Published in final edited form as:

J Mol Biol. 2007 August 10; 371(2): 308–316.

Thermodynamic Analysis of an Antagonistic Folding-Unfolding Equilibrium Between Two Protein Domains

Thomas A. Cutler and Stewart N. Loh

Department of Biochemistry & Molecular Biology, SUNY Upstate Medical University, 750 East Adams Street, Syracuse NY 13210

Abstract

A simple model is formulated for analyzing the coupled folding-unfolding equilibrium present in a unique class of molecular switch proteins. We previously fused two single-domain proteins, barnase and ubiquitin, such that the free energy stored in the folded structure of one subunit is used to drive unfolding of the other. Here, we present a thermodynamic test of that mechanism. The antagonistic interaction is represented by a coupling free energy term ΔG_X . ΔG_X is the penalty imposed on folding of one domain by the native structure of the other. If $\Delta G_X = 0$, then neither domain senses the other and they fold and unfold independently. If $\Delta G_X > 0$, then destabilizing one domain will stabilize the other, and *vice versa*. In the limit where ΔG_X is greater than the intrinsic stability of either protein, then only one domain can be folded at any given time. We estimate ΔG_X by measuring stability parameters for a series of mutants which destabilize either the barnase or ubiquitin domains. Fitting the data to the model leads to a ΔG_X value of $\sim 4 \text{ kcal} \cdot \text{mol}^{-1}$. ΔG_X is proposed to depend on both the length of the linker peptides used to join the two proteins, and on the inherent structural plasticity of each domain. We predict that shortening the linkers from their current lengths of two and three amino acids will increase structural and thermodynamic coupling.

Multidomain proteins are most often comprised of two or more subunits connected in an end-to-end fashion¹. When these chimeras are created in the laboratory, the objective is usually for each domain to retain its original structure, stability and function, and to perturb the other as little as possible. For example, proteins are frequently tagged in this fashion to aid in their purification, crystallization, or detection. We previously took the opposite approach and fused two single-domain proteins so that they fold antagonistically^{2,3}. The goal is to create a bifunctional protein in which folding of one domain is coupled to unfolding of the other. According to this molecular switching mechanism, dubbed “mutually exclusive folding”, the molecule converts between two partially folded forms. The position of the switch, and hence the overall function of the protein, is controlled by factors that affect stability of the individual domains; for example, ligand binding, temperature, or pH.

In our first study, we inserted human ubiquitin (Ub, 76 amino acids) between residues 66 and 67 of the bacterial ribonuclease barnase (Bn, 110 amino acids), at the tip of a solvent exposed loop (Fig. 1)². In another study, the DNA binding domain from yeast GCN4 (56 amino acids) was grafted into the same site in Bn³. The coupled folding-unfolding equilibrium is established by a distance incompatibility: the amino to carboxy terminal lengths of Ub and GCN4 (when bound to its cognate DNA sequence) are 38 Å and 75 Å, respectively, while the distance between the ends of the Bn surface loop is only ~ 10 Å. The folded structures of Ub or GCN4

Address Correspondence to S.N.L.: Tel: (315) 464-8731, Fax: (315) 464-8750, Email: lohns@upstate.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

will therefore stretch Bn, splitting it in two if the amino to carboxy terminal distance is sufficiently long. The native form of Bn will compress and similarly destabilize Ub and GCN4. In this way, the free energy stored in the native state of one domain is used to mechanically unfold the other, and vice versa.

Here, we test the mutually exclusive folding mechanism by determining the extent to which the folding free energies of the Ub and Bn domains are coupled, in the barnase-ubiquitin fusion protein (BU) previously described². Known destabilizing mutations are introduced into one domain, and the change in stability of the other is determined by guanidine hydrochloride (GdnHCl) denaturation experiments. We chose to mutate Ub and observe Bn because the conformational state of Bn can be monitored readily by Trp fluorescence (all three Trp residues in BU are in the Bn domain). Results are interpreted using a minimal four-state model in which coupling is described by an energetic penalty that is imposed on the folding of one domain by the folded structure of the other. The model predicts that destabilization of one domain will stabilize the other, and vice versa. This relationship is expected to produce unusual behavior. For example, placing a helix-breaking mutation into an alpha helix may cause a distant region of the molecule to fold. Conversely, introducing a stabilizing interaction in one part of the protein may induce unfolding of another portion. Previous studies have used a similar mechanism to model mutual stabilization of domains by favorable binding interactions⁴. To our knowledge, however, there has been no corresponding analysis of antagonistic folding.

