

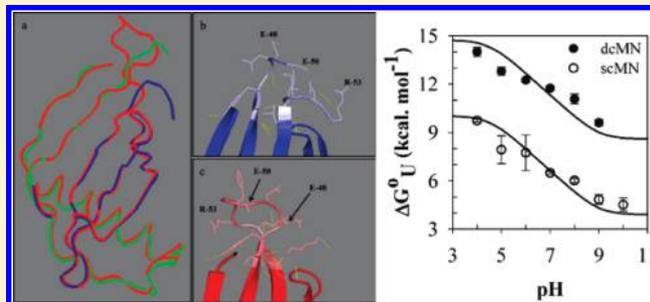
Equilibrium Unfolding Studies of Monellin: The Double-Chain Variant Appears To Be More Stable Than the Single-Chain Variant

Nilesh Aghera,[†] Ninganna Earanna,[‡] and Jayant B. Udgaonkar^{*,†}

[†]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560065, India

[‡]Department of Biotechnology, University of Agricultural Sciences, Bangalore 560065, India

ABSTRACT: To improve our understanding of the contributions of different stabilizing interactions to protein stability, including that of residual structure in the unfolded state, the small sweet protein monellin has been studied in both its two variant forms, the naturally occurring double-chain variant (dcMN) and the artificially created single-chain variant (scMN). Equilibrium guanidine hydrochloride-induced unfolding studies at pH 7 show that the standard free energy of unfolding, ΔG°_U , of dcMN to unfolded chains A and B and its dependence on guanidine hydrochloride (GdnHCl) concentration are both independent of protein concentration, while the midpoint of unfolding has an exponential dependence on protein concentration. Hence, the unfolding of dcMN like that of scMN can be described as two-state unfolding. The free energy of dissociation, ΔG°_d , of the two free chains, A and B, from dcMN, as measured by equilibrium binding studies, is significantly lower than ΔG°_U , apparently because of the presence of residual structure in free chain B. The value of ΔG°_U , at the standard concentration of 1 M, is found to be ~ 5.5 kcal mol⁻¹ higher for dcMN than for scMN in the range from pH 4 to 9, over which unfolding appears to be two-state. Hence, dcMN appears to be more stable than scMN. It seems that unfolded scMN is stabilized by residual structure that is absent in unfolded dcMN and/or that native scMN is destabilized by strain that is relieved in native dcMN. The value of ΔG°_U for both protein variants decreases with an increase in pH from 4 to 9, apparently because of the thermodynamic coupling of unfolding to the protonation of a buried carboxylate side chain whose pK_a shifts from 4.5 in the unfolded state to 9 in the native state. Finally, it is shown that although the thermodynamic stabilities of dcMN and scMN are very different, their kinetic stabilities with respect to unfolding in GdnHCl are very similar.



Proteins seem to be evolutionarily selected to possess marginal stability, which not only is critical in many aspects of cellular physiology^{1,2} but also allows proteins to respond rapidly to a change in environmental conditions.³ A consequence of the marginal stability of proteins is that understanding the relative contributions made by different stabilizing interactions becomes difficult, especially in the context in which intrachain and chain-solvent interactions are similar in type and similarly large in magnitude in the folded and unfolded states. A large body of work has implicated hydrophobic interactions⁴⁻⁶ and the loss of conformational entropy^{7,8} as the predominant determinants of protein stability, although more recently, the importance of hydrogen bonding and packing interactions in maintaining the integrity of protein structure is being recognized again.⁹⁻¹² Multimeric proteins are stabilized by the same type of intrachain and chain-solvent interactions as monomeric proteins, but in addition, interchain interactions contribute significantly to multimeric protein stability.¹³⁻¹⁶ Understanding the contribution of interchain interactions to protein stability is important because cooperative intersubunit communication plays an important role in allowing many of the regulatory aspects of protein function that appear to be unique to multimeric proteins and oligomeric protein assemblies.¹⁷⁻¹⁹

Equilibrium unfolding studies of multimeric proteins can delineate the energetics of the processes of chain association and folding²⁰⁻²⁴ and are therefore of potential utility in understanding the folding of intrinsically disordered proteins that fold only upon association with their binding partners.²⁵⁻²⁷ Thermodynamic characterization of the unfolding of multimeric proteins is often not possible because of the difficulty in finding conditions under which they unfold reversibly. Nevertheless, many multimeric proteins, including homodimeric P22 Arc repressor,¹³ GCN4,²⁸ Trp repressor²⁹ and YibK,³⁰ heterodimeric leucine zipper AB,³¹ histone H2A/H2B,³² and mannose-binding lectin,³³ as well as the larger oligomers secB³⁴ and GroES/cpn10^{35,36} have proven to be amenable to detailed thermodynamic analysis. In particular, much has been learned from these studies about how enthalpy and entropy can change in unforeseen ways during subunit association and folding.

Studies comparing the thermodynamics of unfolding of a multimeric protein and its monomeric (single-chain) counterpart with the same sequence can provide insight into the contribution

Received: December 8, 2010

Revised: February 17, 2011

Published: February 25, 2011

to protein stability of the loss in chain entropy that accompanies folding and association, because the loss will be much greater for the multimeric protein with the same sequence.³⁷ Such studies can also provide insight into the contribution to native state stability of residual structure in the unfolded state³⁸ because the multimeric protein is less likely to possess such structure than its single-chain counterpart. Such studies have been restricted either to monomeric proteins for which heterodimeric counterparts have been created by chain cleavage and/or fragmentation^{39–43} or to multimeric proteins for which monomeric variants have been created by chain fusion at the genetic level.^{15,44,45} An attractive model system for understanding the roles that may be played by chain entropy and residual unfolded state structure in protein stability is the sweet plant protein monellin.

Naturally occurring monellin (dcMN), isolated originally from the berry of an African plant, *Dioscoreophyllum cuminisii*,⁴⁶ is heterodimeric. Its two chains, A and B, are held together by interchain hydrophobic interactions, H-bonds, and salt bridges,^{47–49} and its biological activity (sweetness) is dependent on the integrity of its secondary, tertiary, and quaternary structure.^{50–55} A sweet single-chain variant of monellin (MNEI or scMN) has been created through genetic fusion, by joining the C-terminal end of chain B to the N-terminal end of chain A with a Gly-Phe linker. Single-chain monellin has been used extensively as a model protein for folding studies.^{11,56–60} There have been relatively few studies of the folding of dcMN,^{49,61} partly because of the heterogeneity observed in the commercially available protein,^{61,62} but the recent availability of a good bacterial expression system⁶³ has made dcMN an attractive model protein for the study of the folding of a heterodimeric protein. The structures of dcMN and scMN are nearly identical (Figure 1a), except at the linker region of scMN (Figure 1b,c), where there appears to be a rearrangement of hydrogen bonds. The double-chain and single-chain variants of monellin appear to be an ideal system for investigating the contributions of intrachain and interchain contributions to protein stability and folding.

We conducted equilibrium binding and unfolding studies with dcMN. The GdnHCl-induced unfolding studies indicate that unfolding is two-state: the fluorescence and far-UV circular dichroism (CD)-monitored transitions are coincident, and the free energy of unfolding as well as its dependence on GdnHCl concentration is independent of protein concentration. It appears that unbound B chain, but not unbound A chain, possesses residual structure in isolation. We also conducted equilibrium unfolding studies with scMN and found that scMN is substantially less stable than dcMN, even though their structures and spectroscopic properties are very similar. The higher stability of dcMN over scMN is seen over the entire range from pH 4 to 9 that was studied. The pH dependencies of the stabilities of the two proteins are found to be similar. It appears that one reason why dcMN is more stable than scMN could be that the unfolded state of scMN may be stabilized by residual structure.

MATERIALS AND METHODS

Reagents. All the reagents used were from Sigma and were of the highest purity. GdnHCl was purchased from USB and was of the highest purity. All the experiments were conducted at pH 7 (unless otherwise noted) in 50 mM phosphate buffer containing 0.25 mM EDTA and 1 mM DTT.

