4040 *Biochemistry* **2009,** *48,* 4040–4049
DOI:10.1021/bi9000936

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Characterization of DLC1-SAM Equilibrium Unfolding at the Amino Acid Residue Level†

Shuai Yang,‡ Christian G. Noble,∥ and Daiwen Yang*,§

‡*Department of Chemistry, 3 Science Drive 3, Faculty of Science, National University of Singapore, Singapore 117543,* §*Department of Biological Sciences, 14 Science Drive 4, Faculty of Science, National University of Singapore, Singapore 117543, and* ∥*Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673*

Received January 21, 2009; Revised Manuscript Received March 21, 2009

ABSTRACT: Sterile α motif (SAM) domains are found in many different proteins and shown to play important roles in various biological processes. The N-terminal domain of deleted in liver cancer 1 (DLC1) protein is a SAM domain which exists in a monomeric form in aqueous solution and facilitates the distribution of EF1A1 to the membrane periphery and ruffles upon growth factor stimulation. Here, we report the structure of an N-terminal truncated DLC1 SAM domain (DLC1-SAM) and its urea-induced equilibrium unfolding investigated with various biophysical methods such as CD, fluorescence emission spectroscopy, and NMR. CD and tryptophan intrinsic fluorescence emission data imply that the unfolding of DLC1-SAM follows a simple two-state mechanism, yet the NMR data suggest the presence of at least one intermediate state. The intermediate cannot be detected by NMR, but it does not exist in large aggregates as shown by analytical ultracentrifugation experiments. Analysis of the free energy values for different residues shows that in the transition from the native state to non-native states the C-terminal helix is somewhat more stable than the other parts of the protein, whereas in the transition from the native and intermediate states to the denatured state, the stabilities of different residues are similar except for that of the region surrounding residues D37−F40 which has lower stability and is more readily denatured at high urea concentrations. Analysis of the midpoints of the transitions shows that the unfolding of the native state and formation of the denatured state are not cooperative and the unfolding of only a few residues seems to follow a two-state mechanism.

Equilibrium unfolding is commonly used to determine the conformational stability of a protein and the sequence of events in a protein folding process. Equilibrium (un) folding is a reversible process of folding and unfolding in which equilibrium is maintained by gradually changing the environment of proteins. Optical methods such as circular dichroism (CD) and fluorescence spectroscopy are powerful in monitoring changes in the secondary and tertiary structures upon unfolding and give a macroscopic view of equilibrium unfolding of a protein. On the other hand, nuclear magnetic resonance (NMR) spectroscopy can provide equilibrium unfolding information on a per residue basis.

NMR has been applied to study the equilibrium (un) folding of a number of proteins (*1*−*12*). In the unfolding study of apoflavodoxin, the coincidence of transition midpoints that were obtained from monitoring the intensity changes of 21 $^{1}H-^{15}N$ HSQC correlations (cross-peaks) with denaturant concentrations suggested that the transition between the native and intermediate states was cooperative, whereas the data for the denatured species indicated the noncooperative transition between the intermediate and unfolded states (*1*). Similarly, a detailed picture of the equilibrium unfolding of a CheY mutant was obtained using NMR. The transition from the native to the molten globule-like intermediate was shown to be highly cooperative, yet as far as the transition between the intermediate and unfolded states was concerned, there seemed to be two folding subdomains in the sequence of unfolding, with the C-terminal subdomain experiencing unfolding first and the N-terminal subdomain remaining a collapsed globular conformation in the intermediate state (*2*). An unfolding study of ORF56 by NMR showed that both the unfolding of the native state and the formation of the unfolded state were cooperative and the population decrease of the native state mirrored the population increase of the denatured state in the unfolding process, strongly emphasizing a two-state folding mechanism for

†This work was supported by a grant from the Biomedical Research Council (BMRC) of the Agency for Science, Technology and Research, of Singapore. S.Y. thanks the National University of Singapore for financial assistance.

*To whom correspondence should be addressed. Phone: +65-65161014. Fax: +65-67792486. E-mail: dbsydw@nus.edu.sg.

 Published on Web 3/24/2009

ORF56 (*3*). In these unfolding studies, the cross-peaks from both the native and denatured states in the HSQC spectrum at each denaturant concentration or temperature were observable. However, in the unfolding study of $P19^{INK4d}$, $^{1}H-^{15}N$ cross-peaks from neither the native nor the denatured state appeared at moderate urea concentrations, clearly and directly revealing the presence of a third species at moderate urea concentrations (*4*).

Sterile α motif (SAM) domains are present in a number of proteins, including scaffolding proteins, transcription regulators, translational regulators, tyrosine kinases, and serine/threonine kinases (*13*−*15*). SAM domains have been shown to homo- and hetero-oligomerize and also to bind various non-SAM domain-containing proteins (*16*). In addition, recent studies have shown that SAM domains are also able to bind RNA (*17*,*18*). Nevertheless, the study of the (un)folding of SAM domain is really sparse. So far, temperature-, chemical-, and pH-induced unfolding of the SAM domain of p73 and urea-induced unfolding of the SAM domain of MAPKKK Ste11 were studied with different biophysical techniques (*19*,*20*).

In this paper, we present a comprehensive study of the equilibrium unfolding of an N-terminal truncated SAM domain of human DLC1 (hereafter called DLC1-SAM) using various biophysical methods, such as CD, fluorescence emission spectroscopy, and NMR. DLC1-SAM exists in a monomeric form in aqueous solution and facilitates the distribution of elongation factor-1A1 (EF1A1) to the membrane periphery and ruffles upon growth factor stimulation (*21*). The combination of CD and tryptophan fluorescence emission indicated that the unfolding of DLC1-SAM followed a simple two-state mechanism, yet a more detailed study on DLC1-SAM unfolding by NMR spectroscopy suggested the presence of at least one intermediate state.

MATERIALS AND METHODS

Sample Preparation. DLC1-SAM (N-terminal truncated form, K17−K76) was subcloned to a modified pET-H vector and overexpressed as a His_6 tag fusion protein in BL21(DE3) growing in M9 minimal medium. Isotope-enriched protein was prepared using $^{15}NH_4Cl$ and ^{13}C-labeled glucose as the sole nitrogen and carbon sources, respectively. After elution from the column and cleavage by thrombin, the protein was further purified by size-exclusion chromatography.

Denaturation Experiments. To ensure protein concentration is the same at different urea concentrations, we used one sample for one urea concentration. For CD and fluorescence experiments, each sample was prepared by adding a 1 mM protein stock solution to premixed solutions of urea and buffer. For NMR experiments, each sample was prepared by dissolving lyophilized protein in premixed solutions of urea and buffer. All experiments were performed at 25 °C in a buffer containing 70 mM sodium phosphate and 3 mM DTT at pH 7.0. The samples contained 20 μM, 3 μM, and 0.5 mM protein for CD, fluorescence, and NMR measurements, respectively. The samples were allowed to equilibrate for 2 h before spectra were recorded. For each NMR experiment, the probe was tuned and matched, the magnetic field was shimmed, and pulse widths were calibrated. Before and after acquisitions of two-dimensional (2D) spectra, one-dimensional (1D) spectra were recorded to check if there were any changes in the samples.

