J|A|C|S ARTICLES

Published on Web 04/08/2010

Analyzing Protein Folding Cooperativity by Differential Scanning Calorimetry and NMR Spectroscopy

Patrick Farber,[†] Hariyanto Darmawan,[†] Tara Sprules,[‡] and Anthony Mittermaier*,[†]

Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6, and Quebec/Eastern Canada High Field NMR Facility, 3420 University Street, Room 023, Montreal, Quebec, Canada H3A 2A7

Received January 29, 2010; E-mail: anthony.mittermaier@mcgill.ca

Abstract: Some marginally stable proteins undergo microsecond time scale folding reactions that involve significant populations of partly ordered forms, making it difficult to discern individual steps in their folding pathways. It has been suggested that many of these proteins fold non-cooperatively, with no significant barriers to separate the energy landscape into distinct thermodynamic states. Here we present an approach for studying the cooperativity of rapid protein folding with a combination of differential scanning calorimetry (DSC), nuclear magnetic resonance (NMR) relaxation dispersion experiments, and an analysis of the temperature dependence of amide ^{1}H and ^{15}N chemical shifts. We applied this method to the PBX homeodomain (PBX-HD), which folds on the microsecond time scale and produces a broad DSC thermogram with an elevated and steeply sloping native-state heat capacity baseline, making it a candidate for barrierless folding. However, by globally fitting the NMR thermal melt and DSC data, and by comparing these results to those obtained from the NMR relaxation dispersion experiments, we show that the native form of the protein undergoes two-state exchange with a small population of the thermally denatured form, well below the melting temperature. This result directly demonstrates the coexistence of distinct folded and unfolded forms and firmly establishes that folding of PBX-HD is cooperative. Further, we see evidence of large-scale structural and dynamical changes within the native state by NMR, which helps to explain the broad and shallow DSC profile. This study illustrates the potential of combining calorimetry with NMR dynamics experiments to dissect mechanisms of protein folding.

Introduction

Protein folding has traditionally been modeled as a chemical kinetic process involving cooperative transitions between discrete states that are separated by energy barriers.[1,2] However, it has recently been proposed that some fast-folding, marginally stable proteins fold in the absence of significant (>kT) energy barriers under all conditions, leading to essentially noncooperative accretion of structure.[3−5] In this model, referred to here as downhill folding, the energy landscape contains a single minimum that gradually shifts from a narrow, native-like well to a broad, unfolded-like basin as the protein is destabilized by temperature or denaturants. The cooperative and downhill descriptions of protein folding are quite different, and it is therefore of great interest to establish experimentally whether specific fast-folding proteins undergo cooperative folding. This has proven to be quite challenging. The downhill model can produce exponential folding kinetics and sigmoidal equilibrium unfolding curves, which are characteristic of cooperative two-state folding.[6,7] Conversely, cooperative transitions among multiple states can produce multiexponential folding kinetics and equilibrium unfolding behavior that clearly deviates from a simple two-state model.[8] Here we show that differential scanning calorimetry (DSC) combined with nuclear magnetic resonance (NMR) relaxation dispersion and thermal melt experiments can establish the existence of folding energy barriers and provide site-specific information on temperature-dependent changes in the structure and dynamics of exchanging states.

NMR thermal melts, which comprise chemical shifts measured as a function of temperature, give an "atom-by-atom" description of the folding process.[5] DSC measurements directly detect the heat absorbed during denaturation.[9] NMR relaxation dispersion experiments quantify rapid motions on the millisecond to microsecond time scale, provide site-specific spectroscopic (chemical shift) information on weakly populated states, and are sensitive to the presence of folding intermediates.[10,11] We used these methods to study the folding reaction of the

[†] McGill University.
[‡] Quebec/Eastern Canada High Field NMR Facility.