An additional motivation for this study arises from the observation that domains are comprised of non-sequential parts of the polypeptide chain in ~25 % of naturally occurring multidomain proteins^{1; 5}. This finding suggests that these proteins may have evolved by inserting one domain into a loop of another domain, as we have done here. It has been shown that proteins are generally tolerant to insertions of polypeptide sequences—even long ones—into loops, provided that the inserted sequences are unstructured^{6; 7; 8; 9; 10; 11}. The energetic and structural consequences of inserting folded domains into loops are less well known^{12; 13}. Our results may help to understand how proteins adapt to domain insertions and provide guidance for rational design of allosteric protein chimeras in the future¹⁴.

Choice of mutations

To retain consistency with the original literature, mutations are numbered according to the sequence of the parent protein in which they are found. Mutations V26A, V26G, K29N, K29Q and R42E are in the Ub domain. Val26 is one of the most highly buried side chains in the hydrophobic core of Ub. The V26A and V26G mutations were reported to reduce the folding free energy of free Ub by 3.3 kcal·mol⁻¹ and 6.3 kcal·mol⁻¹, respectively¹⁵. Lys29 is located on the surface where it participates in favorable electrostatic interactions¹⁶. The K29Q and K29N mutations were found to reduce stability of free Ub by 1.5 kcal·mol⁻¹ and 1.7 kcal·mol⁻¹, respectively¹⁶. Arg42 is also on the surface, but it is involved in unfavorable charge-charge interactions. The R42E mutation reverses the charge and was shown to stabilize free Ub by 1.6 kcal·mol⁻¹¹⁶. The I96V and I96A mutations are in the Bn domain. Ile96 is buried in the hydrophobic core. I96V and I96A were reported to be less stable than free WT Bn by 0.95 kcal·mol⁻¹ and 3.3 kcal·mol⁻¹, respectively¹⁷.

Model and simulations

The minimal folding mechanism is shown in Eq. 1, where underlined letters and non-underlined letters denote folded and unfolded domains, respectively.

$$\begin{array}{ccc}
 & \Delta G_U, m_U & \\
 bu & \rightleftharpoons & bu \\
 \Delta G_B, m_B \updownarrow & & \updownarrow (\Delta G_B - \Delta G_X), m_B \\
 bu & \rightleftharpoons & bu \\
 - & & - \\
 & (\Delta G_U - \Delta G_X), m_U &
 \end{array} \quad (1)$$

ΔG_U and ΔG_B are folding free energy changes of the Ub and Bn domains when the other is unfolded. Physically, this scenario corresponds to Ub folding with unstructured polypeptides at both ends, and Bn folding with an extended surface loop. The stabilities of the domains decrease with denaturant according to the linear extrapolation equation (Eqs. 2 and 3):

$$\Delta G_B = \Delta G_B^{H_2O} - m_B[\text{GdnHCl}] \quad (2)$$

$$\Delta G_U = \Delta G_U^{H_2O} - m_U[\text{GdnHCl}] \quad (3)$$

$\Delta G_B^{H_2O}$ and $\Delta G_U^{H_2O}$ are the values in the absence of denaturant, and m is proportional to the difference in accessible surface area between folded and unfolded states of each domain. m_B and m_U are expected to be close to the values measured for the respective free proteins. ΔG_X is assumed to remain constant as a function of denaturant. The fraction of each species is given by Eqs. 4–7, where $K_i = \exp(-\Delta G_i/RT)$:

$$f(bu) = 1 / (1 + K_B + K_U + K_B \cdot K_U \cdot K_X^{-1}) \quad (4)$$

$$f(bu) = K_U / (1 + K_B + K_U + K_B \cdot K_U \cdot K_X^{-1}) \quad (5)$$

$$f(bu) = K_B / (1 + K_B + K_U + K_B \cdot K_U \cdot K_X^{-1}) \quad (6)$$

$$f(bu) = K_B \cdot K_U \cdot K_X^{-1} / (1 + K_B + K_U + K_B \cdot K_U \cdot K_X^{-1}) \quad (7)$$

ΔG_X represents the energetic penalty that is imposed on the folding of one domain by the native structure of the other, according to the distance metric described above. If very long linker peptides are used to join the two proteins, then neither domain will be able to sense whether the other is folded or unfolded. In this case, $\Delta G_X = 0$, and each domain folds independently. As the linker peptides shorten, ΔG_X is expected to increase. If ΔG_X exceeds ΔG_U and ΔG_B , then both domains cannot exist simultaneously in their native states, and mutual exclusivity is attained. The linkers used in this study consist of Gly-Thr and Gly-Ala-Ser at the N and C-termini of Ub, respectively. We surmise that ΔG_X depends not only on distance considerations, but also on the properties of the native and unfolded states. It may contain information on the inherent plasticity of the folded domains, *i.e.*, the degree to which they can distort without unfolding. It is also important to recognize that the unfolded states produced by mechanical disruption and by chemical denaturation may be quite different. For example, when the nested Ub domain unfolds in the presence of folded Bn, its ends remain fixed in close proximity. The reduced entropy is expected to raise the free energy of unfolded Ub, relative to the unfolded conformation that would exist if the Bn domain were itself unstructured. On the other hand, forcing the termini together may not require Ub to unfold completely. One can envision partially folded or molten globule-like structures that are compatible with shorter N and C-terminal distances.