The urea concentration was calculated from the refractive index of each solution using the equation [urea] = $117.66n + 29.753n^2 + 185.56n^3$, where n represents the difference between the refractive indexes of the urea solution and the buffer in which urea was dissolved (*22*).

Fluorescence Emission Spectroscopy and Data Analysis. Tryptophan fluorescence spectra were recorded between 300 and 450 nm upon excitation at 295 nm on a Perkin-Elmer LS-50B luminescence spectrometer. A cuvette with a path length of 0.5 cm was used. Excitation and emission slit widths were both set to 4 nm. The fluorescence emission intensities at 358 nm when excited at 295 nm were plotted against urea concentration to obtain a transition curve. The experimental data were fitted to a two-state equation developed previously (*1*), based on the "linear dependent" model of free energy (*23*):

$$Y_{obs} = \frac{\alpha_N + \beta_N c}{1 + \exp\left[-\frac{m}{RT}(c_m - c)\right]} \quad (1)$$

where Y_{obs} is the experimental signal intensity in the presence of c molar urea, α_N and β_N are the intercept and slope of the pretransition zone, respectively, R is the gas constant, T is the absolute temperature in kelvin, c_m is the urea concentration at the transition midpoint, and m is the slope at the transition midpoint.

CD Spectroscopy and Data Analysis. Far-UV CD spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a thermal controller. Urea-induced unfolding was monitored at 222 nm with a 0.1 cm path length cuvette at a 0.1 nm spectral resolution. Each spectrum represented an average of 10 scans, and the scan rate was 20 nm/min. The transition curve was obtained by plotting the change in ellipticity at 222 nm against urea concentration. Experimental data obtained from CD spectroscopy were also analyzed by the aforementioned two-state equation.

NMR Spectroscopy and Data Processing for Unfolding Studies. All NMR experiments were performed using a Bruker Avance spectrometer with a ^{1}H frequency of 800.15 MHz. For each NMR sample used for the equilibrium unfolding study, one 2D gradient-enhanced HSQC spectrum was recorded; 640 and 128 complex data points were acquired in the ^{1}H and ^{15}N dimensions, respectively. The spectral width was 9615 and 1460 Hz in ^{1}H and ^{15}N dimensions, respectively. The data were apodized using a Gaussian multiplication in both t_2 and t_1, zero-filled to yield a final point-to-point resolution of 1.2 Hz in F_2 and 0.7 Hz in F_1, and then Fourier transformed. The resulting spectra were baseline-corrected in the F_2 dimension. Cross-peak volumes were determined using the nonlinear spectral line shape modeling option in NMRpipe/NMRDraw (*24*).

To measure T_2 values of amide protons (H_N), we inserted a spin-echo element (τ-180_{sel}-τ) in the first INEPT period in the HSQC pulse sequence. In the spin-echo element, τ is the relaxation delay and 180_{sel} denotes a selective 180° ^{1}H pulse with a REBURP shape profile which selectively refocuses amide protons. This selective

pulse removes the scalar coupling interaction between H_N and H_α protons. Although the chemical shift anisotropy (CSA)/dipole cross-correlated relaxation effect and conformational exchange contribution could not be suppressed by this simple scheme, the apparent transverse relaxation times of H_N spins during a spin-echo period can be measured. The relaxation times were determined using five different relaxation delays between 2 and 62 ms (Figure S1 of the Supporting Information). The measured H_N T_2 values were used to correct HSQC peak volumes. After correction, the volumes of each cross-peak from the native species under various urea concentrations were normalized with respect to the volume in the absence of urea. Similarly, the volumes of each peak from the denatured species were normalized with respect to the volume in the presence of 9.57 M urea. In this way, the relative populations of native and denatured forms were obtained.

NMR Data Analysis. The intensity of each $^1H-^{15}N$ cross-peak from the native state decreases with the increase in urea concentration and finally reaches zero. This is similar to the change in CD ellipticity at 222 nm with urea concentration. Therefore, eq 1, which is used for the analysis of CD and fluorescence data, can be applied to analyze the disappearance of each native cross-peak. After relaxation correction, the population of the native species can be considered unaltered in the pretransition zone; namely, the intercept (α_N) and slope (β_N) in the pretransition regime are equal to 1 and 0, respectively. In this case, the disappearance of a native peak can be fitted with a simplified form of eq 1

$$N = \frac{1}{1 + \exp\left[-\frac{m_1}{RT}(c_{m1} - c)\right]} \quad (2)$$

Similarly, the emergence of a cross-peak from the denatured species can be fitted to the following equation:

$$D = \frac{\exp\left[-\frac{m_2}{RT}(c_{m2} - c)\right]}{1 + \exp\left[-\frac{m_2}{RT}(c_{m2} - c)\right]} \quad (3)$$

where N and D denote the normalized populations of the native and denatured species, respectively, c is the urea concentration, c_{m1} and c_{m2} are the urea concentrations at the transition midpoints, and m_1 and m_2 are the slopes at the transition midpoints of the disappearance of the native species and the emergence of the denatured species, respectively.

According to the "linear dependency model" (*25*,*26*), ΔG_1^0 and ΔG_2^u can be related to urea concentrations and transition midpoints c_{m1} and c_{m2}, respectively, using the following equations:

$$\Delta G_1^0 = m c_{m1} \quad (4)$$

$$\Delta G_2^u = m_2(c_{m2} - 9.57) \quad (5)$$

Substituting eqs 4 and 5 into eqs 2 and 3 gives

$$N = \frac{1}{1 + \exp\left[-\frac{1}{RT}(\Delta G_1^0 - m_1 c)\right]} \quad (6)$$

and

$$D = \frac{\exp\left[-\frac{\Delta G_2^u - m_2(c - 9.57)}{RT}\right]}{1 + \exp\left[-\frac{\Delta G_2^u - m_2(c - 9.57)}{RT}\right]} \quad (7)$$

where ΔG_1^0 is the extrapolation of the free energy of the disappearance of the native state to 0 M urea and ΔG_2^u is the extrapolation of the free energy of the emergence of the denatured state to 9.57 M urea where 100% denatured state is populated. Fitting the experimental data N and D to eqs 2, 3, 6, and 7 provides the conformational stability of the region surrounding each amino acid residue and the transition midpoints. The data analyses were performed using ORIGIN 8.0 (Originlab, Northampton, MA).

Sedimentation Velocity Experiments and Data Analysis. DLC1-SAM samples were prepared by dialysis against buffer solutions containing 50 mM sodium phosphate, 200 mM sodium chloride, 0.1% β-mercaptoethanol, and various urea concentrations (0, 4, and 8 M) at pH 7.0. The final protein concentration was ~15 μM. Protein samples (420 μL) and reference solutions (dialysis buffer solutions, 440 μL) were loaded into standard double-sector centerpieces (12 mm optical path length), and the centerpieces were mounted in a Beckman An-50 Ti rotor. Sedimentation velocity experiments were performed at 20 °C on a Beckman Coulter XL-I analytical ultracentrifuge operating at a rotor speed of 42000 rpm. Data were collected in a continuous mode, at a single wavelength of 280 nm. The partial specific volume of the protein was calculated on the basis of the amino acid composition, while the density of the solvent was calculated from the chemical composition of the buffer. Multiple scans at different time points were fitted to a continuous size distribution using SEDFIT (*27*,*28*).