(1) Jackson, S. E. *Folding Des.* **1998**, *3*, R81–R91.
(2) Brockwell, D. J.; Radford, S. E. *Curr. Opin. Struct. Biol.* **2007**, *17*, 30–37.
(3) Garcia-Mira, M. M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J. M.; Munoz, V. *Science* **2002**, *298*, 2191–2195.
(4) Naganathan, A. N.; Sanchez-Ruiz, J. M.; Munoz, V. *J. Am. Chem. Soc.* **2005**, *127*, 17970–17971.
(5) Sadqi, M.; Fushman, D.; Munoz, V. *Nature* **2006**, *442*, 317–321.
(6) Hagen, S. J. *Proteins* **2007**, *68*, 205–217.
(7) Naganathan, A. N.; Doshi, U.; Munoz, V. *J. Am. Chem. Soc.* **2007**, *129*, 5673–5682.
(8) Mayor, U.; Guydosh, N. R.; Johnson, C. M.; Grossmann, J. G.; Sato, S.; Jas, G. S.; Freund, S. M. V.; Alonso, D. O. V.; Daggett, V.; Fersht, A. R. *Nature* **2003**, *421*, 863–867.
(9) Freire, E. In *Methods in Enzymology*; Academic Press Inc.: San Diego, CA, 1995; Vol. 259, pp 144−168.
(10) Korzhnev, D. M.; Salvatella, X.; Vendruscolo, M.; Di Nardo, A. A.; Davidson, A. R.; Dobson, C. M.; Kay, L. E. *Nature* **2004**, *430*, 586–590.

 10.1021/ja100815a

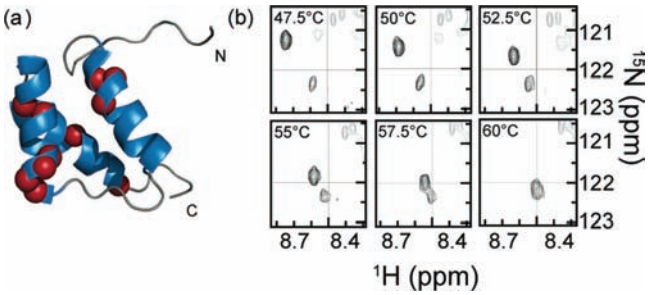

Figure 1. (a) Ribbon structure of PBX-HD.[13] Red spheres indicate the locations of residues with doubled peaks in ^{1}H/^{15}N correlation spectra. All doubled peaks exhibit sigmoidal chemical shift temperature dependences in either one or both dimensions. (b) Regions of PBX-HD ^{1}H/^{15}N correlation spectra collected at temperatures between 47.5 and 60 °C. Major and minor peaks of Leu31 are shown in black, while those of other residues are shown in gray.

homeodomain from the human transcription factor PBX (PBX-HD). Homeodomains are small DNA-binding proteins that are found throughout eukaryotic organisms.[12] The native structure of PBX-HD has been solved by NMR spectroscopy and is typical of homeodomains, comprising three α helices and a well-packed hydrophobic core.[13] PBX-HD folds on the microsecond time scale and undergoes gradual thermal denaturation characteristic of multiple conformational transitions. Using a combination of global fitting and cross-validation of the spectroscopic and calorimetric data sets, we show that PBX-HD folds cooperatively from a denatured state with residual structure to a native state that gradually changes conformation and flexibility as the temperature is raised. The implications for downhill folding are discussed.

Results

NMR Thermal Melts. Two-dimensional ^{15}N/^{1}H correlation spectra were collected for PBX-HD between 10 and 65 °C at intervals of 2.5 °C. The positions of many peaks shift considerably over the temperature range, which is indicative of fast time scale exchange between different local environments of the amide ^{15}N/^{1}H pairs. Each chemical shift corresponds to an ensemble average that gradually changes as the relative populations of the interconverting states vary. Peak trajectories show either linear temperature profiles or sigmoidal transitions near 50 °C. Residues with sigmoidal trajectories are located throughout the protein, as shown in Figure 1a, providing a comprehensive view of the folding transition. The NMR thermal melts contain an unusual feature that, to our knowledge, has not been previously documented. At temperatures slightly below the midpoint of the folding transition, a new set of weak peaks appear in NMR spectra, as shown for Leu31 in Figure 1b. The intensities of the minor peaks increase to approximately 30% of the major peak intensities as the temperature is raised. At the upper end of the thermal melt, each of the sigmoidally shifting major peaks approaches and ultimately overlaps a minor peak, with only one exception as discussed below. We attribute this behavior to cis/trans isomerization of the Tyr28-Pro29 peptide bond. The Tyr28-Pro29 bond is trans in the native protein,[13] while