Preparation of dcMN, scMN, and the Chains of Monellin. The method for the purification of single-chain monellin (MNEI,

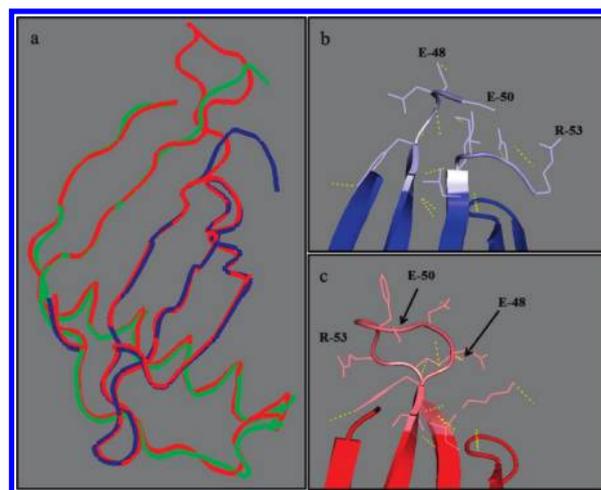


Figure 1. Structural comparison of double-chain monellin (dcMN) and single-chain monellin (scMN). Panel a shows the alignment of the main chain peptide backbones of dcMN (blue and green) and scMN (red). dcMN has two chains: chain A (blue, 5.4 kDa), comprising three antiparallel β -strands, and chain B (green, 6 kDa), comprising one long α -helix intersecting with two antiparallel β -strands. The five β -strands form a single sheet. Panels b and c show the structure adjoining the linker region of scMN in both proteins. The yellow dotted lines represent hydrogen bonds formed within the protein molecule. The linker region of scMN appears to be constrained, with the ϕ and ψ angles of Glu50 falling out of the allowed regions of the Ramachandran plot, and with Glu48 and Arg53 appearing to have unfavorable contacts. The structures were drawn using PYMOL (<http://www.pymol.org>) from Protein Data Bank entries 3MON⁸⁶ and 1IV7 (Kato et al., manuscript to be published) for dcMN and scMN, respectively.

here called scMN) has been described previously.⁵⁸ dcMN was purified by expressing the two subunits of the protein in *Escherichia coli* [BL21*(DE3)] using a pET duet vector system.⁶³ The protein was purified using a protocol similar to that described in ref 58 for the purification of scMN (MNEI). The sequences of the expressed proteins have been reported previously.^{58,63} The purity of each protein was confirmed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and mass spectrometry.

The chains of monellin were purified from dcMN by using reverse phase chromatography (RPC). Separation was performed as described previously.⁶³ For separation, the amount of protein for each run of RPC was restricted to 4 mg of double-chain monellin. After separation, methanol was removed by nitrogen purging. Briefly, nitrogen gas was blown onto the surface of each solution, with care being taken to see that no bubbles formed in the solution. The volume was allowed to reduce from 25 (100 μ M) to 5 mL by evaporation, increased to 10 mL by the addition of water, and then reduced again to 5 mL. The chains were then lyophilized and stored at -20 °C. Before the chains were used, they were incubated in 6 M GdnHCl for 2 h, and the solution was then desalted by the use of a Hi Trap desalting column (GE). Freshly desalted chains were used for all experiments. It was found that this procedure resulted in chains whose CD spectra and GdnHCl-induced unfolding curves did not vary from one preparation to the next. The purity of each protein or polypeptide was confirmed by SDS–PAGE and mass spectrometry.

Protein labeling with TNB was conducted using DTNB [*S'*,*S'*-dithiobis(2-nitrobenzoic acid)], as described previously.⁶⁴

Fluorescence and CD Measurements. Fluorescence spectra were recorded on a Fluoromax-4 spectrofluorimeter (Horiba) at

a protein concentration of 10 μM . The protein was excited at 280 nm with a 0.6 nm bandwidth, and emission was collected using a 10 nm bandwidth. The spectra of the TNB labeled proteins were recorded by exciting the protein at 295 nm using the same bandwidth settings. CD spectra were recorded using a Jasco J-815 spectropolarimeter. Far-UV spectra were recorded using a 0.2 cm cuvette at a protein concentration of 10 μM , using a scan speed of 50 nm/min and a DIT of 2 s. Near-UV CD spectra were recorded at a protein concentration of 100 μM using a 10 mm path length cuvette with the same instrument settings that were used for acquiring the far-UV CD spectra. CD spectra were recorded at a bandwidth of 1 nm.

Equilibrium Binding Study. Chain B (0.25 μM) was mixed with varying concentrations of chain A, in the range of 0–2 μM . Each solution was incubated for 3 h, before measurement of its fluorescence. For fluorescence measurements, the protein was excited at 280 nm and emission was collected at 340 nm using a band-pass filter (Asahi spectra). Fluorescence measurements were taken with the MOS 450 optical system from Biologic.

Equilibrium Unfolding Studies. GdnHCl-induced equilibrium unfolding transitions were monitored by measurement of fluorescence intensity using the MOS 450 optical system. dcMN was incubated in different concentrations of GdnHCl from 0 to 3 M for 48 h, until equilibrium was established, before measurement of its fluorescence signal. For fluorescence measurements, the protein sample was excited at 280 nm and emission was collected at 340 nm using a 10 nm band-pass filter (Asahi spectra). At pH 7, equilibrium unfolding curves were measured at dcMN concentrations ranging from 1 to 40 μM .

Equilibrium unfolding curves were measured at pH 4–10, at protein concentrations of 10 and 5–10 μM for dcMN and scMN, respectively; 50 mM sodium acetate buffer was used at pH 4 and 5, 50 mM sodium phosphate buffer at pH 6 and 7, 50 mM Tris buffer at pH 8 and 9, and 50 mM sodium borate buffer at pH 10. In addition, the solutions contained 0.25 mM EDTA and 1 mM DTT. The equilibrium unfolding transitions were monitored using fluorescence and far-UV CD as probes. At each pH, each protein was incubated in different concentrations of GdnHCl, for 72 h (dcMN) or 6 h (scMN). Fluorescence measurements were taken as described above. CD measurements at 222 nm were taken on a Jasco J-720 instrument, using a 0.2 cm path length cuvette. For each measurement, the CD signal was averaged for 120 s, using a response time of 4 s.

GdnHCl-induced equilibrium unfolding transitions of chain B were performed at a concentration of 20 μM and were monitored by CD at pH 7. Chain B was incubated for 2 h in different concentrations of GdnHCl from 0 to 3.5 M before CD measurements. CD measurements were taken with the same instrument settings described above.

Unfolding Kinetics. The kinetics of unfolding of dcMN and scMN were monitored by measurement of fluorescence at 340 nm upon excitation at 280 nm, using the MOS 450 optical system. Unfolding kinetic studies were performed by manual mixing; the dead time of the experiment was 10 s. Unfolding studies of dcMN and scMN were conducted at protein concentrations of 10 and 5 μM , respectively.

Data Analysis. *Two-State Model for the GdnHCl-Induced Equilibrium Unfolding of scMN.* A two-state $\text{N} \leftrightarrow \text{U}$ model was used for the analysis of the equilibrium unfolding of scMN, as described previously.^{58,65}

Two-State Model for the Dissociation of dcMN into Chains A and B. The equilibrium data for the binding of chain A (A) to

chain B (B) to form native dcMN (N) were analyzed on the basis of the following scheme:



The dissociation constant, K_d , is defined as

$$K_d = \frac{[\text{A}][\text{B}]}{[\text{N}]} \quad (1)$$

Under standard conditions where N, A, and B are all present at concentrations of 1 M, at 25 °C and pH 7, the free energy of dissociation is equal to the standard free energy and, hence, is given by

$$\Delta G^\circ_d = -RT \ln K_d \quad (2)$$

In a binding study in which $[\text{A}_0]$ and $[\text{B}_0]$ are the concentrations of chain A and chain B added to each other

$$K_d = \frac{([\text{A}_0] - [\text{N}])([\text{B}_0] - [\text{N}])}{[\text{N}]} \quad (3)$$

Solving eq 3 for $[\text{N}]$ gives

$$[\text{N}] = \frac{[\text{A}_0] + [\text{B}_0] + K_d \pm \sqrt{([\text{A}_0] + [\text{B}_0] + K_d)^2 - 4[\text{A}_0][\text{B}_0]}}{2} \quad (4)$$

If binding is monitored by the measurement of the change in fluorescence intensity of chain B that occurs upon addition of chain A to a fixed concentration of chain B, then the fluorescence intensity, F , at any concentration of chain A is given by

$$F = F_B + \Delta F[\text{N}] \quad (5)$$

where F_B is the fluorescence intensity of chain B in the absence of any chain A and ΔF is change in fluorescence intensity that occurs upon formation of a unit concentration of N. The equilibrium binding data were analyzed using eqs 4 and 5.

Equilibrium GdnHCl-Induced Unfolding of dcMN. An equilibrium GdnHCl-induced unfolding experiment yields the free energy of unfolding of N to unfolded chain A (A_U) and unfolded chain B (B_U). A two-state model can be used to describe the unfolding of N:



The equilibrium constant for unfolding, K_U is defined as

$$K_U = \frac{[\text{A}_U][\text{B}_U]}{[\text{N}]} \quad (6)$$

The total protein concentration, P_t , expressed as the total monomer concentration is given by

$$[\text{P}_t] = [\text{A}_U] + [\text{B}_U] + 2[\text{N}] \quad (7)$$

$$f_U = \frac{[\text{A}_U] + [\text{B}_U]}{[\text{P}_t]} = \frac{2[\text{A}_U]}{[\text{P}_t]} \quad (8)$$

$$f_N = \frac{2[\text{N}]}{[\text{P}_t]} \quad (9)$$

From eqs 6, 8, and 9

$$K_U = \frac{f_U^2[\text{P}_t]}{2(1 - f_U)} \quad (10)$$

Under standard conditions where N, A_U, and B_U are all present at concentrations of 1 M at 25 °C and pH 7, the free energy of unfolding is equal to the standard free energy. The standard free energy of unfolding is given by

$$\Delta G^\circ_U = -RT \ln K_U \quad (11)$$

Combining eqs 10 and 11

$$\frac{2e^{-\Delta G^\circ_U/RT}}{[P_t]} = \frac{f_U^2}{1-f_U} \quad (12)$$

When the free energy of unfolding is linearly dependent on denaturant concentration [D]

$$-RT \ln \frac{f_U^2 [P_t]}{2(1-f_U)} = \Delta G^\circ_U(\text{H}_2\text{O}) - m_U [D] \quad (13)$$

where $\Delta G^\circ_U(\text{H}_2\text{O})$ is the free energy of unfolding of N to A_U and B_U in water.