RESULTS AND DISCUSSION

Structure of the N-Terminal Truncated SAM Domain of DLC1. The full-length SAM domain (M1−K76) is so stable that it can be completely unfolded only at an extremely high urea concentration (>9 M), and even at ~10 M urea, NMR unfolding curves corresponding to the population increase of the unfolded species still could not reach plateaus. Thus, a truncated form DCL1-SAM (K17−K76) was prepared. Sequence-specific assignments of DLC1-SAM in the absence of urea and in the presence of 7 M urea were obtained using the standard triple-resonance NMR experiments (Supporting Information and Materials and Methods). Experimental restraints and structural statistics for the solution structure of DLC1-SAM [Protein Data Bank (PDB) entry 2KAP] are documented in Table S1. The three-dimensional (3D) structure of DLC1-SAM (Figure 1) has the basic architecture of the full-length SAM domain determined by us [PDB entry 2GYT (*21*)]. Because the first helix of DLC1-SAM is significantly shorter than that of the full-length SAM and the truncated helical region interacts with the last helix, DLC1-SAM is less resistant to urea-induced unfolding than the full-length SAM and suitable for unfolding studies.

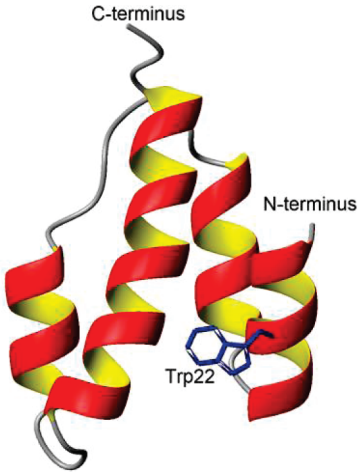

FIGURE 1: Solution structure of DLC1-SAM (PDB entry 2KAP) created with MOLMOL (*29*). The only tryptophan residue is colored blue.

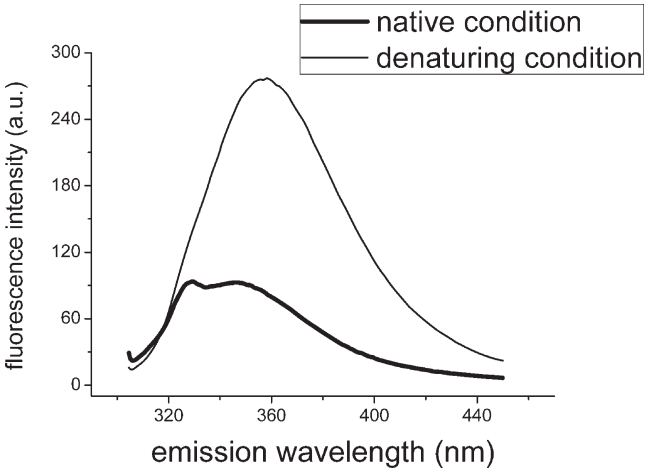

FIGURE 2: Fluorescence spectra of DLC1-SAM in the absence of urea (thick line) and in the presence of 9 M urea (thin line). The samples contained 2 μM protein, 3 mM DTT, and 50 mM sodium phosphate (pH 7).

Equilibrium Unfolding of DLC1-SAM Monitored by CD and Tryptophan Fluorescence. The changes in secondary structures and local geometry of W22 of DLC1-SAM were monitored macroscopically by CD and fluorescence emission spectroscopy, respectively. Since DLC1-SAM has a substantial amount of α-helical secondary structure, the ellipticity at 222 nm was used to track the change in the secondary structure of DLC1-SAM upon protein unfolding. Fluorescence emission spectroscopy allows the probing of the local geometry around W22 as a function of external circumstances. When the native protein was excited at 295 nm, the maximum emission wavelength was at 347 nm, demonstrating that the tryptophan residue was at the protein surface making contact with bound water and other polar groups. Treating DLC1-SAM with 9 M urea leads to a red shift of the fluorescence emission maximum to 358 nm (Figure 2), indicating that W22 was completely exposed to solvent under this condition. The increased fluorescence intensity upon unfolding suggested that tryptophan fluorescence was more severely quenched in the native state than in the unfolded state. The more effective quenching probably

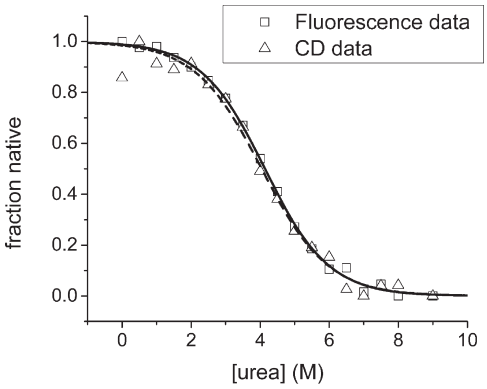

FIGURE 3: Comparison of unfolding curves obtained by CD and fluorescence emission spectroscopy. The ellipticity and fluorescence emission intensity values were normalized, corresponding to the fraction of native characters that remain in various denaturing conditions. CD and fluorescence data are shown with open triangles and squares, respectively. Fits of CD and fluorescence data with the two-state model are shown with solid and dashed curves, respectively.

resulted from a charged arginine residue (R65, a very effective dynamic quencher) that is in the proximity of W22 in space only in the native protein structure. In our study, the fluorescence at 358 nm was used to monitor the alteration in the local geometry of the tryptophan residue upon unfolding.

A combination of CD and fluorescence spectroscopy is often used to demonstrate whether the equilibrium (un) folding of a protein takes place via a two-state mechanism. A protein molecule that unfolds according to a two-state model is characterized by coinciding unfolding curves obtained by different spectroscopic techniques that track the changes in different probes in a protein molecule (*30*). The normalized fluorescence and CD unfolding curves showed similar trends: little change in signals at urea concentrations between 0 and 2 M, followed by a not very cooperative decrease from 2 to 6 M urea, at last reaching a plateau at urea concentrations of >6 M (Figure 3). A two-state unfolding mechanism (eq 1), in which only native and denatured states are populated during the unfolding, fitted remarkably well to the experimental data obtained from CD and fluorescence. c_m and m, determined from the fit of fluorescence data to the two-state equilibrium model, were 4.14 M and 2.67 kJ mol^{-1} M^{-1}, respectively, and 4.04 M and 2.56 kJ mol^{-1} M^{-1} from the fit of the CD data, respectively. The almost identical midpoints and cooperative indices within experimental errors indicate that the unfolding of DLC1-SAM might follow a simple two-state process.

CD and fluorescence spectroscopy monitor denaturant-induced protein unfolding in an only macroscopic way. It is of more interest to gain microscopic information pertinent to structural changes in the unfolding process. Therefore, NMR spectroscopy was used to explore structural changes in a residue-specific manner and to provide a more detailed representation of protein unfolding.