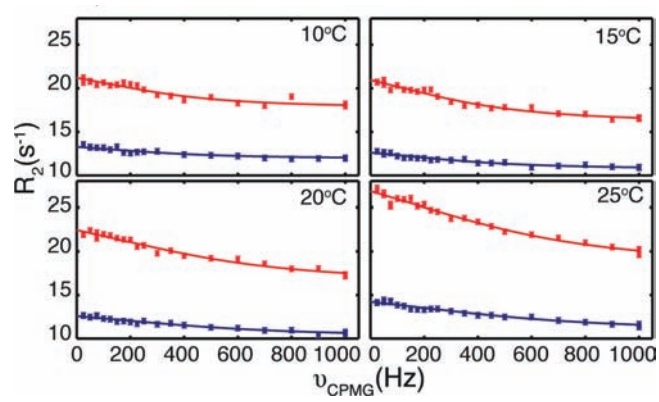

Figure 2. NMR CPMG relaxation dispersion data for the backbone amide ^{15}N of Arg56 in PBX-HD collected at 500 (blue, lower) and 800 (red upper) MHz ^{1}H Larmor frequencies and 10, 15, 20, and 25 °C. Lines indicate best-fit curves obtained with global values of the exchange rate, k_{ex}, according to eq 1.

Tyr-Pro peptide bonds have been found to populate the cis isomer to a level of 21% in studies on small peptides.[14] A similar ratio of cis/trans isomers would be expected for the denatured form of PBX-HD. Cis/trans proline isomerization occurs slowly;[15] thus, the cis-proline denatured state is in slow exchange with the other forms of the protein and produces a distinct set of minor peaks. In contrast, the native and trans-proline denatured states interconvert rapidly and produce a set of major peaks at the population-weighted average chemical shifts. At the highest temperatures, the protein is completely denatured and the major and minor peaks overlap, since the cis- and trans-proline denatured forms have nearly identical chemical shifts. Only the Tyr28 major peak does not ultimately overlap a minor peak, likely because its proximity to Pro29 causes the chemical shifts of this residue to differ in the cis- and trans-proline denatured forms. Slow folding due to proline cis/trans isomerization has also been reported in rapid-mixing and temperature-jump studies of homeodomain-like Myb domains.[16] In what follows, proline isomerization is considered to be separate from the main folding mechanism of PBX-HD, and all of our conclusions apply to the latter process.

NMR Relaxation Dispersion. ^{15}N Carr−Purcell−Meiboom−Gill (CPMG) relaxation dispersion experiments[17] were used to characterize microsecond time scale motions in PBX-HD at 10, 15, 20, and 25 °C, at 11.7 and 18.8 T spectrometer field strengths (500 and 800 MHz). In this method, variable trains of refocusing pulses are applied during a constant relaxation period to suppress spectral broadening due to conformational exchange.[18] Transverse relaxation rates (R_2) were measured as a function of the delay between successive refocusing pulses (τ_{cp}) as illustrated in Figure 2. The magnitudes of the dispersion profiles vary as the square of the spectrometer field strength (2.6-fold larger at 800 MHz that at 500 MHz), which is characteristic of fast time

(11) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021–U11.
(12) Gehring, W. J.; Affolter, M.; Burglin, T. *Annu. Rev. Biochem.* **1994**, *63*, 487–526.
(13) Sprules, T.; Green, N.; Featherstone, M.; Gehring, K. *Biochemistry* **2000**, *39*, 9943–9950.
(14) Wu, W. J.; Raleigh, D. P. *Biopolymers* **1998**, *45*, 381–394.
(15) Grathwohl, C.; Wuthrich, K. *Biopolymers* **1981**, *20*, 2623–2633.
(16) Gianni, S.; Guydosh, N. R.; Khan, F.; Caldas, T. D.; Mayor, U.; White, G. W. N.; DeMarco, M. L.; Daggett, V.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13286–13291.
(17) Hansen, A. F.; Vallurupalli, P.; Kay, L. E. *J. Phys. Chem. B* **2008**, *112*, 5898–5904.
(18) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzymol.* **2001**, *339*, 204–238.