Figure 2 simulates fluorescence-monitored unfolding curves for three representative cases: zero coupling ($\Delta G_X = 0$), strong coupling ($\Delta G_X > \Delta G_B^{H_2O}, \Delta G_U^{H_2O}$), and intermediate coupling ($0 < \Delta G_X < \Delta G_B^{H_2O}, \Delta G_U^{H_2O}$). The fraction of folded Bn domain ($f(\underline{bu}) + f(\underline{bu})$) is plotted on the y-axis. The simulations employ parameters determined experimentally for WT BU ($\Delta G_B^{H_2O} = 7.5 \text{ kcal}\cdot\text{mol}^{-1}$, $m_B = 4.8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$, $\Delta G_U^{H_2O} = 5.2 \text{ kcal}\cdot\text{mol}^{-1}$, $m_U = 1.8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$; see next section). Results are shown for WT BU, and for a series of mutants in which the Ub domain is either stabilized or destabilized by $1 \text{ kcal}\cdot\text{mol}^{-1}$ and $2 \text{ kcal}\cdot\text{mol}^{-1}$.

In the zero coupling limit, the Bn and Ub domains denature independently (Fig. 2A). The observed Bn unfolding transition corresponds to $\underline{bu} \leftrightarrow \underline{bu}$ (for stable Ub variants) or to $\underline{bu} \leftrightarrow bu$ (for unstable Ub variants). In either case, the fitted parameters ($\Delta G_{app}^{H_2O}$ and m_{app}) are simply equal to $\Delta G_B^{H_2O}$ and m_B , respectively. The stability of the Bn domain is not connected to that of the Ub domain, as evidenced by the fact that WT and mutant unfolding curves are identical.

In the strong coupling threshold (Fig. 2B), state $\underline{bu}$ is never populated. The observed transition is described by $\underline{bu} \leftrightarrow \underline{bu}$ for all but the most destabilized Ub mutants. Ub folding now accompanies Bn unfolding. A direct result is that $\Delta G_{app}^{H_2O} = \Delta G_B^{H_2O} - \Delta G_U^{H_2O}$ and $m_{app} = m_B - m_U$. Stabilizing or destabilizing the Ub domain therefore produces an equal but opposite effect on the stability of the Bn domain. As in the zero coupling case, the Bn unfolding transition for highly unstable mutants corresponds to $\underline{bu} \leftrightarrow bu$, and $\Delta G_{app}^{H_2O}$ and m_{app} approach their maximum values of $\Delta G_B^{H_2O}$ and m_B .

More complex behavior is observed in the intermediate coupling regime (Fig. 2C). ΔG_X is low enough so that, for relatively stable Ub mutants, $\underline{bu}$ is the species most populated in the absence of denaturant. The Bn unfolding transition for WT BU is thus attributed largely to $\underline{bu} \leftrightarrow \underline{bu}$. Mutations that further stabilize the Ub domain have little effect on Bn unfolding, as they stabilize both $\underline{bu}$ and $\underline{bu}$ to comparable extents. Mutations that destabilize Ub, however, stabilize Bn significantly. The reason is that these mutations raise the free energy of $\underline{bu}$ but not that of $\underline{bu}$. Consequently, $\underline{bu}$ becomes more populated than $\underline{bu}$ in the absence denaturant. The observed transition switches to $\underline{bu} \leftrightarrow \underline{bu}$, as observed for the strong coupling condition. Finally, a third transition, $\underline{bu} \leftrightarrow bu$, is observed for extremely unstable Ub mutants, as in the previous two cases.

We now consider how to best quantitate stability changes in order to test the model. It is clear from the shift in the midpoints of denaturation (C_m) that, when $\Delta G_X > 0$, stabilizing or destabilizing the Ub domain produces the opposite effect on the Bn domain (Figs. 2B and 2C). However, up to three two-state reactions can contribute to the observed Bn unfolding transition. The individual reactions are not differentiated by Trp fluorescence: all data are fit well by the two-state linear extrapolation equation, as evident from the simulations. $\Delta G_{app}^{H_2O}$ and m_{app} values can oscillate as a consequence. For example, in the intermediate coupling regime (Fig. 2C), m_{app} drops from m_B to $m_B - m_U$, then rises again to m_B as the Ub domain is progressively destabilized by mutation. $\Delta G_{app}^{H_2O}$ fluctuates correspondingly. $\Delta G_{app}^{H_2O}$ is therefore not a suitable parameter for gauging stability of BU (although it is appropriate in the case of free Bn and free Ub, which unfold in a two-state manner). We instead use C_m , as it exhibits a consistent and predictable dependence on stability of the Ub domain in all coupling regimes (Fig. 2D). C_m can also be determined with greater accuracy and precision than $\Delta G_{app}^{H_2O}$. Accordingly, we express the change in stability of mutants as ΔC_m rather than $\Delta \Delta G_{app}^{H_2O}$.