Urea-Induced Equilibrium Unfolding Monitored by NMR. The ^{1}H$-^{15}$N HSQC spectrum of a protein serves as a fingerprint of its conformational state. To monitor the denaturation of DLC1-SAM, a series of ^{1}H$-^{15}$N

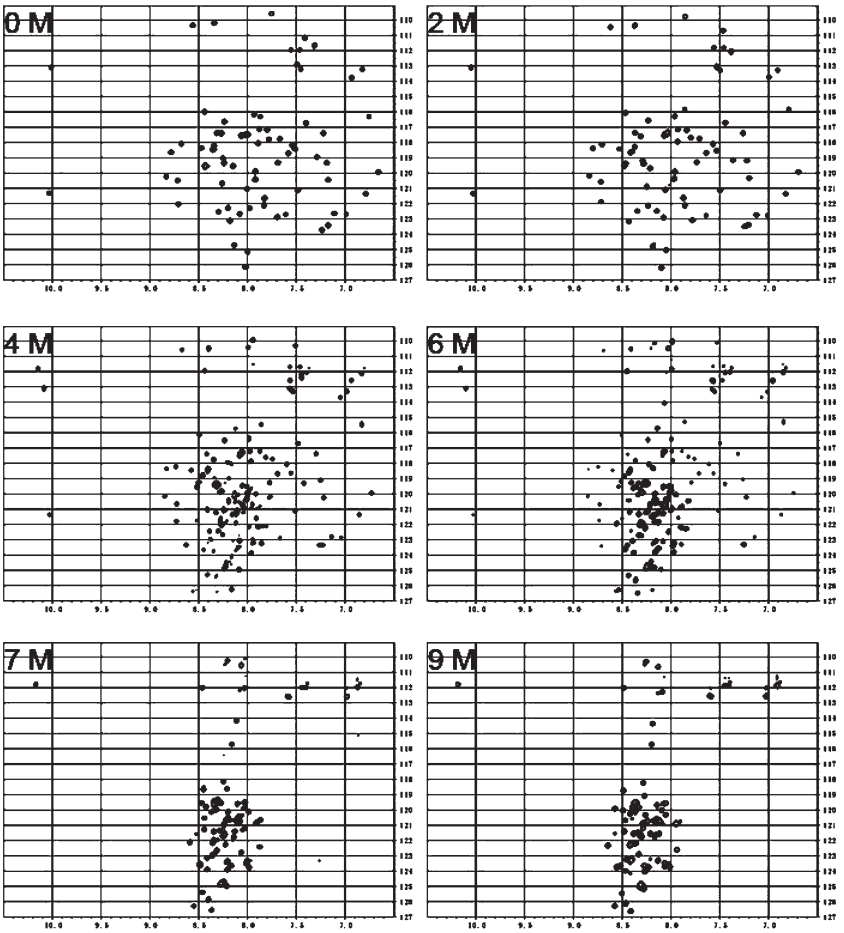

FIGURE 4: HSQC spectra of DLC1-SAM at different urea concentrations which are indicated in each panel. The horizontal dimension in each panel corresponds to ^{1}H chemical shifts, while the vertical dimension corresponds to ^{15}N chemical shifts.

HSQC spectra at different urea concentrations were recorded under equilibrium conditions. At a urea concentration of <2 M, only one set of cross-peaks corresponding to the native state was observed in the HSQC spectra; no additional peaks from other states were found (Figure 4, top panel, pretransition zone). For most native cross-peaks, ^{1}H and ^{15}N chemical shifts did not change significantly, implying that in the pretransition zone the global α-helical structure of DLC1-SAM was not affected. In the spectra recorded at urea concentrations between 2 and 6 M, some extra cross-peaks that resonated at 8.2 ± 0.3 ppm in the ^{1}H dimension (positions expected for unfolded proteins) were emerging. Hence, each spectrum consisted of two sets of cross-peaks, one corresponding to the native state and the other to the unfolded state (in the transition zone of unfolding), indicating that the chemical exchange between the native and unfolded states was rather slow on the NMR time scale. The volumes of native cross-peaks progressively decreased, and the volumes of unfolded cross-peaks gradually increased (Figure 4, middle panel). As the urea concentration increased to >6 M, the chemical shift dispersion in the ^{1}H dimension was strongly reduced to 7.9−8.5 ppm, typical for a denatured protein (Figure 4, bottom panel, post-transition zone). A detailed unfolding study by NMR spectroscopy was made possible by nearly complete assignments of the native and denatured protein. For the native DLC1-SAM in the absence of urea, all non-proline residues were assigned, and for the denatured DLC1-SAM at 7 M urea, all non-proline residues except K17 were assigned. In the unfolding process, the cross-peaks for the native and denatured forms shifted slightly with urea concentration. Assignments of the native cross-peaks at different denaturant concentrations were established on the basis of the prior assignment of the folded state in the absence of urea by following the gradual chemical shift changes. In a similar way, the cross-peaks of the unfolded form at different urea concentrations were obtained from the prior assignment of the unfolded protein in the presence of 7 M urea. The residues that had no obvious peak overlap and could be unambiguously assigned at more than seven urea concentrations are listed in Table S2A,B of the Supporting Information.

Correction of Cross-Peak Volumes. The line widths of cross-peaks were greatly dependent on urea concentration due to the changes in solution viscosity and amide−water exchange rates. Thus, peak volumes instead of peak intensities were used to evaluate the populations of the native and denatured states. Before data acquisition, NMR signals are attenuated during magnetization transfer and gradient selection periods in the gradient-enhanced HSQC experiment (*31*). The attenuation factor (A_f) is given by

$$A_f = 0.5 \exp[-(4\tau + 2\delta_2)/T_{2,H}][\exp(-2\tau/T_{2,MQ}) + \exp(-2\tau/T_{1,H})]$$

where τ (2.5 ms) and δ_2 (1 ms) are the delays used in the INEPT and gradient selection periods, respectively, in the HSQC experiment (Figure 1 in ref *31*), $T_{2,H}$ and $T_{1,H}$ are the transverse and longitudinal relaxation times of an amide proton, respectively, and T_{MQ} is the average relaxation time for amide H−N multiple quantum coherences. For nondeuterated proteins, $T_{2,H}$ and T_{MQ} are dominated by dipolar interactions between the amide proton and its proximal protons; i.e., $T_{MQ} \approx T_{2,H}$. Since $T_{1,H} \gg \tau$, A_f can be approximated as

$$A_f \approx 0.5 \exp[-(4\tau + 2\delta_2)/T_{2,H}][\exp(-2\tau/T_{2,H}) + 1]$$

Since A_f varies from one residue to another in a protein, the peaks in an HSQC spectrum often have different intensities or volumes, although each peak represents only one amide. Only after peak volumes are properly corrected can they be used to quantify relative conformer populations at a series of denaturant concentrations. The corrected volume (V^{cr}) is given by

$$V^{cr} = V^{ex}/A_f$$

where V^{ex} is the volume of a peak measured from an HSQC spectrum. For DLC1-SAM, $T_{2,H}$ values varied from ~50 to ~20 ms and A_f from 0.75 to 0.49 when urea concentrations were changed from 0 to 5 M. Thus, the correction is necessary. Because urea affects not only solution viscosity but also amide−water hydrogen exchange rates, the $T_{2,H}$ value at one urea concentration for a given amide cannot be predicted from the value measured at another urea concentration. To obtain an accurate unfolding profile, it is necessary to measure $T_{2,H}$ at each urea concentration.