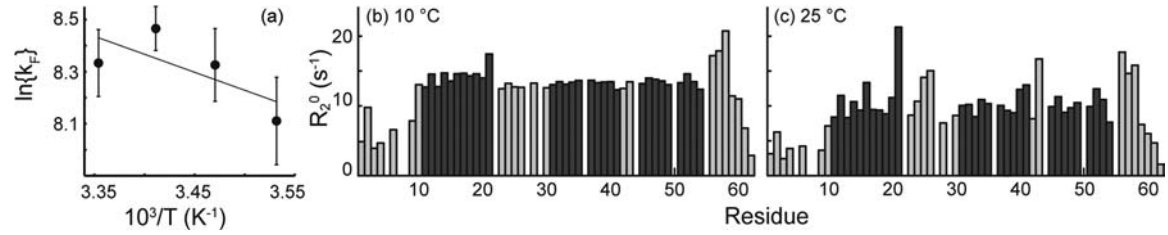

Figure 3. (a) Arrhenius plot of CPMG-derived folding rates, k_F. The slope of the plot is $-E_a/R$, where E_a is the activation energy and R is the ideal gas constant. (b,c) Asymptotic values of transverse relaxation rates at high ν_{CPMG}, R_2^0, obtained at 10 and 25 °C and 18.8 T (800 MHz ^{1}H Larmor frequency). Dark bars indicate the locations of α helices as given in ref 13.

scale exchange.[19] Therefore, an equation describing rapid two-state folding[20] was fitted to the data according to

$$R_2 = R_2^0 + \frac{\Phi_{ex}}{k_{ex}}\left(1 - \frac{4\nu_{CPMG}}{k_{ex}}\right)\tanh\left(\frac{k_{ex}}{4\nu_{CPMG}}\right), \quad \nu_{CPMG} = \frac{1}{2\tau_{cp}} \tag{1}$$

where k_{ex} is the sum of the folding and unfolding rates, $k_{ex} = k_F + k_U$, R_2^0 is the asymptotic value of R_2 at high pulse repetition rates, and the parameter Φ_{ex} is given by $p_N p_D(\Delta\omega)^2$, where p_N and p_D are the relative populations of the native and denatured states and $\Delta\omega$ is the difference in chemical shift between the two forms. Data for all but one residue are compatible with a single, global value of k_{ex} at each temperature, which suggests that folding can be described as a cooperative two-state process. The one outlying residue, Gly44, likely undergoes additional local motions that obscure the global folding reaction. This contrasts sharply with Fyn SH3 domain mutants that undergo multistate folding and produce CPMG data that are incompatible with global two-state models.[10,21] Folding rates were calculated from k_{ex} using values of p_D obtained calorimetrically, as described below. Values of k_F vary from 3300 ± 600 s^{-1} at 10 °C to 4100 ± 500 s^{-1} at 25 °C. An Arrhenius plot is shown in Figure 3a, and yields a folding activation energy of 2.8 kcal mol^{-1}. Essentially identical results are obtained when the CPMG data at all temperatures are analyzed globally with Arrhenius folding kinetics.

When exchange is in the fast time scale regime, as is the case for PBX-HD, the asymptotic values of R_2 at high pulse repetition rates, R_2^0, correspond to the population-weighted average transverse relaxation rates of the native and denatured states.[18,22] We find that the denatured state is populated at less than 2% for the entire CPMG data set; therefore, R_2^0 values largely reflect motions within the native state, with much smaller contributions coming from dynamics in the denatured state. Large-amplitude internal motions on the picosecond time scale lead to small R_2^0 values.[23] Conversely, ultrarapid microsecond time scale motions that are slow enough to cause broadening ($<10^6$ s^{-1}, approximately) but too rapid to be suppressed by the CPMG pulse trains employed ($>10^4$ s^{-1}, approximately) lead to elevated R_2^0 values. Plots of R_2^0 obtained at 18.8 T and at 10 and 25 °C are shown in Figure 3b,c. At 10 °C, the values are fairly homogeneous, which suggests that most positions in PBX-HD experience similar internal motions at that temperature, although the values are slightly larger than would be expected for a protein the size of PBX-HD (see Methods). Lower values, indicating picosecond time scale motions, are obtained near the N-terminus, which is believed to be unstructured[13] and at the extreme C-terminus. Higher values, indicating ultrarapid microsecond time scale dynamics, are obtained immediately after the third helix in a region that becomes ordered upon binding DNA.[13] At 25 °C, these trends are maintained. However, the R_2^0 values are far more heterogeneous, which suggests that the native form of PBX-HD undergoes an increasingly complex array of rapid internal motions as the temperature is raised.