Comparison with experimental results

Figure 3 shows GdnHCl denaturation curves for BU and free Bn. The conformational state of the Bn domain is assessed by fluorescence spectra of the three Trp residues of BU, all of which

are located in the Bn portion of the molecule. Folded Bn is identified by a maximum emission wavelength ($F_{\max}$) value near 336 nm. Plotting $F_{\max}$ as a function of denaturant reveals a cooperative transition to a final value of 356 nm for unfolded Bn (Fig. 3A). $F_{\max}$ is an effective means for ascertaining relative stabilities of BU variants from the primary data, but it is not ideal for obtaining thermodynamic parameters. For that purpose, we fit fluorescence intensities from the full Trp emission spectra. Multi-wavelength analysis is expected to provide better estimates of $\Delta G_{\text{app}}^{\text{H}_2\text{O}}$, m_{app} and C_m , as well as identify intermediates that may not be detected by the $F_{\max}$ method. Each wavelength from 305–405 nm was fit to the two-state linear extrapolation equation¹⁸ using the IGOR Pro software package (WaveMetrics, Inc.). $\Delta G_{\text{app}}^{\text{H}_2\text{O}}$ and m_{app} values were linked and fluorescence intensities of native and unfolded states (as well as baseline slope parameters) were allowed to float independently. Satisfactory fits are obtained for all BU variants, and plotting the residuals finds no systematic deviation at any wavelength. Several representative wavelengths are shown in Fig. 3B. This result confirms that Bn unfolding is adequately described by a two-state transition by fluorescence, despite the possibility that multiple individual reactions may be present.

To characterize unfolding of the entire BU molecule, we employed circular dichroism (CD) spectroscopy. CD spectra of WT and V26G BU are depicted in Fig. 3C. Ellipticity at 230 nm reveals two distinct transitions for WT BU (Fig. 3D). Fitting the data to a three-state unfolding mechanism yields $\Delta G_{\text{app}}^{\text{H}_2\text{O}} = 3.7 \text{ kcal}\cdot\text{mol}^{-1}$, $m_{\text{app}} = 4.3 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$, $C_m = 0.87 \text{ M}$ for the first transition. These values agree with those determined from fluorescence-monitored denaturation experiments (Fig. 3A and Table 1), thus; the first CD transition corresponds to the coupled Bn unfolding/Ub folding reaction. The second CD transition is assigned unambiguously to the $bu \leftrightarrow bu$ equilibrium, as it occurs at GdnHCl concentrations beyond the midpoint of Bn domain unfolding. Fitting the second transition as above yields $\Delta G_{\text{U}}^{\text{H}_2\text{O}} = 5.2 \text{ kcal}\cdot\text{mol}^{-1}$ and $m_{\text{U}} = 1.8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$. By comparison, free WT Ub exhibits $\Delta G_{\text{app}}^{\text{H}_2\text{O}}$ and m_{app} values of $6.9 \text{ kcal}\cdot\text{mol}^{-1}$ and $1.9 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$, respectively (Table 1). The presence of unfolded Bn fragments at both ends destabilizes Ub by $1.7 \text{ kcal}\cdot\text{mol}^{-1}$ and reduces C_m by 0.74 M .

To test the mechanism of Eq. 1, we introduced stabilizing and destabilizing mutations into the Ub domain of BU and monitored unfolding of the Bn domain. The model holds that destabilizing Ub will stabilize Bn. This prediction is supported by the primary CD data. Figure 3D shows that the K29N mutation shifts C_m of the Bn and Ub unfolding transitions to higher and lower denaturant concentrations, respectively. This trend continues so that for V26A and V26G only a single transition is observed (data not shown). Moreover, multi-wavelength analysis of CD data is made largely ineffective by the absorptivity of GdnHCl, which blocks wavelengths below ~215 nm. For these reasons, we used fluorescence data exclusively to obtain thermodynamic parameters. As expected, insertion of Ub strongly destabilizes Bn. C_m of the Bn domain of WT BU is lower than that of free Bn by 1.7 M (Table 1). The stability of Ub variants follows the order $\text{R42E} > \text{WT} > \text{K29Q} > \text{K29N} > \text{V26A} > \text{V26G}$. Figure 3A and Table 1 show that C_m values of the Bn domain increase with the opposite rank order. The stability of the Bn domain is inversely proportional to the stability of the Ub domain, and the zero coupling condition (Fig. 2A) can be excluded.