To the best of our knowledge, no relaxation correction has been applied to quantify the relative populations of folded and unfolded species at different denaturant concentrations, although the changes in relaxation upon addition of denaturant can be significant. In the study on apoflavodoxin unfolding, a correction procedure was proposed to correct the effects of protein dilution and aggregation and Q factor reduction of the NMR receiver coil induced by GndHCl titration on peak volume (*1*). In our study, the dilution and aggregation effects were minimized by using one sample for one denaturant concentration. Urea does not change the conductivity of the sample, and thus, the Q factor is independent of urea concentration. Our correction method should be applicable to other proteins that exhibit changes in relaxation behavior during the unfolding process and allows unfolding parameters to be determined more accurately.

Unfolding Equilibrium Intermediate Revealed by NMR Spectroscopy. A detailed picture of urea-induced unfolding of DLC1-SAM was achieved by analyzing unfolding curves of residues from native and denatured species which were well resolved in most HSQC spectra at a series of urea concentrations. Normalized populations of the native and denatured species at various urea concentrations are documented in Table S2 of the Supporting Information. Transition curves for different amide groups in the native form did not coincide with one another (Figure 5). The experimental data were fitted to a two-state model. Only fitting results with an adjusted

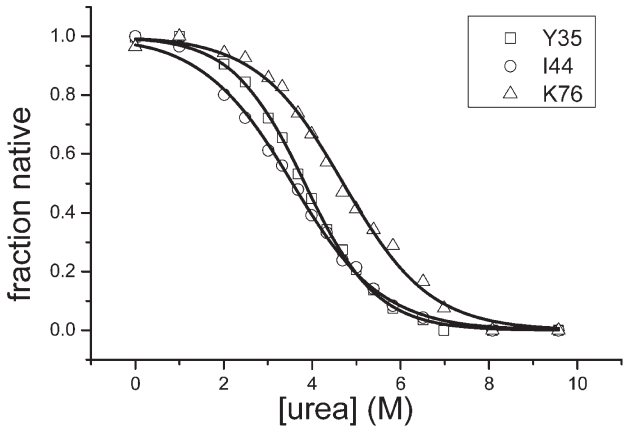

FIGURE 5: Native transition curves for Y35, I44, and K76. The fits of NMR data to a simple two-state mechanism are shown as smooth curves.

R^2 value (coefficient of determination) of ≥0.95 were considered acceptable, yielding unfolding parameters from 47 peaks assigned to the native state and 39 peaks assigned to the denatured state (Table S3 of the Supporting Information). The resultant transition midpoints (c_{m1}) ranged from 3.68 to 4.81 M urea, and cooperative indices (m_1) ranged from 2.11 to 3.23 kJ mol^{-1} M^{-1} (Table S3). Clearly, individual residues in the native protein behaved differently during the unfolding process, indicating that the unfolding of DLC1-SAM does not follow a simple two-state process in which all amino acid residues in the native protein should unfold concurrently (*3*).

When the populations of both native and denatured species were plotted against urea concentrations in one graph, there was a pronounced lag between the disappearance of the native species and the emergence of the denatured species. Hence, these two curves did not intersect at a fraction value of 0.5 (Figure 6), and the sum of the populations of native and denatured forms was not equal to unity in the transition zone. This result indicates that at least one intermediate species should exist in the equilibrium unfolding of DLC1-SAM. Thus, the DLC1-SAM equilibrium unfolding can be described approximately as follows:

$$\text{native} \Leftrightarrow \text{intermediate} \Leftrightarrow \text{denatured}$$

in which the native state first evolves to the intermediate state and then the intermediate state further unfolds to the denatured state. The population of the intermediate can be calculated by subtracting the populations of the native and denatured forms from unity.

DLC1-SAM Unfolding in a Residue-Specific Way. One aim of this study is to explore the behavior of each amino acid residue and the stability of regions surrounding individual residues throughout the equilibrium unfolding process of DLC1-SAM. Gibbs free energy changes signify the stability in different states. The free energy change of the transition between the native state and non-native states (i.e., intermediate and denatured states) in the absence of urea (ΔG_1^0) provides information about the stability of the region surrounding a given residue in the native state compared to the non-native states. Higher ΔG_1^0 values correspond to regions with higher stability in the native state. ΔG_1^0 obtained here exhibits two sets of

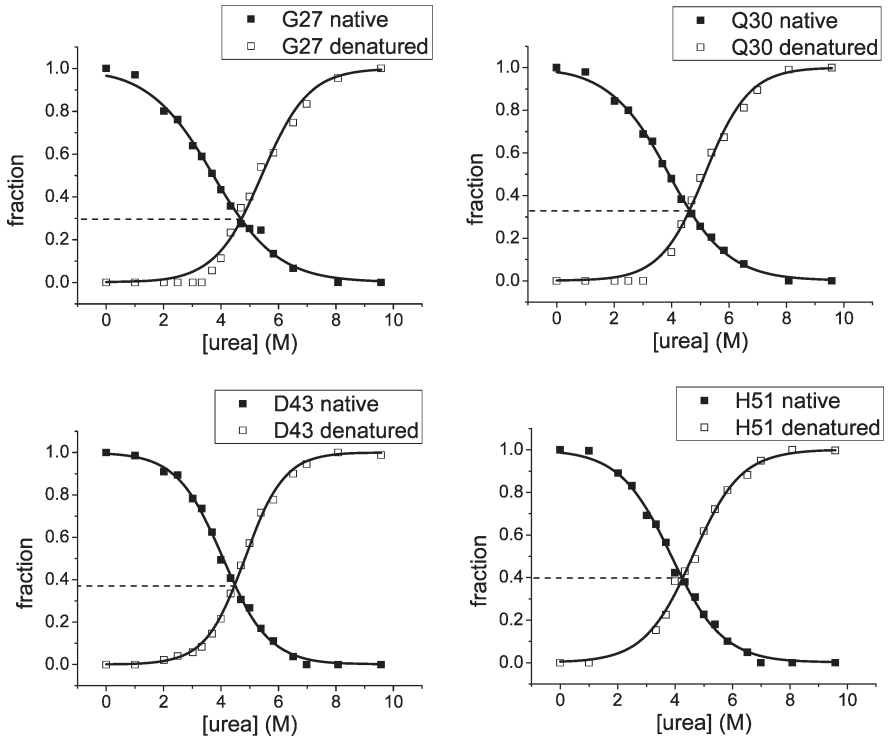

FIGURE 6: Native (■) and denatured (□) transition curves for residues G27, Q30, D43, and H51. The native and denatured transition curves of G27 intersect at a fraction of 0.30, those of Q30 at 0.32, those of D43 at 0.38, and those of H51 at 0.39 (indicated by dashed lines).