At the midpoint (C_m) of the unfolding transition, $f_U = 0.5$ when [D] = C_m. Hence, from eq 13

$$C_m = \frac{RT}{m_U} \ln \frac{[P_t]}{4} + \frac{\Delta G^\circ_U(\text{H}_2\text{O})}{m_U} \quad (14)$$

Substituting $2e^{-\Delta G^\circ_U/RT}/[P_t]$ with s in eq 12 and solving the resultant quadratic equation result in the following equation:

$$f_U = 0.5[-s \pm (s^2 + 4s)^{1/2}] \quad (15)$$

For a two-state transition, any spectroscopic signal Y is the weighted sum of the signals of native (Y_N) and unfolded (Y_U) chains. Hence

$$f_U = \frac{Y - Y_N}{Y_U - Y_N} \quad (16)$$

Equations 15 and 16 were used for analyzing all the equilibrium unfolding transitions of dcMN.

When the overall assembly of N from A_U and B_U can be represented as



the free energy of unfolding of dcMN is related to the free energy of dissociation by

$$\Delta G_U = \Delta G_d + \Delta G_U^A + \Delta G_U^B \quad (17)$$

where ΔG_U^A and ΔG_U^B are the free energies of unfolding of chains A and B, respectively.

pH Dependence of Stability. When a single titratable group in a protein has different pK_a values in the N and U forms (pK_a^N and pK_a^U, respectively), then the linkage between folding and proton binding yields the following expression for the dependence of the apparent free energy of unfolding ($\Delta G^{\circ,app}_U$) on pH:

$$\Delta G^{\circ,app}_U = \Delta G^\circ_U - 2.3RT \log \frac{1 + 10(pK_a^U - \text{pH})}{1 + 10(pK_a^N - \text{pH})} \quad (18)$$

where ΔG°_U is the free energy of unfolding of the protein with the titratable group deprotonated in both N and U states.

Calculation of FRET Efficiency. The efficiency, E , of FRET for any donor (D)–acceptor (A) pair is given by

$$E = 1 - \frac{F_{DA}}{F_D} \quad (19)$$

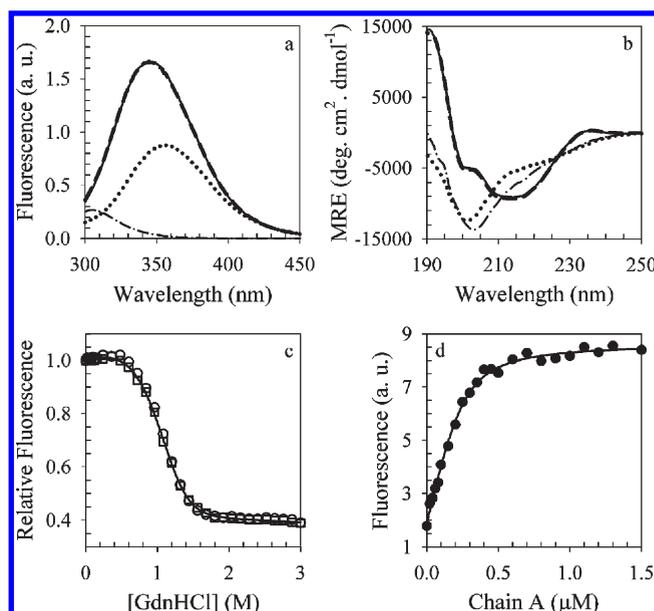


Figure 2. Complementations of chains A and B to form native dcMN at pH 7. (a) Fluorescence spectra of 10 μM native dcMN (—) and 10 μM reconstituted dcMN (---). Reconstituted dcMN was formed via incubation of a mixture of 10 μM chain A and 10 μM chain B together for 1 h (---). (b) Far-UV CD spectra of 10 μM native dcMN (—) and 10 μM reconstituted dcMN (---). Also shown in panels a and b are the fluorescence and far-UV CD spectra, respectively, of 10 μM chain A (— · —) and 10 μM chain B (· · · ·). (c) GdnHCl-induced equilibrium unfolding curves of native dcMN (○) and reconstituted dcMN (□). The solid line through the data was drawn by inspection only. (d) Binding of chain A to chain B. Chain B (0.25 μM) was incubated with the indicated concentrations of chain A for 3 h. Binding was monitored by measurement of the change in fluorescence at 340 nm. The solid line through the data is a nonlinear, least-squares fit to eq 5.

where F_D is the fluorescence intensity of the donor in the absence of the acceptor and is measured in the unlabeled protein and F_{DA} is the fluorescence intensity of the donor in the presence of the acceptor and is measured in the labeled protein.

RESULTS

Complementation of Chain A and Chain B To Form Native dcMN. Reconstitution of dcMN from its constituent chains A and B was conducted via incubation of equimolar concentrations of the two polypeptide chains for 1 h. Figure 2a shows the fluorescence spectra of native dcMN and reconstituted dcMN, obtained upon excitation at 280 nm. Figure 2b shows the far-UV CD spectra of native and reconstituted dcMN. The overlapping spectra in each case suggest that reconstituted dcMN obtained by chain complementation has the same structure as native dcMN. Figure 2c shows that the stabilities of native and reconstituted dcMN are the same: their GdnHCl-induced equilibrium unfolding curves are superimposable. Hence, reconstituted dcMN appears to have not only the same structure but also the same stability as native dcMN. Figure 2a also shows the fluorescence spectra of chain A and chain B. dcMN has one Trp residue and two Tyr residues located in chain B. Hence, chain B shows maximal emission at 350 nm, while chain A containing five Tyr residues and no Trp residues shows maximal fluorescence emission at 305 nm. The binding of chain A to chain B can be conveniently monitored at 340 nm.

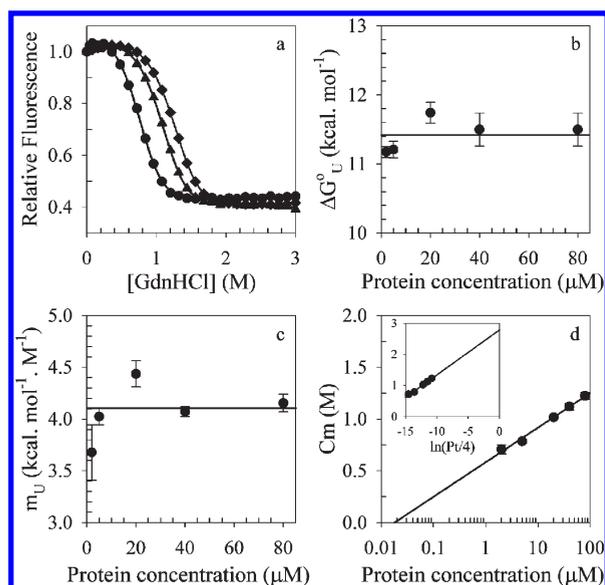


Figure 3. Dependence on protein concentration of the equilibrium unfolding transitions of dcMN at pH 7. (a) Representative GdnHCl-induced equilibrium unfolding transitions at protein concentrations $[[P_t]/2$ (see eq 7)] of 1 (●), 10 (▲) and 40 μM (◆). The solid lines through the data are nonlinear, least-squares fits to eq 15. The fits yielded values for $\Delta G^{\circ}_U(\text{H}_2\text{O})$ and m_U that are shown in panels b and c. (b) Dependence of $\Delta G^{\circ}_U(\text{H}_2\text{O})$ on protein concentration. (c) Dependence of m_U on protein concentration. In panels b and c, the lines have been drawn by inspection only. (d) Dependence of C_m on total monomer protein concentration ($[P_t]$). The inset shows $\ln([P_t]/4)$ vs C_m . The lines in panel d are least-squares fits to eq 14. In panels b–d, the error bars represent standard deviations determined from measurements taken in three separate experiments.

Figure 2d shows the fluorescence-monitored binding curve obtained via addition of varying concentrations of chain A to solutions containing a fixed (0.25 μM) concentration of chain B. The binding curve is hyperbolic, reaching a saturating fluorescence value. The data were analyzed on the basis of a simple two-state binding model (Scheme 1). The value obtained for the dissociation constant, K_d , is 0.04 μM. The standard free energy of dissociation of the two chains of dcMN is therefore 10 kcal mol⁻¹. It should be noted that because it was possible to titrate chain A against a very low concentration of chain B, the binding curve does not yield merely the 1:1 binding stoichiometry but a reliable value for the dissociation constant.