Differential Scanning Calorimetry. DSC data for PBX-HD are plotted in Figure 4a and are well fit by a two-state folding model with linear heat capacity baselines. However, this analysis predicts that the folded and denatured baselines intersect near the midpoint of the transition and therefore that the folded state has increasingly greater heat capacity than the denatured state, which is physically unrealistic (Supporting Information Figure S4). Protein denaturation is expected to increase the solvent exposure of hydrophobic residues, leading to an increase rather than a decrease in heat capacity.[24] The intersection of the linear baselines is related to the heat capacity profile at low temperatures, which lies above and is more steeply sloped than the baseline expected for a typical protein the size of PBX-HD.[25] For example, the average specific heat capacities at 20 °C for PBX and a representative set of nine globular proteins[26] are 0.419 and 0.343(0.014) cal K^{-1} g^{-1}, respectively, while the corresponding values of dC_p/dT are 6.29×10^{-3} and $1.45 \times 10^{-3}(0.29 \times 10^{-3})$ cal K^{-2} g^{-1}, where parentheses indicate the standard deviations. This behavior has been observed for other marginally stable proteins and is believed to result from pre-unfolding conformational transitions (that may or may not be cooperative), such that endothermic structural melting and exposure of hydrophobic surface to the solvent lead to a temperature-dependent elevation of the heat capacity.[3,25,27,28] Recent reports have suggested that it is possible to identify proteins undergoing downhill folding using DSC data alone.[7,29] When this approach is applied to PBX-HD, it predicts that there is no barrier separating the native and denatured states (Supporting Information Figures S2 and S3). However, we find that

(19) Millet, O.; Loria, J. P.; Kroenke, C. D.; Pons, M.; Palmer, A. G. *J. Am. Chem. Soc.* **2000**, *122*, 2867–2877.
(20) Luz, Z.; Meiboom, S. *J. Chem. Phys.* **1963**, *39*, 366–370.
(21) Neudecker, P.; Zarrine-Afsar, A.; Davidson, A. R.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15717–15722.
(22) Ishima, R.; Torchia, D. A. *J. Biomol. NMR* **2006**, *34*, 209–219.
(23) Palmer, A. G.; Williams, J.; McDermott, A. *J. Phys. Chem.* **1996**, *100*, 13293–13310.
(24) Prabhu, N. V.; Sharp, K. A. *Annu. Rev. Phys. Chem.* **2005**, *56*, 521–548.
(25) Crane-Robinson, C.; Read, C. M.; Cary, P. D.; Driscoll, P. C.; Dragan, A. I.; Privalov, P. L. *J. Mol. Biol.* **1998**, *281*, 705–717.
(26) Makhatadze, G. I. *Biophys. Chem.* **1998**, *71*, 133–156.
(27) Dragan, A. I.; Privalov, P. L. *J. Mol. Biol.* **2002**, *321*, 891–908.
(28) Dragan, A. I.; Read, C. M.; Makeyeva, E. N.; Milgotina, E. I.; Crane-Robinson, C.; Privalov, P. L. *Biophys. J.* **2004**, *86*, 623A–623A.
(29) Munoz, V.; Sanchez-Ruiz, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17646–17651.