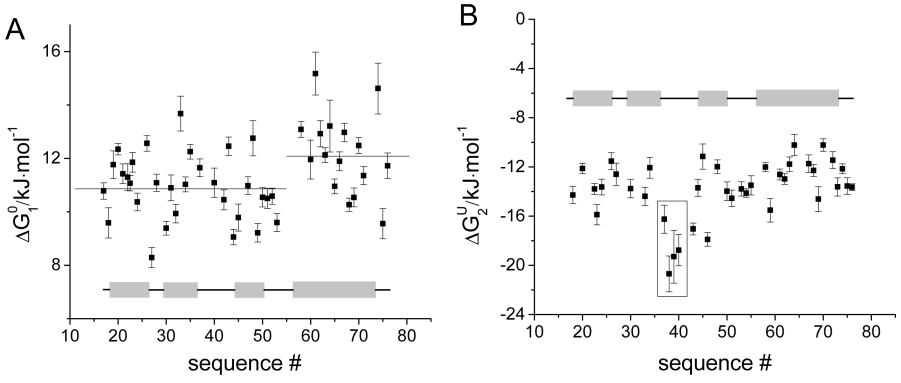

FIGURE 7: Variation of free energy values (ΔG_1^0 and ΔG_2^u) along the protein sequence. (A) Average ΔG_1^0 values of the N- and C-terminal parts of the protein (straight lines). (B) Residues 37−40 are enclosed in a box. The helical secondary structures are illustrated as gray rectangles.

values (Figure 7A). The average ΔG_1^0 value for residues from the N-terminal part of the sequence (from K17 to F53) was 10.91 kJ/mol, whereas the average ΔG_1^0 value for residues from the C-terminal part of the sequence (A58 to K76) was 12.18 kJ/mol. Residues A58−K76 correspond to the long C-terminal helix. The higher average ΔG_1^0 value for this part of the protein suggests that the C-terminal helix is slightly more stable than the other parts of the protein, which is consistent with the solution structure of native DLC1-SAM in which all three other helices have extensive contacts and interactions with the long C-terminal helix yet not many interactions are found among these three helices. Also, the greater number of H-bonds formed in this longer C-terminal helix could contribute to its greater stability.

The free energy value obtained from the fluorescence data was 11.03 kJ/mol, which is nearly the same as the free energy obtained from the disappearance of the native W22 indole amide cross-peak, 11.07 kJ/mol (Table S3). This suggests that the intermediate species has a fluorescence property more similar to that of the denatured species than that of the native species. The free energy obtained from the CD data was 10.33 kJ/mol, while the average free energy obtained from the disappearance of native NMR signals was 11.34 kJ/mol. The difference may result from the fact that CD measures the changes in the secondary structure while NMR measures the changes in both secondary and tertiary structures.

The free energy change of the transition from the nondenatured states (i.e., native and intermediate states) to the denatured state, ΔG_2^u, was determined to evaluate the stability of the nondenatured state relative to the denatured state in different regions of DLC1-SAM under an extreme denaturing condition. Low ΔG_2^u values refer to residues located in regions which more readily unfold or which have a low stability in the nondenatured states. The variation of ΔG_2^u values along the protein sequence is shown in Figure 7B. It is noteworthy that the regions

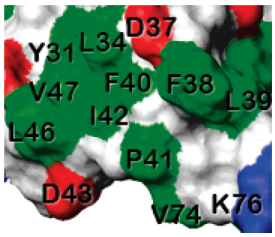

FIGURE 8: Part of the van der Waals surface of DLC1-SAM created with MOLMOL (*29*), illustrating potential protein−protein association interfaces. The side chains are colored green for hydrophobic, blue for basic, red for acidic, and gray for polar (and main chain).

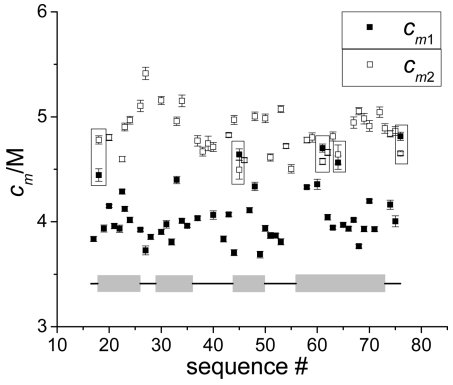

FIGURE 9: Variation of c_{m1} (■) and c_{m2} (□) values along the protein sequence. E18, S45, A61, R64, and K760 are enclosed in a box. The helical secondary structures are illustrated as gray rectangles.

surrounding residues D37−F40 have lower ΔG_2^u values, indicating that these regions are less stable in the native and intermediate states and much more readily denatured at high urea concentrations. These residues happen to be located around the FLF motif of DLC1-SAM (F38, L39, and F40) which is responsible for the binding to EF1A1 (*21*), insinuating that the stability of this region might have something to do with the biological function of the protein. However, this assumption has yet to be verified. Some other residues, i.e., D43 and L46, also have low ΔG_2^u values, but they appear to be slightly distant from the aforementioned region in space (Figure 8) and, therefore, are not likely a part of the motif.

Transition midpoints are indicators of the sequence of events in equilibrium unfolding. The midpoints of the transition from the native state to non-native states (c_{m1}) for different residues are shown in Figure 9 (■). These residues exhibited quite diverse transition midpoints, indicating that the unfolding of the native state was noncooperative. Most of the residues have c_{m1} values of ~4 M urea, the same as those obtained by the global analysis of the optical data. On the other hand, a minority of residues (E18, S45, A61, R64, and K76) have larger c_{m1} values. For these residues, their c_{m1} values are comparable to their respective c_{m2} values. It seems that the unfolding of these residues is governed by an apparent two-state mechanism, and the decrease in the native population of these residues mirrors the increase in the denatured population of the same residue (Figure 10). It is possible that for these residues the population of the intermediate form is so small that it could not be reflected by the transition curves. It is also possible that these residues maintain a nativelike conformation in the intermediate state; hence, the apparent population of the native species corresponds to the sum of the native and intermediate states. Similar to the unfolding of the native state, the formation of the denatured state is not cooperative since c_{m2} values were distributed in the range of 4.5−5.5 M (Figure 9). The noncooperative equilibrium unfolding probably results from the extensive tertiary interactions in the globular DLC1-SAM. Such noncooperative unfolding is frequently observed for globular α-helical proteins (*32*). However, for some elongated nonglobular proteins which lack significant tertiary interactions, the equilibrium unfolding is usually concerted for all amino acid residues (*4*).

No Aggregation for the Equilibrium Unfolding Intermediate in the Unfolding Process. NMR data show that at least one intermediate state exists in the unfolding of DLC1-SAM. To find out whether the intermediate state exists in an aggregated form, sedimentation velocity experiments were conducted. The samples at 0 and 8 M urea correspond to native and completely denatured forms, respectively, and the sample at 4 M urea contains a large amount of the intermediate form based on NMR results. Analytical ultracentrifugation data are shown in Figure 11. Data analysis for all three samples showed low root-mean-square deviation (rmsd) values (< 0.05) and no systematic deviation (data not shown). In the sedimentation coefficient distribution plots of all three samples, we can see only one distinct peak and no sign of the presence of other faster sedimenting species, indicating that large protein aggregates did not exist in any of these three protein samples. The native sample had a frictional ratio of 1.28, suggesting that the native protein had a compact globular structure. The peak of the native protein was found at a sedimentation coefficient of 1.01 S (Figure 11, black), corresponding to a protein size of 7564 ± 1312 Da, which matches the expected value (7738 Da) well within 3% error. The protein sample in the presence of 8 M urea had a frictional ratio of 2.03, characteristic of a denatured polypeptide chain. The peak was found at a sedimentation coefficient of ~0.1 S (Figure 11, red), corresponding to a molecular mass of 5633 ± 4537 Da. At 4 M urea, the protein sample was a mixture of the native, intermediate, and unfolded species. The sedimentation velocity method is supposed to be capable of detecting changes in protein conformation under native and extremely denaturing conditions. However, because of the small size of DLC1-SAM, the peaks for the native, intermediate, and denatured species overlap and become indistinguishable from one another at the intermediate urea concentration (4 M). The frictional ratio of this sample was 1.61, and the peak was centered at a sedimentation coefficient of ~0.7 S (Figure 11, blue), corresponding to a molecular mass of 10022 ± 1284 Da.