Determination of the Stability of dcMN Determined from Equilibrium Unfolding Studies at Different Protein Concentrations. GdnHCl-induced equilibrium unfolding studies were conducted at various protein concentrations ranging from 1 to 40 μM $[[P_t] = 2\text{--}80\ \mu\text{M}$ (see eq 7)]. Figure 3a shows the fluorescence-monitored GdnHCl-induced equilibrium unfolding curves obtained for dcMN at protein concentrations of 1, 10, and 40 μM. The GdnHCl-induced equilibrium unfolding curve for 10 μM protein was also monitored by far-UV CD (see below). The observation that the plots of f_U versus GdnHCl concentration, when determined by the two different probes, are coincident suggested that the unfolding of dcMN is a two-state process (Scheme 2). It was also found that the unfolding of dcMN was completely reversible: when completely unfolded dcMN was refolded by dilution of GdnHCl, the fluorescence and far-UV CD properties of native dcMN were fully recovered (data not shown). Hence, it became possible to conduct a thermodynamic

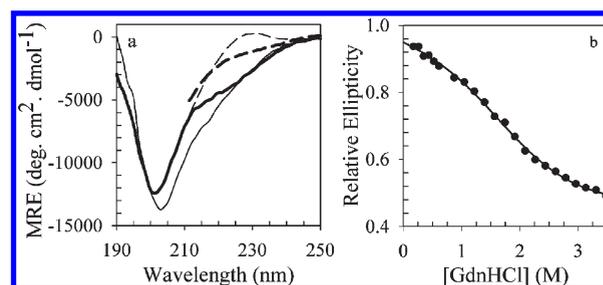


Figure 4. Characterization of structure in isolated chains A and B at pH 7. (a) Far-UV CD spectra of chain A (thin solid line), unfolded chain A in 3 M GdnHCl (thin dashed line), chain B (thick solid line), and unfolded chain B in 3 M GdnHCl (thick dashed line). All spectra were recorded at chain concentrations of 20 μM. (b) GdnHCl-induced equilibrium unfolding curve obtained for chain B. The transition was monitored by measurement of CD at 222 nm. The solid line through the data is a nonlinear least-squares fit to the equation for a two-state $N \leftrightarrow U$ transition,⁶⁵ using values for $\Delta G^{\circ,B}_U$ and m_U^B of 2 kcal mol⁻¹ and 1.1 kcal mol⁻¹ M⁻¹, respectively.

analysis of its stability. The values for $\Delta G^{\circ}_U(\text{H}_2\text{O})$ (Figure 3b) and m_U (Figure 3c) do not depend on protein concentration, and the average values of $\Delta G^{\circ}_U(\text{H}_2\text{O})$ and m_U are 11.4 ± 0.2 kcal mol⁻¹ and 4.2 ± 0.2 kcal mol⁻¹ M⁻¹, respectively. The value of C_m increases with an increase in protein concentration (Figure 3a), which is a characteristic feature of multimeric proteins. The dependence of C_m on protein concentration expressed as total monomer concentration, $[P_t]$ (eq 7) (Figure 3d), was analyzed according to eq 14. The intercept and slope of the plot of C_m versus $\ln([P_t]/4)$ (Figure 3d, inset) yield values for $\Delta G^{\circ}_U(\text{H}_2\text{O})$ and m_U of 11.1 kcal mol⁻¹ and 3.8 kcal mol⁻¹ M⁻¹, respectively. These values are in good agreement with the average $\Delta G^{\circ}_U(\text{H}_2\text{O})$ and m_U values determined from panels b and c of Figure 3, especially given the long extrapolation that is evident in the inset of Figure 3d.

Residual Structure Is Present in Chain B of dcMN. Figure 4a shows the far-UV CD spectra of individual chains A and B, in their native states as well as in the unfolded states. The spectra indicate that chain B possesses residual structure that disappears upon addition of 3 M GdnHCl. It should be noted that great care had to be taken during the isolation of chain B (see Materials and Methods) to obtain reproducible spectra in the native state. Figure 4b shows that the far-UV CD-monitored GdnHCl-induced equilibrium unfolding transition of chain B is broad but still appears sigmoidal. Two-state analysis of the transition yields a value of 2.0 kcal mol⁻¹ for $\Delta G^{\circ,B}_U(\text{H}_2\text{O})$. This value does not, however, appear to be reliable as other combinations of values for $\Delta G^{\circ,B}_U(\text{H}_2\text{O})$ and m_U^B also yielded equally satisfactory fits. Nevertheless, the data suggest that chain B does contain residual structure that is melted out upon addition of denaturant. It should also be noted that even though chain A displays significant ellipticity at ~220 nm, it was difficult to establish the presence of residual structure in chain A, because of its propensity to aggregate.

pH Dependence of the Equilibrium Unfolding Reactions of dcMN and scMN. To compare the relative stabilities of dcMN and scMN, equilibrium unfolding studies of dcMN and scMN were conducted over pH range of 4–9 and 4–10, respectively. Both fluorescence at 340 nm and far-UV CD at 222 nm were used as probes to determine whether unfolding is two-state over this pH range and then to determine the stabilities of the two proteins. Figure 5a shows that far-UV CD and fluorescence-

monitored equilibrium unfolding transitions of 10 μM dcMN are coincident at pH 4, 7, and 9. Figure 5b does likewise for 10 μM scMN at pH 4, 7, and 10. Previous studies of scMN had shown that the equilibrium unfolding transition was independent of protein concentration⁵⁸ as expected for a monomeric protein. Equilibrium unfolding studies were also conducted at pH <4 for both proteins, but it was not obvious that the transitions were two-state for either protein (data not shown).

Two-state analysis of the equilibrium unfolding transitions in the pH range of 4–10, of the single-chain and double-chain variants of monellin, as described in Materials and Methods, yielded values for $\Delta G^\circ_{\text{U}}(\text{H}_2\text{O})$, m_{U} , and C_{m} as a function of pH (Figure 6). Both dcMN and scMN are seen to have maximal stability at pH 4, which decreases with an increase in pH to pH 9. dcMN appears to be more stable than scMN, by a $\Delta\Delta G^\circ_{\text{U}}(\text{H}_2\text{O})$ of $\sim 5.5 \pm 0.5 \text{ kcal mol}^{-1}$, over the entire pH range (Figure 6a). The m_{U} values also appear to be independent of pH in the range of 4–10, with the values for scMN being significantly lower than those for dcMN (Figure 6b). At each pH value, the C_{m} value is higher for scMN than for dcMN, but it should be noted that the C_{m} for the dimeric protein increases with protein concentration (Figure 3).

The Spectroscopic Properties of dcMN and scMN Are Similar. It was important to determine whether the difference in the stabilities of dcMN and scMN is due to differences in their structures, in the folded or unfolded states. Figure 7 compares the

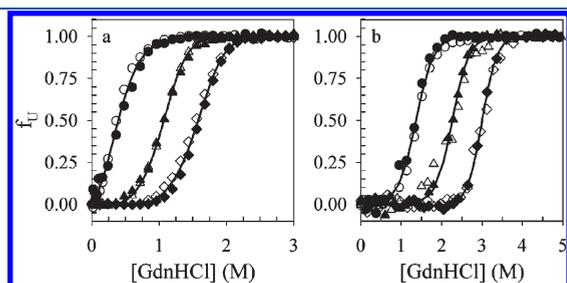


Figure 5. pH dependencies of the equilibrium unfolding transitions of dcMN and scMN. (a) Equilibrium unfolding of dcMN. The GdnHCl-induced transitions were monitored by measurement of the fluorescence at 340 nm (filled symbols) and far-UV CD at 222 nm (empty symbols) at pH 9 (●), 7 (▲), and 4 (◆). The displayed plots of f_{U} vs GdnHCl concentration were obtained from the raw data using eq 16. The solid lines through the data are nonlinear least-squares fits to eq 15. (b) Equilibrium unfolding of scMN at pH 10 (●), 7 (▲), and 4 (◆). The solid lines through the data are nonlinear least-squares fits to the equation for a two-state $\text{N} \leftrightarrow \text{U}$ transition.⁶⁵ The values obtained for $\Delta G^\circ_{\text{U}}(\text{H}_2\text{O})$ and m_{U} are shown in Figure 6.

far-UV CD (Figure 7a), near-UV CD (Figure 7b), and fluorescence (Figure 7c) spectra of the two proteins in their folded and unfolded states. For the native states of the two proteins, all three spectra are very similar for both proteins. For the unfolded states, the far-UV CD and fluorescence spectra are very similar, but the near-UV CD spectra, while similar in shape, appear to report reproducibly different values for the mean residue ellipticity (MRE). The difference in the near-UV CD spectra of the unfolded proteins is indeed very small but was seen consistently in multiple acquisitions of the spectra. From the spectroscopic studies, it therefore appears that the native states of dcMN and scMN are very similar in structure, while the unfolded states might be marginally different in conformation.