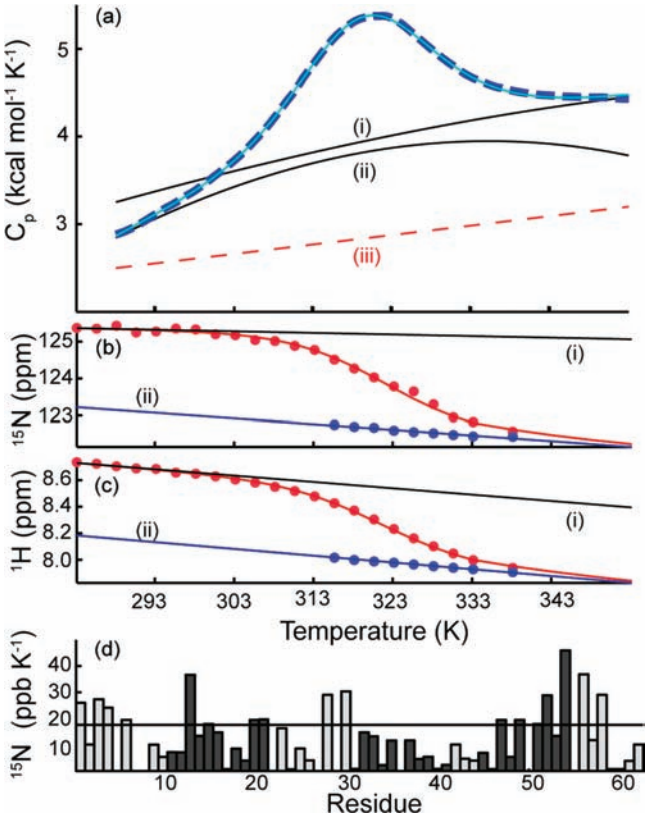

Figure 4. (a) Experimental PBX-HD DSC data (thick dashed line) and fit to a global folding model (thin continuous line). (i), (ii), and (iii) correspond to denatured- and native-state heat capacity baselines and a theoretical baseline calculated using set of representative globular proteins, respectively. (i) and (ii) were extracted directly in the global fit, while (iii) was calculated using $(0.343 + 0.00145(T - 293.15\ \text{K})$ cal mol^{-1} K^{-1} g^{-1}) and the molecular mass of PBX-HD, 7438 g mol^{-1}, as described in the text. (b,c) NMR thermal melt data for Ala38. Red and blue points correspond to the positions of the major and minor peaks, while lines (i) and (ii) are the native- and denatured-state chemical shift baselines extracted from the global fit, respectively. The red curve indicates the population-weighted average of the two baselines. (d) Temperature dependence of native-state ^{15}N chemical shifts. Absolute slopes were obtained by linear regression of data at 10, 12.5, 15, 17.5, and 20 °C and are plotted as a function of residue number. Dark bars indicate the locations of α helices.[13] The horizontal line indicates the maximum ^{15}N slope reported for ubiquitin (17.6 ppb K^{-1}).[30]

this conclusion is not supported by our global analysis of the DSC and NMR data.

Global Fits of DSC and NMR Thermal Melt Data. We analyzed the DSC and NMR thermal melt data simultaneously using a global model of protein folding. As discussed above, the DSC thermogram strongly suggests that PBX-HD undergoes endothermic pre-unfolding changes. Typically, pre-unfolding transitions are modeled with one or more discrete intermediates in the folding pathway.[25,27,28] However in this case, we find that models with a small number of folding intermediates do not provide satisfactory agreement with the combined DSC and NMR data set, suggesting that the number of conformational substates involved may be large, as discussed in the Supporting Information. We have therefore employed a phenomenological two-state model in which all folding intermediates are effectively subsumed into the native- or denatured-state conformational ensembles, whose heat capacity profiles are represented by third-order polynomial functions of temperature. This approach provides nearly perfect agreement with both DSC and NMR data sets, as shown in Figure 4a−c and Supporting Information Figure S1. The analysis yields a denaturation midpoint (at which 50% of the proteins are in the native state) of 46.4 ± 0.3 °C. At this temperature, the differences in enthalpy and entropy between the native and trans-proline denatured states are $\Delta H_{ND} = 35 \pm 1.4$ kcal mol^{-1} and $\Delta S_{ND} = 108 \pm 4$ cal mol^{-1} K^{-1}. Notably, the curvilinear baselines do not intersect, even though no explicit constraints were applied in the calculation to obtain this physically realistic result. The NMR thermal melt data provide some indication of the nature of the pre-unfolding transition. The native-state chemical shifts (δ_{native}) of several residues exhibit strong temperature dependences, indicating that they experience large changes in local environment as the temperature is varied. Notably, $\delta_{native}(^{15}N)$ values of PBX-HD are significantly more sensitive to temperature than are those of ubiquitin, which does not undergo a large pre-unfolding transition.[30,31] The temperature dependences of $\delta_{native}(^{1}H)$ values were not examined in detail, since these tend to show more natural variability, depending on patterns of hydrogen bonding.[32] This work builds on a recent study in which DSC and circular dichroism (CD) spectroscopic data were fitted globally[33] and illustrates the potential of combined calorimetric and spectroscopic analyses. NMR has much to offer in this regard, since it provides many independent probes (nuclei) of the folding process. Furthermore, in the case of PBX-HD, the slowly exchanging cis-proline unfolded form permits direct detection of denatured-state chemical shifts throughout the thermal transition region, facilitating precise analyses of peak positions in terms of native- and denatured-state populations.