The molecular mass of the native protein calculated by analytical ultracentrifugation data was quite close to the actual native protein molecular mass. Nevertheless, the calculated molecular mass of the denatured species deviated significantly from the actual molecular mass of DLC1-SAM. This is because the peak position is so close to zero that the sedimentation coefficient of the denatured protein becomes inaccurate. This is also true for the protein sample in the presence of 4 M urea. The big difference

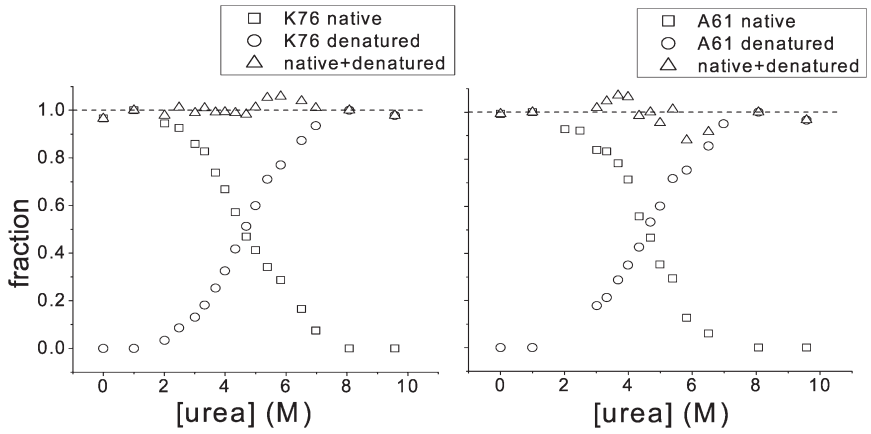

FIGURE 10: Plots of the fractions of native (□) and denatured (○) species and the sum of the populations of both species (△) for residues A61 and K76 vs urea concentration. Unity is highlighted using dashed lines.

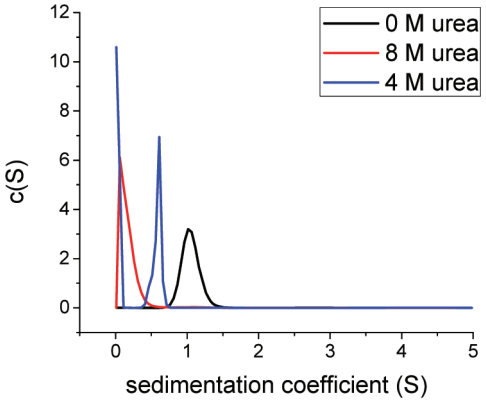

FIGURE 11: Distribution of the sedimentation coefficients of DLC1-SAM in the native form (black) and in the denatured form (red) and in the presence of 4 M urea (blue).

between the measured and expected molecular masses for the 4 M urea sample may also result from oligomerization of the intermediate species at 4 M urea. As can be seen from the distribution plot of this sample, the right end of the sample peak is at a sedimentation coefficient of ~0.9 (Figure 11, blue), corresponding to a molecular mass of 16.7 kDa. Therefore, even if the intermediate oligomer indeed existed, the size of the oligomer would not exceed a trimer.

Altogether, from sedimentation velocity results, we confirm that large fast sedimenting protein aggregates were not present in any of the protein samples. In particular, even if the intermediate species self-associated, the intermediate oligomer(s) could not be larger than trimer. Therefore, the absence of peaks from the intermediate species in HSQC spectra at intermediate urea concentrations is not caused by the formation of large protein aggregates. Instead, it is most probably that the intermediate state is an ensemble of monomeric structures which interconvert on millisecond time scales. The conformational interconversion gives rise to dramatic contributions to the transverse relaxation rates of amide ^{1}H and/or ^{15}N (*33*) and results in no HSQC cross-peaks for the intermediate species.

Equilibrium Unfolding of Different SAM Domains. In addition to the unfolding of DLC1 SAM domain, the equilibrium unfolding of the SAM domains of p73 (p73-SAM) and Ste11 (Ste11-SAM) have also been studied previously (*19*,*20*). Temperature-, chemical-, and pH-induced unfolding of p73-SAM was studied macroscopically using CD, fluorescence, Fourier transform infrared spectroscopy, and differential scanning calorimetry (*19*). Those probes have indicated that p73-SAM folded via a two-state mechanism, other than the three-state mechanism governing the unfolding of DLC1-SAM. This is not surprising since DLC1-SAM is a unique member of the SAM superfamily. It has a bundled structure instead of a globular structure that is adopted by many other SAM domains, including p73-SAM (*21*). The variation in the 3D structure of these SAM domains is expected to result in the difference in the unfolding properties. All the probes for studying the unfolding of p73-SAM monitor only the changes in biophysical properties in the unfolding process macroscopically; however, it is of more interest to examine the unfolding of SAM domains in detail, at a higher resolution that can be achieved by NMR spectroscopy. The residue-specific unfolding of Ste11-SAM was investigated using NMR. On the basis of the coinciding unfolding transition curves of many residues corresponding to different regions of the native conformation, the authors claimed that the unfolding of Ste11-SAM was highly cooperative without any intermediate species. The transition may be cooperative, yet the presence of the intermediate state cannot be ruled out, since it has been suggested that sometimes the native state unfolds and evolves into the intermediate state cooperatively (*1*,*2*). In addition, the characterization of the unfolding of a protein at the residue level requires the complete assignment of the backbone amide resonances of the fully denatured form. Without monitoring the denatured form, an incomplete or even biased description of the molecular details of the unfolding transitions would be inevitable. That is why we have examined not only the population decrease of the native species but also the population increase of the denatured species. The lag between the disappearance of the native species and the emergence of the denatured species revealed the presence of at least one relatively stable intermediate state at moderate urea concentrations. Our study has provided a very comprehensive interpretation of the experimental data concerning the unfolding process of DLC1-SAM.

CONCLUSION

In this paper, we have studied the stability and equilibrium unfolding of the truncated SAM domain of DLC1. The urea-induced unfolding of DLC1-SAM was investigated using various biophysical methods, such as CD, fluorescence emission spectroscopy, and NMR. The unfolding curves obtained by CD and tryptophan intrinsic fluorescence emission coincided, suggesting a simple two-state folding mechanism. However, NMR data revealed the presence of at least one intermediate state involved in equilibrium unfolding. The intermediate does not form large aggregates and cannot be detected by NMR. The detailed information about the unfolding transitions was obtained by fitting NMR experimental data to the simplest two-state model for each residue. In the transition from the native state to non-native states, the C-terminal helix is slightly more stable than the other parts of the protein, whereas in the transition from the native and intermediate states to the denatured state, the stabilities of regions surrounding different residues were similar except for the region surrounding residues in the FLF motif which are much more readily denatured at high urea concentrations. The transition from the native state to non-native states is not cooperative: very few residues have large transition midpoints and seem to unfold via a two-state mechanism. The exploration of the unfolding of DLC1-SAM should have a profound impact on the unfolding studies of proteins with a DLC1-SAM-like fold and provide new insights into the elementary equilibrium folding processes of small proteins.