We applied the following strategy to estimate ΔG_X . The V26G mutation lowers the stability of free Ub to $1.8 \text{ kcal}\cdot\text{mol}^{-1}$ and its C_m to 0.72 M (Table 1). Figure 3D suggests that the presence of Bn will further decrease these values by approximately $1.7 \text{ kcal}\cdot\text{mol}^{-1}$ and 0.74 M , respectively. The Ub domain of V26G BU is therefore predicted to be intrinsically unstable in the absence of denaturant. Consequently, the observed unfolding transition of V26G BU should closely approximate $bu \leftrightarrow bu$. $\Delta G_{\text{B}}^{\text{H}_2\text{O}}$ and m_{B} can thus be assigned to the $\Delta G_{\text{app}}^{\text{H}_2\text{O}}$ ($7.5 \text{ kcal}\cdot\text{mol}^{-1}$) and m_{app} ($4.8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$) values measured for V26G BU. Bn appears to be destabilized by $4.3 \text{ kcal}\cdot\text{mol}^{-1}$ by the presence of unfolded Ub (Table 1). This value is

considerably greater than the amount that Ub is destabilized by unfolded Bn ($1.7 \text{ kcal}\cdot\text{mol}^{-1}$). The result is expected because in the former case, Bn folding requires closing of a lengthy, unstructured surface loop. The decrease in entropy associated with loop closure can be approximated by Eq. 8:

$$\Delta S = a - (3/2)R\ln N \quad (8)$$

where a is related to the distance between ends of the closed loop (-2 to $-8 \text{ cal}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$) and N is the number of amino acids in the loop^{8; 19}. The observed decrease in $\Delta G_{\text{app}}^{\text{H}_2\text{O}}$ falls within the range of $T\Delta S$ values defined by Eq. 8.

We then plotted experimental C_m values of the Bn domain against the change in stability of the Ub domain, expressed as ΔC_m (Fig. 4A). The latter quantity was determined by separate experiments on the same mutants in free Ub (Table 1). $\Delta G_U^{\text{H}_2\text{O}}$ ($5.2 \text{ kcal}\cdot\text{mol}^{-1}$) and m_U ($1.8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$) were obtained from Fig. 3D as described. ΔG_X was then adjusted manually, and theoretical C_m values were calculated from simulated BU unfolding curves such as those in Fig. 2. The lines in Fig. 4A are the results of simulations using ΔG_X values of 3, 4 and 5 $\text{kcal}\cdot\text{mol}^{-1}$.

A reasonable fit is obtained with $\Delta G_X \sim 4.0 \text{ kcal}\cdot\text{mol}^{-1}$ (Fig. 4A). All values of ΔG_X shown in the figure are able to reproduce the behavior of the most unstable Ub variants. To determine ΔG_X more accurately, it is necessary to extend the data set to the regime where Ub is stabilized relative to WT. Unfortunately, few such variants have been reported. The R42E mutation stabilizes WT Ub by $1.6 \text{ kcal}\cdot\text{mol}^{-1}$ when measured by urea denaturation¹⁶. GdnHCl appears to screen the favorable effect of charge reversal, however, and R42E Ub is only $0.3 \text{ kcal}\cdot\text{mol}^{-1}$ more stable than WT in our experiments (Table 1). An alternate approach is to measure the extent to which mutations destabilize the Bn domain of BU relative to free Bn. We introduced the I96V and I96A mutations into BU and free Bn and repeated the analysis of Fig. 4A. Because native baselines are not well defined for these unstable mutants, $\Delta G_{\text{app}}^{\text{H}_2\text{O}}$ and m_{app} values could not be determined confidently by the procedure used for the other mutants. C_m values, however, could be obtained reasonably accurately by fitting F_{max} data while constraining the native F_{max} to the average value of the other variants (337 nm) (Fig. 3A). The results (Fig. 4B) confirm that a ΔG_X value of $3\text{--}4 \text{ kcal}\cdot\text{mol}^{-1}$ fits the data adequately.

In our earlier paper we concluded that Bn and Ub domain folding are mutually exclusive in WT BU². That deduction was drawn from variable temperature CD experiments. As temperature increased from $5\text{--}50^\circ\text{C}$, a cooperative positive shift in the 230 nm peak was observed, signifying Bn unfolding (the 230 nm minimum arises from Trp94 of Bn and is diagnostic of folded Bn structure²⁰). The 208 nm peak showed a corresponding negative change and an isodichroic point was noted at 223 nm . Since the molar ellipticity of Ub at 208 nm is more negative than that of Bn, those findings argue that the Ub domain is unfolded at temperatures below 30°C , and that it folds as Bn is denatured by heat or GdnHCl. The current results, however, place ΔG_X in the intermediate range which implies that folding is not fully exclusive. For example, ΔG_X values of $3\text{--}5 \text{ kcal}\cdot\text{mol}^{-1}$ predict that $98\text{--}58\%$ of the WT protein exists in conformation *bu* in the absence of denaturant. Which view is correct?