Cross-Validation of the Two-State Folding Model. Both the CPMG and the NMR thermal melt data sets are related to the differences in ^{15}N chemical shift between the native and denatured states. In the case of the NMR thermal melts, this relationship is simply the difference between the fitted values of δ_{native} and $\delta_{denatured}$, $\Delta\delta$ (Figure 4c), for each residue. For CPMG experiments, this relationship is contained in the fitted parameter $\Phi_{ex} = p_N p_D (\Delta\omega)^2$ (eq 1), where $\Delta\omega$ is the difference in chemical shift between the native and denatured states. If the broadening in NMR spectra is due to cooperative exchange between the native and denatured states, then plots of Φ_{ex} versus $(\Delta\delta)^2$ should be linear with a slopes of $p_N p_D$. As shown in Figure 5, this is indeed the case at all four temperatures. Residues exhibiting large differences in ^{15}N chemical shift between the native and denatured states in NMR thermal melts also produce large-amplitude relaxation dispersions. Conversely, residues with ^{15}N chemical shifts that are nearly identical in the native and denatured states ($\Delta\delta \approx 0$) have essentially flat dispersion profiles ($\Phi_{ex} \approx 0$, see Methods). Furthermore, the populations of the denatured state (p_D) calculated from the slopes of these plots (0.19 ± 0.04, 0.31 ± 0.07, 0.55 ± 0.07, and 0.61 ± 0.12%) are similar to those extracted from the global DSC and NMR thermal melt analysis (0.13 ± 0.05, 0.30 ± 0.09, 0.70 ± 0.15, 1.6 ± 0.3%) at 10, 15, 20, and 25 °C. This provides strong evidence that the folding of PBX-HD can be described as a cooperative transition between a denatured and native state. Notably, the chemical shifts of the denatured state differ significantly from those predicted for a completely unstructured

(30) Cordier, F.; Grzesiek, S. *J. Mol. Biol.* **2002**, *317*, 739–752.
(31) Wintrode, P. L.; Makhatadze, G. I.; Privalov, P. L. *Proteins* **1994**, *18*, 246–253.
(32) Baxter, N. J.; Williamson, M. P. *J. Biomol. NMR* **1997**, *9*, 359–369.
(33) Streicher, W. W.; Makhatadze, G. I. *Biochemistry* **2007**, *46*, 2876–2880.