Measurement of an Intramolecular Distance in dcMN and scMN. To determine whether there is indeed a difference in the unfolded state conformations of dcMN and scMN, a single intramolecular distance between Trp4 and Cys42 was measured in both proteins by fluorescence resonance energy transfer (FRET) between Trp4 (the FRET donor) and a TNB adduct (the FRET acceptor) covalently attached to Cys42. Both Trp4 and Cys42 are located in chain B of dcMN, and the goal was to measure whether there is a difference in the efficiency in FRET from Trp4 to Cys42TNB as a consequence of the distance separating the donor and acceptor being different in the two proteins. In previous work with scMN, FRET had been shown to be powerful tool for measuring conformational change during unfolding, and the utility of monellin for FRET measurements has been established.^{11,60} Figure 8 shows the fluorescence spectra of the single-chain and double-chain variants of monellin, in the folded and unfolded states. It is seen that the FRET efficiency¹¹ is identical for the native states of the two proteins, again indicating that the native states have very similar structures. The FRET efficiencies are different for the unfolded states of the two proteins. The difference is small (see the legend of Figure 8) but was consistently obtained and indicates that free unfolded chain B (in unfolded dcMN) may have a conformation different from that of the chain B segment of unfolded scMN.

dcMN and scMN Unfold at Similar Rates. Both scMN⁵⁸ and dcMN (data not shown) unfold with single-exponential kinetics, when unfolding is monitored by measurement of fluorescence at 340 nm. Figure 9 shows that the apparent rate constants of unfolding in the range of GdnHCl concentrations from 2 to 6 M are very similar for the two proteins. This result indicates that even though the two proteins have very different stabilities, the free energies of activation for unfolding are very similar in value in the two cases. Kinetic stabilities are similar even though

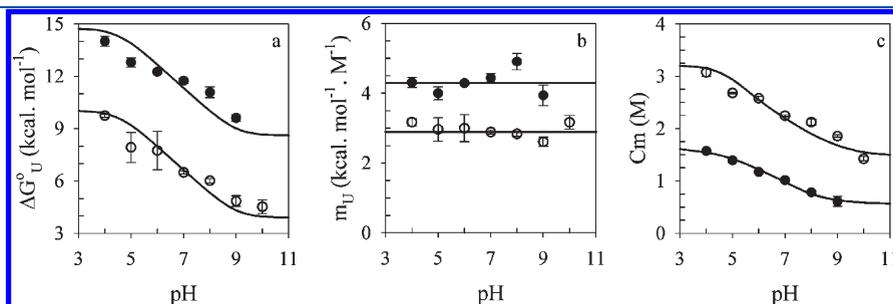


Figure 6. pH dependence of the thermodynamic parameters governing equilibrium unfolding of (●) dcMN and (○) scMN. Panels a–c show the pH dependencies of $\Delta G^\circ_{\text{U}}(\text{H}_2\text{O})$, m_{U} , and C_{m} , respectively. The solid lines through the data in panel a are described by eq 18, with values for pK_a^{U} and pK_a^{N} of 4.5 and 9, respectively. The lines through the data in panels b and c have been drawn by inspection only. The error bars represent the standard errors of measurement obtained from two separate experiments.

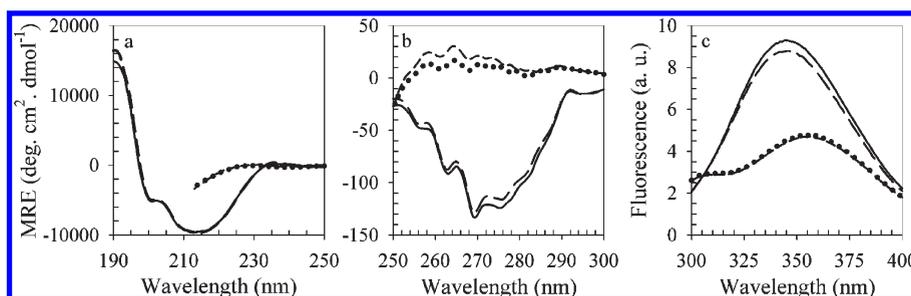


Figure 7. Comparison of the spectroscopic properties of dcMN and scMN at pH 7. (a) Far-UV CD spectra. (b) Near-UV CD spectra. (c) Fluorescence spectra. The solid and long-dash lines represent the native state spectra of dcMN and scMN, respectively, while the short-dash line and the dotted line represent the spectra of unfolded dcMN and scMN, respectively, in 4 M GdnHCl.

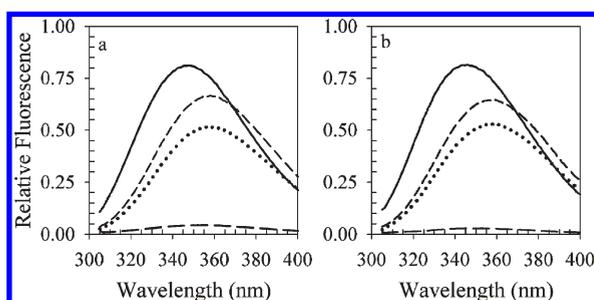


Figure 8. Fluorescence spectra of TNB-labeled and unlabeled proteins at pH 7. Panels a and b show the spectra for dcMN and wild-type scMN, respectively. dcMN was labeled at the sole cysteine residue (C42) in chain B, and scMN was also labeled at the corresponding cysteine residue (C42). The solid and long-dash lines represent the fluorescence spectra of the unlabeled and labeled native proteins, respectively, while the short-dash line and the dotted line represent the fluorescence spectra of unfolded proteins in 4 M GdnHCl. All spectra were recorded at a protein concentration of 10 μ M. The FRET efficiencies were calculated from the fluorescence spectra, using eq 19; they are 0.95 for native dcMN, 0.23 for unfolded dcMN, 0.96 for native scMN, and 0.18 for unfolded scMN.

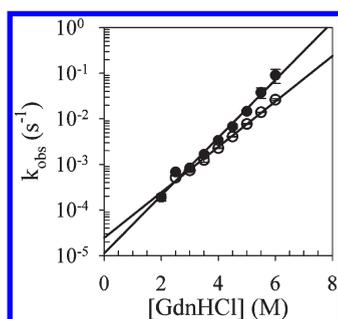


Figure 9. Unfolding kinetics of dcMN and scMN at pH 7. The apparent rate constants of unfolding of dcMN (●) and scMN (○) were obtained by fitting the fluorescence-monitored kinetic traces of unfolding in different concentrations of GdnHCl to a single-exponential equation. The error bars reflect the spreads in measurements from two separate experiments.

thermodynamic stabilities are very different. It is unlikely that this result would have been observed if the two native states differed significantly in their structures and stabilizing free energies.

DISCUSSION

Apparent Two-State Unfolding of Monellin. Earlier equilibrium unfolding studies^{11,58} had shown that scMN undergoes

apparent two-state unfolding at pH 7 and 8. In this study, by the criterion that the equilibrium unfolding curves monitored by fluorescence and CD are coincident (Figure 5), it is seen that the equilibrium unfolding of scMN can be described as two-state in the pH range of 4–10. This study also shows, using the same criterion of coincident fluorescence and CD-monitored equilibrium unfolding curves (Figure 5), that equilibrium unfolding of dcMN can be approximated as two-state unfolding in the pH range of 4–9.

Several other dimeric proteins also appear to display two-state unfolding and folding behavior in equilibrium studies.^{13,28,32} The folding of a dimeric protein involves both binding and structure formation. The probability of both events happening simultaneously is negligible, given that both are stochastic events. It is possible that the binding event is silent to and, hence, cannot be detected by spectroscopic change, and that the entire fluorescence and CD change occurs only during the structure formation (folding) step. In that case, the structure formation step would appear to be a two-state step. It is more likely that unfolding or folding occurs in multiple steps but that partially folded intermediates are unstable and, hence, populated too sparsely at equilibrium to be detected for some dimeric proteins. Partially folded intermediates can be seen to accumulate transiently during the folding of dimeric proteins, whose folding and unfolding appear to be two-state steps in equilibrium studies.^{66,67}

Coupling of Binding and Folding. Denaturant-induced equilibrium unfolding studies measure the free energy difference between the native and unfolded states of a protein under refolding conditions. Thus, for dcMN, the measured free energy difference is between state N and states A_U and B_U (Scheme 2). If chains A and B were to exist in the completely unfolded conformations, A_U and B_U, respectively, when free in solution, then the free energy of unfolding of dcMN, ΔG°_U , would be the same as the free energy of dissociation of dcMN into chains A and B, ΔG°_d . The value of ΔG°_d is 10 kcal mol⁻¹, similar to that reported previously.⁶¹ This is significantly lower than the value of 11.4 kcal mol⁻¹ obtained for ΔG°_U . The difference of 1.4 kcal mol⁻¹ would represent the free energy of unfolding of residual structure present in free chains A and B (eq 17).

Chain B does indeed appear to possess residual structure that is lost upon addition of GdnHCl (Figure 4). It is less clear whether chain A also possesses residual structure. The equilibrium unfolding transition of chain B has a broad but sigmoidal dependence on GdnHCl concentration. Although it can be fit well to a two-state B \leftrightarrow B_U model for unfolding, the value of \sim 2 kcal mol⁻¹ obtained for $\Delta G^{\circ,B}_U$, the free energy of unfolding of free chain B, is not very robust: other values also fit the data

equally well. Nevertheless, the data in Figure 4 strongly suggest that the origin of the difference observed in $\Delta G^{\circ}_{\text{U}}$ and $\Delta G^{\circ}_{\text{d}}$ must lie in the residual structure present in free chain B. Hence, the overall assembly of unfolded chains A and B into native protein can be represented by Scheme 3 (see Materials and Methods). In Scheme 3, A and B represent monomeric intermediates in the overall assembly–folding reaction starting from A_{U} and B_{U} , respectively. In the equilibrium unfolding experiments, A and B are populated too sparsely to be detected; hence, the unfolding of dcMN appears to be a two-state process. It is expected that kinetic studies, now in progress, will be able to determine whether A and B do indeed accumulate transiently during the folding of A_{U} and B_{U} to dcMN.

