfolded in swapping. By measuring the equilibrium of the partially unfolded forms by techniques such as native-state hydrogen exchange,[22] the connection between swapping kinetics and the partial unfolding thermodynamics may be established and be tested similarly.

Materials and Methods

Details on the formalism of the unfolding mechanism for domain swapping are described in the Supporting Information.

Conclusions

In conclusion, we have established formulism of domain swapping under the unfolding mechanism, and used the obtained criteria to test a number of protein systems by combining their swapping kinetics with available folding/unfolding equilibrium data. The results suggest that the domain swapping of CV-N, M^{pro}-C, and Trx_W28A^{ox} is dominated by the unfolding mechanism.

Acknowledgment

The authors gratefully acknowledge Prof. Bing Xia and Dr. Xue Kang for their insightful discussions which stimulated this work.

References

1. Bennett MJ, Schlunegger MP, Eisenberg D (1995) 3D domain swapping—a mechanism for oligomer assembly. Protein Sci 4:2455–2468.
2. Huang YQ, Cao HQ, Liu ZR (2012) Three-dimensional domain swapping in the protein structure space. Proteins 80:1610–1619.
3. Gronenborn AM (2009) Protein acrobatics in pairs—dimerization via domain swapping. Curr Opin Struct Biol 19:39–49.
4. Schreiber G, Haran G, Zhou HX (2009) Fundamental aspects of protein-protein association kinetics. Chem Rev 109:839–860.
5. 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.
6. Rousseau F, Schymkowitz JWH, 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.
7. 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.
8. Liu L, Byeon IJL, 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.
9. Kang X, Zhong N, Zou P, Zhang SN, Jin CW, Xia B (2012) Foldon unfolding mediates the interconversion between M^{pro}-C monomer and 3D domain-swapped dimer. Proc Natl Acad Sci USA 109:14900–14905.
10. Rousseau F, Schymkowitz JWH, Itzhaki LS (2003) The unfolding story of three-dimensional domain swapping. Structure 11:243–251.
11. 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.
12. Boyd MR, Gustafson KR, McMahon JB, Shoemaker RH, Okeefe BR, Mori T, Gulakowski RJ, Wu L, Rivera MI, Laurencot CM, Currens MJ, Cardellina JH, Buckheit RW, Nara PL, Pannell LK, Sowder RC, Henderson LE (1997) Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. Antimicrob Agents Chemother 41:1521–1530.
13. Yang F, Bewley CA, Louis JM, Gustafson KR, Boyd MR, Gronenborn AM, Clore GM, Wlodawer A (1999) Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. J Mol Biol 288:403–412.
14. Myers JK, Pace CN, Scholtz JM (1995) Denaturant m-values and heat-capacity changes—relation to changes in accessible surface-areas of protein unfolding. Protein Sci 4:2138–2148.
15. Barrientos LG, Lasala F, Delgado R, Sanchez A, Gronenborn AM (2004) Flipping the switch from monomeric to dimeric CV-N has little effect on antiviral activity. Structure 12:1799–1807.
16. Barrientos LG, Louis JM, Botos I, Mori T, Han ZZ, O'Keefe BR, Boyd MR, Wlodawer A, Gronenborn AM (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.
17. Fan KQ, Wei P, Feng Q, Chen SD, Huang CK, Ma L, Lai B, Pei JF, Liu Y, Chen JG, Lai LH (2004) Biosynthesis, purification, and substrate specificity of severe

acute respiratory syndrome coronavirus 3C-like proteinase. J Biol Chem 279:1637–1642.
18. Zhong N, Zhang SN, Xue F, Kang X, Zou P, Chen JX, Liang C, Rao ZH, Jin CW, Lou ZY, Xia B (2009) C-terminal domain of SARS-CoV main protease can form a 3D domain-swapped dimer. Protein Sci 18:839–844.
19. Huang YQ, Kang X, Xia B, Liu ZR (2012) How a core helix is swapped in the C-terminal domain of SARS coronavirus main protease: clues from simulations. Acta Phys-Chim Sinica 28:2411–2417.
20. Lillig CH, Holmgren A (2007) Thioredoxin and related molecules—from biology to health and disease. Antioxid Redox Signal 9:25–47.
21. Garcia-Pino A, Martinez-Rodriguez S, Wahni K, Wyns L, Loris R, Messens J (2009) Coupling of domain swapping to kinetic stability in a thioredoxin mutant. J Mol Biol 385:1590–1599.
22. Bai YW, Sosnick TR, Mayne L, Englander SW (1995) Protein-folding intermediates—native-state hydrogen-exchange. Science 269:192–197.