SUPPORTING INFORMATION AVAILABLE

Supporting information and methods, one figure showing amide proton T_2 decay profiles, one table listing experimental restraints and structural statistics for the DLC1-SAM structure, one table listing the relative populations of residues in the native and denatured states at a series of urea concentrations, and one table listing the residue level thermodynamic information. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. van Mierlo, C. P., van den Oever, J. M., and Steensma, E. (2000) Apoflavodoxin (un)folding followed at the residue level by NMR. *Protein Sci. 9*, 145–157.
2. Garcia, P., Serrano, L., Rico, M., and Bruix, M. (2002) An NMR view of the folding process of a CheY mutant at the residue level. *Structure 10*, 1173–1185.
3. Zeeb, M., Lipps, G., Lille, H., and Balbach, J. (2004) Folding and association of an extremely stable dimeric protein from *Sulfolobus islandicus*. *J. Mol. Biol. 336*, 227–240.
4. Zeeb, M., Rosner, H., Zeslawski, W., Canet, D., Holak, T. A., and Balbach, J. (2002) Protein folding and stability of human CDK inhibitor p19(INK4d). *J. Mol. Biol. 315*, 447–457.
5. Tang, Y., Grey, M. J., McKnight, J., Palmer, A. G.III, and Raleigh, D. P. (2006) Multistate folding of the villin headpiece domain. *J. Mol. Biol. 355*, 1066–1077.
6. Latypov, R. F., Cheng, H., Roder, N. A., Zhang, J., and Roder, H. (2006) Structural characterization of an equilibrium unfolding intermediate in cytochrome c. *J. Mol. Biol. 357*, 1009–1025.
7. Kumar, A., Srivastava, S., Kumar Mishra, R., Mittal, R., and Hosur, R. V. (2006) Residue-level NMR view of the urea-driven equilibrium folding transition of SUMO-1 (1−97): Native preferences do not increase monotonously. *J. Mol. Biol. 361*, 180–194.
8. Watters, A. L., Deka, P., Corrent, C., Callender, D., Varani, G., Sosnick, T., and Baker, D. (2007) The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell 128*, 613–624.
9. Latypov, R. F., Harvey, T. S., Liu, D., Bondarenko, P. V., Kohno, T., Fachini, R. A.II, Rosenfeld, R. D., Ketchem, R. R., Brems, D. N., and Raibekas, A. A. (2007) Biophysical characterization of structural properties and folding of interleukin-1 receptor antagonist. *J. Mol. Biol. 368*, 1187–1201.
10. Vadrevu, R., Wu, Y., and Matthews, C. R. (2008) NMR analysis of partially folded states and persistent structure in the α subunit of tryptophan synthase: Implications for the equilibrium folding mechanism of a 29-kDa TIM barrel protein. *J. Mol. Biol. 377*, 294–306.
11. Latypov, R. F., Maki, K., Cheng, H., Luck, S. D., and Roder, H. (2008) Folding mechanism of reduced cytochrome c: Equilibrium and kinetic properties in the presence of carbon monoxide. *J. Mol. Biol. 383*, 437–453.
12. Mohan, P. M., and Hosur, R. V. (2008) pH dependent unfolding characteristics of DLC8 dimer: Residue level details from NMR. *Biochim. Biophys. Acta 1784*, 1795–1803.
13. Kim, C. A., and Bowie, J. U. (2003) SAM domains: Uniform structure, diversity of function. *Trends Biochem. Sci. 28*, 625–628.
14. Schultz, J., Ponting, C. P., Hofmann, K., and Bork, P. (1997) SAM as a protein interaction domain involved in developmental regulation. *Protein Sci. 6*, 249–253.
15. Ponting, C. P. (1995) SAM: A novel motif in yeast sterile and *Drosophila* polyhomeotic proteins. *Protein Sci. 4*, 1928–1930.
16. Peterson, A. J., Kyba, M., Bornemann, D., Morgan, K., Brock, H. W., and Simon, J. (1997) A domain shared by the Polycomb group proteins Scm and ph mediates heterotypic and homotypic interactions. *Mol. Cell. Biol. 17*, 6683–6692.
17. Aviv, T., Lin, Z., Lau, S., Rendl, L. M., Sicheri, F., and Smibert, C. A. (2003) The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. *Nat. Struct. Biol. 10*, 614–621.
18. Green, J. B., Gardner, C. D., Wharton, R. P., and Aggarwal, A. K. (2003) RNA recognition via the SAM domain of Smaug. *Mol. Cell 11*, 1537–1548.
19. Barrera, F. N., Garzon, M. T., Gomez, J., and Neira, J. L. (2002) Equilibrium unfolding of the C-terminal SAM domain of p73. *Biochemistry 41*, 5743–5753.
20. Bhunia, A., Domadia, P. N., Xu, X., Gingras, R., Ni, F., and Bhattacharjya, S. (2008) Equilibrium unfolding of the dimeric SAM domain of MAPKKK Ste11 from the budding yeast: Role of the interfacial residues in structural stability and binding. *Biochemistry 47*, 651–659.
21. Zhong, D., Zhang, J., Yang, S., Soh, U. J., Buschdorf, J. P., Zhou, Y. T., Yang, D., and Low, B. C. (2009) The SAM domain of the RhoGAP DLC1 binds EF1A1 to regulate cell migration. *J. Cell Sci. 122*, 414–424.
22. Warren, J. R., and Gordon, J. A. (1966) On the refractive indices of aqueous solutions of urea. *J. Phys. Chem. 70*, 297–300.
23. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol. 131*, 266–280.
24. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277–293.
25. Tanford, C. (1968) Protein denaturation. *Adv. Protein Chem. 23*, 121–282.
26. Schellman, J. A. (1978) Solvent denaturation. *Biopolymers 17*, 1305–1322.
27. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J. 78*, 1606–1619.
28. Schuck, P., and Rossmanith, P. (2000) Determination of the sedimentation coefficient distribution by least-squares boundary modeling. *Biopolymers 54*, 328–341.
29. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 51–55.
30. van Mierlo, C. P., and Steensma, E. (2000) Protein folding and stability investigated by fluorescence, circular dichroism (CD), and nuclear magnetic resonance (NMR) spectroscopy: The flavodoxin story. *J. Biotechnol. 79*, 281–298.
31. Kay, L., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc. 114*, 10663–10665.
32. Wijesinha-Bettoni, R., Dobson, C. M., and Redfield, C. (2001) Comparison of the denaturant-induced unfolding of the bovine and human α-lactalbumin molten globules. *J. Mol. Biol. 312*, 261–273.
33. Schulman, B. A., Kim, P. S., Dobson, C. M., and Redfield, C. (1997) A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat. Struct. Biol. 4*, 630–634.