Monomeric intermediates in the folding of a dimeric protein are likely to be observable only at a low protein concentration in the range of the equilibrium constant for the dissociation of the dimeric protein. Not surprisingly then, instances of monomeric intermediates being reported to accumulate in the assembly of a dimeric protein are rare.³⁰ There are more instances reported of dimeric intermediates populated during the assembly of multimeric proteins^{29,32,68} because they are populated at higher protein concentrations where equilibrium unfolding studies are amenable. When intermediates in the assembly of multimeric proteins cannot be directly detected, their presence may manifest itself in measured $\Delta G^{\circ}_{\text{U}}$ and m_{U} values being dependent on protein concentration.^{24,30} In the case of dcMN, the observation (Figure 5) that $\Delta G^{\circ}_{\text{U}}$ and m_{U} do not change with a change in protein concentration suggests that any unfolding intermediate present at equilibrium is so sparsely populated that a single $\Delta G^{\circ}_{\text{U}}$ and m_{U} value can adequately account for the dependence of the midpoint of the unfolding transition on protein concentration (Figure 3d).

pH Dependence of the Unfolding of dcMN and scMN. A difference in the pK_{a} values for an ionizable group in the folded and unfolded states of a protein can contribute to the dependence of its stability on pH.⁶⁹ In the case of monellin, the observed decrease in stability with an increase in pH from 4 to 10 (Figure 6) can be ascribed to a linkage of folding to the protonation of a group whose pK_{a} increases from ~ 4.5 in U to ~ 9 in N. Because there is no His residue in monellin, the ionizable group is likely to be the carboxylate side chain of a Glu or Asp residue, whose pK_{a} is expected to be ~ 4.5 in U, and whose pK_{a} can be expected to be high if it is either buried in a nonpolar environment or near a negative charge in N. In monellin, the side chains of Glu26 and Asp27 in chain A and the side chain of Glu23 in chain B are buried in N: they expose $\leq 10\%$ of their solvent-accessible surface areas in the native state of dcMN as well as that of scMN. Of these three residues, only the side chain of Glu23 in chain B does not have the positively charged side chain of a Lys or Arg residue in its immediate proximity that could neutralize its charge, nor does it appear to be hydrogen-bonded to any other residue. Hence, it is likely that it is the side chain of Glu23 that has an abnormally high pK_{a} in N; consequently, its side chain does not deprotonate and become charged in the pH range of 4–10. Large increases^{70–72} as well as decreases⁷³ in the pK_{a} values of buried ionizable side chains have been reported to account for the pH dependencies of the stabilities of several other proteins, and factors determining the pK_{a} values of buried ionizable groups have been described previously.⁷⁴

Comparison of the Stabilities of dcMN and scMN. It is of interest to compare the stabilities of dcMN and scMN. The value of $\Delta G^{\circ}_{\text{U}}$ for dcMN is ~ 5.5 kcal mol⁻¹ larger than that for scMN over the entire pH range from 4 to 9. Both scMN and dcMN have

virtually identical structures; hence, the enthalpic contributions to stability should be very similar. The entropic contributions should, however, be very different. In the case of dcMN, the standard free energy of unfolding, $\Delta G^{\circ}_{\text{U}}$, is the sum of the intrinsic stability, $\Delta G^{\circ}_{\text{i}}$, which is generally positive, and an entropic free energy, $\Delta G^{\circ}_{\text{s}}$, which is generally negative because of the gain in translational and rotational entropy as a result of dissociation.⁷⁵ The gain in translational entropy upon dissociation has been estimated as $R \ln 55.5$;^{4,76,77} hence, this cratic entropic contribution of $RT \ln 55.5$ to $\Delta G^{\circ}_{\text{s}}$ in the case of the unfolding of dcMN is calculated to be 2.4 kcal mol⁻¹ at 25 °C. Thus, in the case of monellin, even the intrinsic stability of dcMN is ~ 3.1 kcal mol⁻¹ greater than that of scMN. This result suggests that when stability is measured as $\Delta G^{\circ}_{\text{U}}$ at a standard concentration of 1 M, which is usually done, it is not possible to exclude the possibility that single-chain variants of other proteins may also be less intrinsically stable than their multichain counterparts.

Nevertheless, this possibility rarely appears to manifest itself when stability is measured as $\Delta G^{\circ}_{\text{U}}$ at the standard concentration of 1 M. For example, the artificially created single-chain variants of gene V protein, bacteriophage f1⁴⁴ and GroES¹⁵ are more stable than their naturally occurring multimeric forms, the disulfide-bridged coil–coil domain of laminin is more stable than its unbridged counterpart,⁷⁸ and ribonuclease A is more stable than its artificially created two-chain variant, ribonuclease S.^{79,80} In the case of thioredoxin, while one artificially created two-chain variant is nearly as stable as the naturally occurring single-chain variant,⁴³ another is less stable.⁸¹ Clearly, for a multimeric protein whose chains are connected covalently, the decrease in entropy when the unfolded chains of the multichain variant are joined together in a single unfolded chain of the single chain variant usually leads to $\Delta G^{\circ}_{\text{U}}$ for the multichain variant being smaller than that for the single-chain variant. It is therefore surprising that the unfolding of dcMN is seen to be accompanied by a larger standard free energy change than the unfolding of scMN.

One reason why $\Delta G^{\circ}_{\text{U}}$ for scMN is less than that for dcMN could be that the native state of scMN is strained, and the strain in the native state^{82–84} is released when there is a break in the chain as in dcMN. The increased entropy in the native state of dcMN would lower its free energy relative to that of scMN. The structure of scMN indicates that at least one residue in the region of the linker joining chain A to chain B is constrained (Figure 1). The observations that the observed unfolding rate constants of dcMN and scMN are similar and that their dependencies on denaturant concentration are not very different suggest that the greater stability of dcMN is not derived from a slower unfolding rate constant. The rate constants of unfolding are expected to be similar when the specific interactions maintaining the native state structure are identical in dcMN and scMN. Hence, the structure of scMN is unlikely to be strained in a manner in which the structure of dcMN is not. Alternatively, there could be strain in scMN, and that strain could be relieved before the rate-limiting step of unfolding. It is known that the unfolding of scMN begins with the formation of a dry molten globule¹¹ in which packing interactions become looser, and it is possible that strain in the structure of scMN is absent in the dry molten globule.

Another reason could be that the unfolded state of scMN is stabilized with respect to that of dcMN because of the presence of residual stabilizing structure in the former but not in the latter. For example, stabilizing structure in the unfolded form of barstar has been shown to reduce its stability.³⁸ There is some evidence

of residual structure in the unfolded state of scMN. One indicator is that the value of m_U is significantly larger for dcMN than for scMN (Figure 5), indicating that much more surface area becomes solvent-accessible when dcMN unfolds than when scMN unfolds. Another indicator is that the near-UV CD spectrum of unfolded dcMN is different from that of scMN, albeit to a small extent (Figure 7). A third pointer is that several intramolecular distances in scMN are shorter than the distances expected for a random coil.¹¹ In this study, it is seen that an intramolecular distance measured by FRET is different in B_U (from dcMN) and U (from scMN) (Figure 8), albeit to a small extent. Although the data indicate that the distance is shorter in B_U than in U , it also suggests that the conformation of the B_U is different from that of the B chain segment in U . At present, it is not possible to determine why the unfolding of dcMN is accompanied by a larger free energy change than the unfolding of scMN, but it appears that monellin may have evolved to exist in a two-chain form in nature because the two-chain form is more stable than the one-chain form.

Finally, it is important to understand the meaning of the difference in the standard free energies of unfolding of dcMN and scMN. At pH 7, the value of ΔG°_U for scMN (6.2 kcal mol⁻¹) indicates that the value of f_U is 3×10^{-5} . For a dimeric protein, the value of f_U depends on the value of $[P_t]$ (eq 12). For dcMN at pH 7, the value of ΔG°_U (11.4 kcal mol⁻¹) indicates that the value of f_U is 0.07 when $[P_t] = 1 \mu\text{M}$, 0.02 when $[P_t] = 10 \mu\text{M}$, and 0.008 when $[P_t] = 100 \mu\text{M}$. The value of f_U for dcMN becomes smaller than that for scMN only for protein concentrations in the molar range. It is important to remember that the stabilities that are being compared are standard free energies that are determined under the arbitrarily chosen standard conditions when reactants and products are all present at concentrations of 1 M at 25 °C and pH 7 (see Data Analysis). In this context, two points need to be considered when comparing the stabilities of the single-chain and two-chain variants of the same protein. (1) The free energy of unfolding of the double-chain protein decreases with an increase in concentration. The stability of dcMN becomes lower than that of scMN at micromolar protein concentrations, when $[P_t] < 100 \mu\text{M}$. (2) The value of the free energy of unfolding of a double-chain protein depends on the choice of the standard concentration.⁸⁵ For example, if 1 μM had been chosen as the standard concentration instead of the conventional value of 1 M, then the standard free energy of unfolding of double-chain monellin would have been found to be lower than that of its single-chain counterpart. Clearly, comparing the stabilities of the two-chain and single-chain variants of a protein is not straightforward.