As a first test we compared CD spectra of WT and V26G BU (Fig. 3C). If Bn folding causes Ub to unfold completely, then the two spectra should be identical. The region below 200 nm is particularly diagnostic of unfolded structure. The ellipticity of V26G is markedly lower at those wavelengths, indicating a greater amount of random coil conformation. The more positive signal observed above 205 nm is also consistent with a lower α -helix and/or β -sheet content in V26G. We attempted to estimate secondary structure content using the SELCON3²¹ and CONTIN/LL^{22; 23} programs. Both methods agree that V26G contains a greater percentage of unfolded structure compared to WT ($7\text{--}15\%$ increase, using the SDP42 reference protein

set). This figure should be considered as only a rough estimate, as agreement between experimental and theoretical data is poor above ~205 nm (presumably due to the anomalous minimum at 230 nm arising from Trp94). Nevertheless, CD results suggest that the Ub domain is folded in a fraction of WT BU molecules.

A second test is provided by m_{app} values. The model holds that m_{app} of BU variants with unstable Ub domains (*e.g.* V26G) should be similar to m_{app} of free Bn. That result is indeed found (Table 1). If Bn and Ub folding are completely exclusive in WT BU, then m_{app} is expected to decrease by the full value of $1.8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$ observed for free Ub. m_{app} of WT BU is smaller than that of V26G, but only by $0.5 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$. Thus, CD and thermodynamic data together suggest that both domains are folded in a measurable percentage of WT BU molecules. Denaturing the Bn domain by GdnHCl or temperature appears to allow the fraction of unfolded Ub domains to refold, which forms the basis of our earlier conclusion.

We now turn the discussion to limitations of the model. Although it reproduces the important trends in the experimental data, it does not capture some details. The dependence of C_m on Ub stability appears to be steeper and more cooperative than the model predicts (Fig. 4A). In addition, while Bn m_{app} values exhibit the expected dependence on Ub stability (a decrease then an increase as the Ub domain is progressively destabilized), the experimental figures exceed theoretical values by $0.2\text{--}0.4 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$ for most mutants. One reason may be that the four-state model neglects the possibility of the two domains associating in their folded states. Additional molecular surface could be buried as a result of such an interaction. A related assumption is that the individual domains unfold in a two-state manner, *i.e.*, that the unfolded states produced by mechanical disruption and by GdnHCl denaturation are structurally and energetically equivalent. This is most likely not the case, as pointed out earlier. The unfolded conformations of the Ub domain in *bu* (populated in the absence of GdnHCl) and *bu* (present in high GdnHCl concentrations) are different, if for no other reason than the termini of Ub are constrained in the former. There may also be a significant amount of residual structure present in mechanically disrupted Ub. These issues can be addressed by incorporating additional states into the folding model. Such a mechanism would contain too many parameters to be characterized by the present data set. It may be possible, however, to test it using higher resolution structural techniques such as nuclear magnetic resonance studies of BU variants in different concentrations of denaturant.

In conclusion, we have established a thermodynamic framework for analyzing the coupled folding-unfolding equilibrium of the new class of bi-functional molecules exemplified by BU and barnase-GCN4. The coupling free energy of $\sim 4 \text{ kcal}\cdot\text{mol}^{-1}$ is less than or comparable to $\Delta G_B^{\text{H}_2\text{O}}$ and $\Delta G_U^{\text{H}_2\text{O}}$ values of most mutants, suggesting that coupling is in the intermediate regime. Surprisingly, the model predicts that both domains are folded in a significant fraction of WT BU molecules. We propose that shortening the linker peptides from their current lengths of two and three amino acids will intensify the strain that the folded structures of Bn and Ub impose on each other, and thus increase ΔG_X . We are currently testing that hypothesis by experimental and computational studies of BU variants with progressively shortened linkers.

Acknowledgements

We thank Drs. James S. Butler and Jeung-Hoi Ha for discussions. This work was supported by N.I.H. Grant R01GM069755.

References

1. Jones S, Stewart M, Michie A, Swindells MB, Orengo C, Thornton JM. Domain assignment for protein structures using a consensus approach: characterization and analysis. *Protein Sci* 1998;7:233–242. [PubMed: 9521098]