AUTHOR INFORMATION

Corresponding Author

*Fax: 91-80-23636662. Phone: 91-080-23666150. E-mail: jayant@ncbs.res.in.

Funding Sources

This work was funded by the Tata Institute of Fundamental Research and by the Department of Biotechnology, Government of India. J.B.U. is the recipient of a JC Bose National Fellowship from the Government of India.

ACKNOWLEDGMENT

We thank Ashish Patra for his help and discussion and Deepak Nair, Shachi Gosavi, and Raghavan Varadarajan for discussions.

Mass spectra were recorded at the Mass Spectrometry Facility of the National Centre for Biological Sciences.

ABBREVIATIONS

dcMN, double-chain monellin; scMN, single-chain monellin; CD, circular dichroism; RPC, reverse phase chromatography; FRET, fluorescence resonance energy transfer; GdnHCl, guanidine hydrochloride; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol.

REFERENCES

- (1) Shoichet, B. K., Baase, W. A., Kuroki, R., and Matthews, B. W. (1995) A relationship between protein stability and protein function. *Proc. Natl. Acad. Sci. U.S.A.* 92, 452–456.
- (2) Greene, L. H., Grobler, J. A., Malinovsky, V. A., Tian, J., Acharya, K. R., and Brew, K. (1999) Stability, activity and flexibility in α -lactalbumin. *Protein Eng.* 12, 581–587.
- (3) Lazaridis, T., and Karplus, M. (2003) Thermodynamics of protein folding: A microscopic view. *Biophys. Chem.* 100, 367–395.
- (4) Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* 14, 1–63.
- (5) Tanford, C. (1962) Contribution of Hydrophobic Interactions to the Stability of the Globular Conformation of Proteins. *J. Am. Chem. Soc.* 84, 4240–4247.
- (6) Dill, K. A. (1990) Dominant forces in protein folding. *Biochemistry* 29, 7133–7155.
- (7) Privalov, P. L. (1979) Stability of proteins: Small globular proteins. *Adv. Protein Chem.* 33, 167–241.
- (8) Brady, G. P., and Sharp, K. A. (1997) Entropy in protein folding and in protein-protein interactions. *Curr. Opin. Struct. Biol.* 7, 215–221.
- (9) Pace, C. N. (2001) Polar group burial contributes more to protein stability than nonpolar group burial. *Biochemistry* 40, 310–313.
- (10) Bolen, D. W., and Rose, G. D. (2008) Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annu. Rev. Biochem.* 77, 339–362.
- (11) Jha, S. K., and Udgaonkar, J. B. (2009) Direct evidence for a dry molten globule intermediate during the unfolding of a small protein. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12289–12294.
- (12) Baldwin, R. L., Frieden, C., and Rose, G. D. (2010) Dry molten globule intermediates and the mechanism of protein unfolding. *Proteins* 78, 2725–2737.
- (13) Bowie, J. U., and Sauer, R. T. (1989) Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry* 28, 7139–7143.
- (14) Mei, G., Di Venere, A., Rosato, N., and Finazzi-Agro, A. (2005) The importance of being dimeric. *FEBS J.* 272, 16–27.
- (15) Sakane, I., Hongo, K., Motojima, F., Murayama, S., Mizobata, T., and Kawata, Y. (2007) Structural stability of covalently linked GroES heptamer: Advantages in the formation of oligomeric structure. *J. Mol. Biol.* 367, 1171–1185.
- (16) Baez, M., and Babul, J. (2009) Reversible unfolding of dimeric phosphofructokinase-2 from *Escherichia coli* reveals a dominant role of inter-subunit contacts for stability. *FEBS Lett.* 583, 2054–2060.
- (17) Schachman, H. K. (1988) Can a simple model account for the allosteric transition of aspartate transcarbamoylase?. *J. Biol. Chem.* 263, 18583–18586.
- (18) Neet, K. E. (1995) Cooperativity in enzyme function: Equilibrium and kinetic aspects. *Methods Enzymol.* 249, 519–567.
- (19) Perham, R. N. (2000) Swinging arms and swinging domains in multifunctional enzymes: Catalytic machines for multistep reactions. *Annu. Rev. Biochem.* 69, 961–1004.
- (20) Jaenicke, R. (1987) Folding and association of proteins. *Prog. Biophys. Mol. Biol.* 49, 117–237.
- (21) Neet, K. E., and Timm, D. E. (1994) Conformational stability of dimeric proteins: Quantitative studies by equilibrium denaturation. *Protein Sci.* 3, 2167–2174.

- (22) Ragone, R. (2000) How the protein concentration affects unfolding curves of oligomers. *Biopolymers* 53, 221–225.
- (23) Park, C., and Marqusee, S. (2004) Analysis of the stability of multimeric proteins by effective ΔG and effective m -values. *Protein Sci.* 13, 2553–2558.
- (24) Rumpf, J. A., Galvagnion, C., Vassall, K. A., and Meiring, E. M. (2008) Conformational stability and folding mechanisms of dimeric proteins. *Prog. Biophys. Mol. Biol.* 98, 61–84.
- (25) Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447, 1021–1025.
- (26) Espinoza-Fonseca, L. M. (2009) Reconciling binding mechanisms of intrinsically disordered proteins. *Biochem. Biophys. Res. Commun.* 382, 479–482.
- (27) Welker, S., Rudolph, B., Frenzel, E., Hagn, F., Liebisch, G., Schmitz, G., Scheuring, J., Kerth, A., Blume, A., Weinkauff, S., Haslbeck, M., Kessler, H., and Buchner, J. (2010) Hsp12 is an intrinsically unstructured stress protein that folds upon membrane association and modulates membrane function. *Mol. Cell* 39, 507–520.
- (28) Thompson, K. S., Vinson, C. R., and Freire, E. (1993) Thermodynamic characterization of the structural stability of the coiled-coil region of the bZIP transcription factor GCN4. *Biochemistry* 32, 5491–5496.
- (29) Gloss, L. M., and Matthews, C. R. (1997) Urea and thermal equilibrium denaturation studies on the dimerization domain of *Escherichia coli* Trp repressor. *Biochemistry* 36, 5612–5623.
- (30) Mallam, A. L., and Jackson, S. E. (2005) Folding studies on a knotted protein. *J. Mol. Biol.* 346, 1409–1421.
- (31) Jelesarov, I., and Bosshard, H. R. (1996) Thermodynamic characterization of the coupled folding and association of heterodimeric coiled coils (leucine zippers). *J. Mol. Biol.* 263, 344–358.
- (32) Banks, D. D., and Gloss, L. M. (2003) Equilibrium folding of the core histones: The H3-H4 tetramer is less stable than the H2A-H2B dimer. *Biochemistry* 42, 6827–6839.
- (33) Bachhawat, K., Kapoor, M., Dam, T. K., and Surolia, A. (2001) The reversible two-state unfolding of a monocoat mannose-binding lectin from garlic bulbs reveals the dominant role of the dimeric interface in its stabilization. *Biochemistry* 40, 7291–7300.
- (34) Panse, V. G., Swaminathan, C. P., Aloor, J. J., Surolia, A., and Varadarajan, R. (2000) Unfolding thermodynamics of the tetrameric chaperone, SecB. *Biochemistry* 39, 2362–2369.
- (35) Boudker, O., Todd, M. J., and Freire, E. (1997) The structural stability of the co-chaperonin GroES. *J. Mol. Biol.* 272, 770–779.
- (36) Luke, K., Apiyo, D., and Wittung-Stafshede, P. (2005) Dissecting homo-heptamer thermodynamics by isothermal titration calorimetry: Entropy-driven assembly of co-chaperonin protein 10. *Biophys. J.* 89, 3332–3336.
- (37) Zhou, H. X. (2001) Single-chain versus dimeric protein folding: Thermodynamic and kinetic consequences of covalent linkage. *J. Am. Chem. Soc.* 123, 6730–6731.
- (38) Pradeep, L., and Udgaonkar, J. B. (2004) Effect of salt on the urea-unfolded form of barstar probed by m value measurements. *Biochemistry* 43, 11393–11402.
- (39) Richards, F. M., and Vithayathil, P. J. (1959) The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein components. *J. Biol. Chem.* 234, 1459–1465.
- (40) Hartley, R. W. (1977) Complementation of peptides of barnase, extracellular ribonuclease of *Bacillus amyloliquefaciens*. *J. Biol. Chem.* 252, 3252–3254.
- (41) Parr, G. R., and Taniuchi, H. (1981) Ordered complexes of cytochrome *c* fragments. Kinetics of formation of the reduced (ferrous) forms. *J. Biol. Chem.* 256, 125–132.
- (42) de Prat Gay, G., and Fersht, A. R. (1994) Generation of a family of protein fragments for structure-folding studies. I. Folding complementation of two fragments of chymotrypsin inhibitor-2 formed by cleavage at its unique methionine residue. *Biochemistry* 33, 7957–7963.
- (43) Georgescu, R. E., Braswell, E. H., Zhu, D., and Tasayco, M. L. (1999) Energetics of assembling an artificial heterodimer with an α/β motif: Cleaved versus uncleaved *Escherichia coli* thioredoxin. *Biochemistry* 38, 13355–13366.
- (44) Liang, H., Sandberg, W. S., and Terwilliger, T. C. (1993) Genetic fusion of subunits of a dimeric protein substantially enhances its stability and rate of folding. *Proc. Natl. Acad. Sci. U.S.A.* 90, 7010–7014.
- (45) Robinson, C. R., and Sauer, R. T. (1996) Equilibrium stability and sub-millisecond refolding of a designed single-chain Arc repressor. *Biochemistry* 35, 13878–13884.
- (46) Morris, J. A., and Cagan, R. H. (1972) Purification of monellin, the sweet principle of *Dioscoreophyllum cumminsii*. *Biochim. Biophys. Acta* 261, 114–122.
- (47) Ogata, C., Hatada, M., Tomlinson, G., Shin, W. C., and Kim, S. H. (1987) Crystal structure of the intensely sweet protein monellin. *Nature* 328, 739–742.
- (48) Somoza, J. R., Jiang, F., Tong, L., Kang, C. H., Cho, J. M., and Kim, S. H. (1993) Two crystal structures of a potentially sweet protein. Natural monellin at 2.75 Å resolution and single-chain monellin at 1.7 Å resolution. *J. Mol. Biol.* 234, 390–404.
- (49) Xue, W. F., Szczepankiewicz, O., Bauer, M. C., Thulin, E., and Linse, S. (2006) Intra- versus intermolecular interactions in monellin: Contribution of surface charges to protein assembly. *J. Mol. Biol.* 358, 1244–1255.
- (50) Morris, J. A., Martenson, R., Deibler, G., and Cagan, R. H. (1973) Characterization of monellin, a protein that tastes sweet. *J. Biol. Chem.* 248, 534–539.
- (51) Morris, J. A., and Cagan, R. H. (1975) Effects of denaturants on the sweet-tasting protein monellin. *Proc. Soc. Exp. Biol. Med.* 150, 265–270.
- (52) Morris, J. A., and Cagan, R. H. (1980) Formation of oligomeric monellin in protein denaturants. *Proc. Soc. Exp. Biol. Med.* 164, 351–354.
- (53) Brand, J. G., and Cagan, R. H. (1977) Fluorescence characteristics of native and denatured monellin. *Biochim. Biophys. Acta* 493, 178–187.
- (54) Brand, J. G., Cagan, R. H., and Bayley, D. L. (1985) Conformational changes of the sweet protein monellin as measured by fluorescence emission. *Proc. Soc. Exp. Biol. Med.* 179, 76–82.
- (55) Jirgensons, B. (1976) Conformational transitions of monellin, an intensely sweet protein. *Biochim. Biophys. Acta* 446, 255–261.
- (56) Kimura, T., Uzawa, T., Ishimori, K., Morishima, I., Takahashi, S., Konno, T., Akiyama, S., and Fujisawa, T. (2005) Specific collapse followed by slow hydrogen-bond formation of β -sheet in the folding of single-chain monellin. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2748–2753.
- (57) Kimura, T., Maeda, A., Nishiguchi, S., Ishimori, K., Morishima, I., Konno, T., Goto, Y., and Takahashi, S. (2008) Dehydration of main-chain amides in the final folding step of single-chain monellin revealed by time-resolved infrared spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13391–13396.
- (58) Patra, A. K., and Udgaonkar, J. B. (2007) Characterization of the folding and unfolding reactions of single-chain monellin: Evidence for multiple intermediates and competing pathways. *Biochemistry* 46, 11727–11743.
- (59) Patra, A. K., and Udgaonkar, J. B. (2009) GroEL can unfold late intermediates populated on the folding pathways of monellin. *J. Mol. Biol.* 389, 759–775.
- (60) Jha, S. K., Dhar, D., Krishnamoorthy, G., and Udgaonkar, J. B. (2009) Continuous dissolution of structure during the unfolding of a small protein. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11113–11118.
- (61) Xue, W. F., Carey, J., and Linse, S. (2004) Multi-method global analysis of thermodynamics and kinetics in reconstitution of monellin. *Proteins* 57, 586–595.
- (62) Kohmura, M., Nio, N., and Ariyoshi, Y. (1990) Complete amino acid sequence of the sweet protein monellin. *Agric. Biol. Chem.* 54, 2219–2224.
- (63) Aghera, N., and Udgaonkar, J. B. (2011) Heterologous expression, purification and characterization of heterodimeric monellin. *Protein Expression Purif.* 76, 248–253.
- (64) Ramachandran, S., and Udgaonkar, J. B. (1996) Stabilization of barstar by chemical modification of the buried cysteines. *Biochemistry* 35, 8776–8785.

- (65) Agashe, V. R., and Udgaonkar, J. B. (1995) Thermodynamics of denaturation of barstar: Evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride. *Biochemistry* 34, 3286–3299.
- (66) Zitzewitz, J. A., Ibarra-Molero, B., Fishel, D. R., Terry, K. L., and Matthews, C. R. (2000) Preformed secondary structure drives the association reaction of GCN4-p1, a model coiled-coil system. *J. Mol. Biol.* 296, 1105–1116.
- (67) Placek, B. J., and Gloss, L. M. (2005) Three-state kinetic folding mechanism of the H2A/H2B histone heterodimer: The N-terminal tails affect the transition state between a dimeric intermediate and the native dimer. *J. Mol. Biol.* 345, 827–836.
- (68) Shaw, G. S., Marlatt, N. M., Ferguson, P. L., Barber, K. R., and Bottomley, S. P. (2008) Identification of a dimeric intermediate in the unfolding pathway for the calcium-binding protein S100B. *J. Mol. Biol.* 382, 1075–1088.
- (69) Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.* 24, 1–95.
- (70) Langsetmo, K., Fuchs, J. A., and Woodward, C. (1991) The conserved, buried aspartic acid in oxidized *Escherichia coli* thioredoxin has a pKa of 7.5. Its titration produces a related shift in global stability. *Biochemistry* 30, 7603–7609.
- (71) Dwyer, J. J., Gittis, A. G., Karp, D. A., Lattman, E. E., Spencer, D. S., Stites, W. E., and Garcia-Moreno, E. B. (2000) High apparent dielectric constants in the interior of a protein reflect water penetration. *Biophys. J.* 79, 1610–1620.
- (72) Grimsley, G. R., Scholtz, J. M., and Pace, C. N. (2009) A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci.* 18, 247–251.
- (73) Giletto, A., and Pace, C. N. (1999) Buried, charged, non-ion-paired aspartic acid 76 contributes favorably to the conformational stability of ribonuclease T1. *Biochemistry* 38, 13379–13384.
- (74) Harms, M. J., Castaneda, C. A., Schlessman, J. L., Sue, G. R., Isom, D. G., Cannon, B. R., and Garcia-Moreno, E. B. (2009) The pK_a values of acidic and basic residues buried at the same internal location in a protein are governed by different factors. *J. Mol. Biol.* 389, 34–47.
- (75) Jencks, W. P. (1981) On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U.S.A.* 78, 4046–4050.
- (76) Luque, I., and Freire, E. (1998) Structure-based prediction of binding affinities and molecular design of peptide ligands. *Methods Enzymol.* 295, 100–127.
- (77) Dell’Orco, D., Xue, W. F., Thulin, E., and Linse, S. (2005) Electrostatic contributions to the kinetics and thermodynamics of protein assembly. *Biophys. J.* 88, 1991–2002.
- (78) Antonsson, P., Kammerer, R. A., Schulthess, T., Hanisch, G., and Engel, J. (1995) Stabilization of the α -helical coiled-coil domain in laminin by C-terminal disulfide bonds. *J. Mol. Biol.* 250, 74–79.
- (79) Catanzano, F., Giancola, C., Graziano, G., and Barone, G. (1996) Temperature-induced denaturation of ribonuclease S: A thermodynamic study. *Biochemistry* 35, 13378–13385.
- (80) Neira, J. L., Sevilla, P., Menendez, M., Bruix, M., and Rico, M. (1999) Hydrogen exchange in ribonuclease A and ribonuclease S: Evidence for residual structure in the unfolded state under native conditions. *J. Mol. Biol.* 285, 627–643.
- (81) Santos, J., Sica, M. P., Buslje, C. M., Garrote, A. M., Ermacor, M. R., and Delfino, J. M. (2009) Structural selection of a native fold by peptide recognition. Insights into the thioredoxin folding mechanism. *Biochemistry* 48, 595–607.
- (82) Herzberg, O., and Moulton, J. (1991) Analysis of the steric strain in the polypeptide backbone of protein molecules. *Proteins* 11, 223–229.
- (83) Karplus, P. A. (1996) Experimentally observed conformation-dependent geometry and hidden strain in proteins. *Protein Sci.* 5, 1406–1420.
- (84) Penel, S., and Doig, A. J. (2001) Rotamer strain energy in protein helices: Quantification of a major force opposing protein folding. *J. Mol. Biol.* 305, 961–968.
- (85) Zhou, H. X., and Gilson, M. K. (2009) Theory of free energy and entropy in noncovalent binding. *Chem. Rev.* 109, 4092–4107.
- (86) Kim, S. H., de Vos, A., and Ogata, C. (1988) Crystal structures of two intensely sweet proteins. *Trends Biochem. Sci.* 13, 13–15.