2. Radley TL, Markowska AI, Bettinger BT, Ha JH, Loh SN. Allosteric switching by mutually exclusive folding of protein domains. *J Mol Biol* 2003;332:529–536. [PubMed: 12963365]
3. Ha JH, Butler JS, Mitrea DM, Loh SN. Modular enzyme design: regulation by mutually exclusive protein folding. *J Mol Biol* 2006;357:1058–1062. [PubMed: 16483603]
4. Brandts JF, Hu CQ, Lin LN, Mas MT. A simple model for proteins with interacting domains. Applications to scanning calorimetry data. *Biochemistry* 1989;28:8588–8596. [PubMed: 2690944]
5. Holm L, Sander C. The FSSP database of structurally aligned protein fold families. *Nucleic Acids Res* 1994;22:3600–3609. [PubMed: 7937067]
6. Ferraro DM, Hope EK, Robertson AD. Site-specific reflex response of ubiquitin to loop insertions. *J Mol Biol* 2005;352:575–584. [PubMed: 16081099]
7. Ferraro DM, Ferraro DJ, Ramaswamy S, Robertson AD. Structures of ubiquitin insertion mutants support site-specific reflex response to insertions hypothesis. *J Mol Biol* 2006;359:390–402. [PubMed: 16647719]
8. Scalley-Kim M, Minard P, Baker D. Low free energy cost of very long loop insertions in proteins. *Protein Sci* 2003;12:197–206. [PubMed: 12538883]
9. Ladurner AG, Fersht AR. Glutamine, alanine or glycine repeats inserted into the loop of a protein have minimal effects on stability and folding rates. *J Mol Biol* 1997;273:330–337. [PubMed: 9367765]
10. Nagi AD, Regan L. An inverse correlation between loop length and stability in a four-helix-bundle protein. *Folding Design* 1997;20:67–75. [PubMed: 9080200]
11. Doi N, Itaya M, Yomo T, Tokura S, Yanagawa H. Insertion of foreign random sequences of 120 amino acid residues into an active enzyme. *FEBS Lett* 1997;402:177–180. [PubMed: 9037190]
12. Collinet B, Hervé M, Pecorari F, Minard P, Eder O, Desmadril M. Functionally accepted insertions of proteins within protein domains. *J Biol Chem* 2000;275:17428–17433. [PubMed: 10747943]
13. Betton JM, Jacob JP, Hofnung M, Broome-Smith JK. Creating a bifunctional protein by insertion of β -lactamase into the maltodextrin-binding protein. *Nature Biotechnol* 1997;15:1276–1279. [PubMed: 9359111]
14. Ostermeier M. Engineering allosteric protein switches by domain insertion. *Protein Eng Des Sel* 2005;18:359–364. [PubMed: 16043448]
15. Khorasanizadeh S, Peters ID, Roder H. Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nature Struct Biol* 1996;3:193–205. [PubMed: 8564547]
16. Loladze VV, Ibarra-Molero B, Sanchez-Ruiz JM, Makhataдзе GI. Engineering a thermostable protein via optimization of charge-charge interactions on a protein surface. *Biochemistry* 1999;38
17. Kellis JT, Nyberg K, Fersht AR. Energetics of complementary side-chain packing in a protein hydrophobic core. *Biochemistry* 1989;28:4914–4922. [PubMed: 2669964]
18. Santoro MM, Bolen DW. Unfolding free energy changes determined by the linear extrapolation method. 1 Unfolding of phenylmethanesulfonyl alpha chymotrypsin using different denaturants. *Biochemistry* 1988;27:8063–8068. [PubMed: 3233195]
19. Chan HS, Dill KA. Intrachain loops in polymers. *J Chem Phys* 1988;90:492–509.
20. Vuilleumier S, Sancho J, Loewenthal R, Fersht AR. Circular dichroism studies of barnase and its mutants: characterization of the contribution of aromatic side chains. *Biochemistry* 1993;32:10303–10313. [PubMed: 8399173]
21. Sreerama N, Woody RW. A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal Biochem* 1993;209:32–44. [PubMed: 8465960]
22. Sreerama N, Woody RW. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSTTR methods with an expanded reference set. *Anal Biochem* 2000;287:252–260. [PubMed: 11112271]
23. Provencereh SW, Glockner J. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* 1981;30:33–37.

Figure 1.

Design of the BU fusion protein. Ub (top) was inserted into a surface loop of Bn (bottom) between residues 66 and 67 (indicated by asterisk). Locations of mutated residues are shown in black. BU variants were cloned, expressed, and purified as described previously².

Figure 2.

Simulated Bn domain unfolding curves for three representative coupling regimes: (A) zero coupling ($\Delta G_X = 0$), (B) strong coupling ($\Delta G_X = 20 \text{ kcal}\cdot\text{mol}^{-1}$), (C) intermediate coupling ($\Delta G_X = 4 \text{ kcal}\cdot\text{mol}^{-1}$). WT BU is depicted by the heavy black line. Variants in which the Ub domain is stabilized are shown in gray. Gray curves to the immediate left and far left of WT are mutants stabilized by $1 \text{ kcal}\cdot\text{mol}^{-1}$ and $2 \text{ kcal}\cdot\text{mol}^{-1}$, respectively. Variants in which the Ub domain is destabilized are shown in black. Black curves to the immediate right and far right of WT are mutants destabilized by $1 \text{ kcal}\cdot\text{mol}^{-1}$ and $2 \text{ kcal}\cdot\text{mol}^{-1}$, respectively. Free Bn is indicated by the dashed line. Curves are generated from the model in Figure 2 using the following parameters: $\Delta G_B = 7.5 \text{ kcal}\cdot\text{mol}^{-1}$, $m_B = 4.8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$, $\Delta G_U = 5.2 \text{ kcal}\cdot\text{mol}^{-1}$ (WT), $m_B = 1.8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$. Panel D shows the dependence Bn domain stability on Ub domain stability, as gauged by C_m values. The zero, strong, and intermediate coupling cases are represented by gray, solid black, and dashed black lines, respectively. Bn domain C_m values are derived from fitting the data in panels A–C to the two-state linear extrapolation equation. ΔC_m and $\Delta\Delta G^{\text{H}_2\text{O}}$ values on the x-axes are relative to WT Ub.

Figure 3.

(A) GdnHCl-induced unfolding curves for BU variants, monitored by Trp fluorescence maximum. F_{max} values were calculated using the DATAMAX software package (Jobin-Yvon/SPEX). Lines are best fits to the linear extrapolation equation. Symbols are as follows: filled black circles, WT; crosses, R42E; diamonds, K29Q; inverted triangles, K29N; filled triangles, V26A; squares, V26G; crosshatched squares, I96V; half-filled squares, I96A; open circles and dashed line, free WT Bn. (B) GdnHCl-induced unfolding of K29N BU monitored by

fluorescence intensity at 320 nm (open circles), 340 nm (open squares), 360 nm (closed circles), 380 nm (closed squares). Lines are best fits to the linear extrapolation equation, as determined by multi-wavelength fits of Trp emission spectra from 305–405 nm. Curve-fit residuals are shown in the upper panel. (C) CD spectra of WT BU (solid line) and V26G BU (dashed line). (D) GdnHCl-induced unfolding of WT (closed circles) and K29N (inverted triangles) BU monitored by circular dichroism signal at 230 nm. Lines are best fit to the three-state linear extrapolation equation. Sample conditions are listed in Table 1.

Figure 4.

(A) Stability of the Bn domain of BU, reported by C_m values, as a function of the change in stability of free Ub, determined by ΔC_m values. ΔC_m is relative to free WT Ub. Circles represent the following variants, from left to right: V26G, V26A, K29N, K29Q, WT, R42E. Lines are simulated results using ΔG_X values of 3 kcal·mol⁻¹, 4 kcal·mol⁻¹ and 5 kcal·mol⁻¹ (from upper to lower curves); see text for other parameters. (B) Stability of the Bn domain of BU as a function of the stability of free Bn. ΔC_m is relative to free WT Bn. Data points are, from left to right: I96A, I96V, WT. Solid lines are simulated results for ΔG_X values of 3 kcal·mol⁻¹, 4 kcal·mol⁻¹ and 5 kcal·mol⁻¹ (from left to right). Dashed line is the simulated result for $\Delta G_X = 0$.

Table 1

Thermodynamic parameters for unfolding of free Ub, free Bn, and the Bn domain of BU. Bn and BU parameters are determined from multi-wavelength fits of fluorescence data. Ub parameters are obtained from fitting CD ellipticity at 230 nm. Sample conditions are 1–2 μ M protein, 10 mM sodium acetate (pH 5.0), 0.1 M NaCl, 10 °C. Reported values are the averages of three measurements. Standard deviations are typically ± 0.2 – 0.4 kcal·mol⁻¹ ($\Delta G_{\text{app}}^{\text{H}_2\text{O}}$), ± 0.1 – 0.3 kcal·mol⁻¹·M⁻¹ (m_{app}), ± 0.01 – 0.02 M (C_m).

Protein	Variant	$\Delta G_{\text{app}}^{\text{H}_2\text{O}}$ (kcal·mol ⁻¹)	m_{app} (kcal·mol ⁻¹ ·M ⁻¹)	C_m (M)
Ub	WT	6.9	1.9	3.63
Ub	R42E	7.2	1.8	4.00
Ub	K29Q	6.7	1.8	3.72
Ub	K29N	5.7	1.9	3.00
Ub	V26A	4.4	2.1	2.10
Ub	V26G	1.8	2.5	0.72
Bn	WT	11.8	4.6	2.58
Bn	I96V	8.9	3.8	2.38
Bn	I96A	8.7	4.5	1.96
BU	WT	3.8	4.3	0.87
BU	R42E	3.7	4.2	0.87
BU	K29Q	3.6	4.0	0.90
BU	K29N	4.0	3.8	1.04
BU	V26A	7.3	4.6	1.57
BU	V26G	7.5	4.8	1.56
BU	I96V	n.d. ^a	n.d.	0.72 ^b
BU	I96A	n.d.	n.d.	0.35 ^b

^a Not determined due to insufficient native baseline.

^b Obtained from constrained fits of F_{max} values (see text).