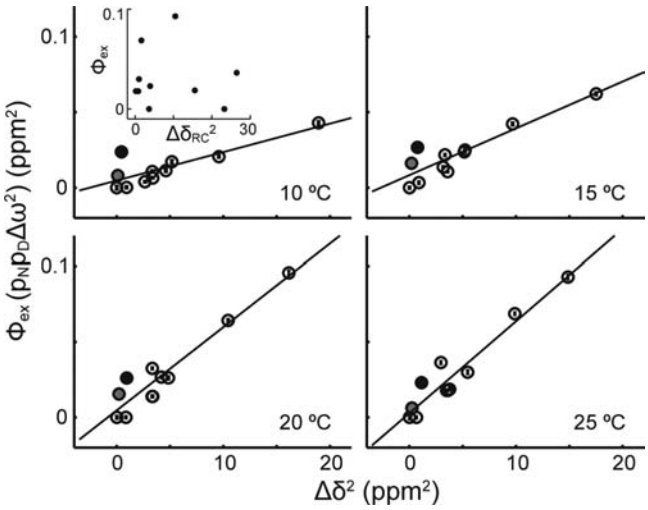

Figure 5. CPMG-derived Φ_{ex} parameters for all residues with doubled peaks, plotted as a function of the squared differences in chemical shift between the native and denatured states, $\Delta\delta^2$, estimated from NMR thermal melts. Lines were obtained by linear regression and have slopes of $p_N p_D$, where p_N and p_D are the relative populations of the native and denatured states. The inset shows Φ_{ex} at 25 °C plotted as a function of squared differences in chemical shift between the native state and those predicted for a random-coil chain, $\Delta\delta_{RC}^2$. Points for Thr13 (backbone) and Trp51 (side chain) are indicated with filled black and gray circles, respectively.

polypeptide chain. Chemical shift differences between the native state and a theoretical random coil[34,35] were calculated, $\Delta\delta_{RC} = \delta_{RC} - \delta_N$. A plot of Φ_{ex} versus $(\Delta\delta_{RC})^2$, shown in the inset to Figure 5, illustrates that there is essentially no correlation between the two parameters. This suggests that the denatured state contains some residual structure, as discussed below. There are some small deviations from the overall good agreement between $\Delta\delta$ and Φ_{ex} values. For example, the ^{15}N resonances of Thr13 (backbone) and Trp51 (side chain), indicated by filled circles in Figure 5, exhibit slightly elevated Φ_{ex} values, which likely reflect additional millisecond to microsecond time scale dynamics within the native state at these sites. This is particularly evident at low temperatures, where there is less broadening due to global folding/unfolding dynamics.

Discussion

Evidence for Cooperative Folding. The DSC thermogram and all peaks in the NMR thermal melt analysis represent independent probes of the equilibrium unfolding of PBX-HD. Therefore, the simultaneous agreement of the global two-state model with the NMR melting curves and the DSC data strongly supports the idea that PBX-HD folds in a cooperative manner. The CPMG data likewise agree with a global two-state model of exchange. However, the key piece of evidence demonstrating that folding is cooperative comes from the cross-validation of the NMR thermal melt and CPMG data. Both analyses produce independent estimates of the chemical shift differences between the native and denatured states on a per-residue basis, in the forms of $\Delta\delta$ and Φ_{ex} for thermal melts and CPMG fits, respectively. We observe excellent agreement between these two sets of parameters, which implies that, between 10 and 25 °C, the native form of the protein undergoes cooperative exchange with a small population (<2%) of the denatured state on the microsecond time scale. This directly contradicts the downhill folding model, for which there is only a single minimum in the energy landscape. It has been proposed that conformational diffusion within this single energy well can produce broadening in NMR spectra.[5] However, the downhill folding model does not suggest that the fluctuations in chemical shift, and hence broadening, should correlate with the spectral properties of the thermally denatured state, as is the case for PBX-HD. In fact, little or no correlation would be expected, since local structures melt independently over a wide range of temperatures, according to this model.[5,36] It must be emphasized that our results do not rule out the existence of intermediates along the PBX-HD folding pathway. In fact, the pre-unfolding transition likely involves a substantial number of conformational substates. However, the data indicate that any intermediates are weakly populated, exchange very rapidly with the native or denatured states, or both. In other words, the energy landscape of PBX-HD is dominated by a single free energy barrier. On one side of the barrier lie the unstructured conformations and any pre-transition-state intermediates that we collectively term the denatured state. On the other side lie the relatively ordered conformations and any post-transition-state intermediates we collectively term the native state. Conformational exchange within the native and denatured ensembles occurs far more rapidly than exchange between ensembles.

One of the great advantages of CPMG dynamics experiments is that exchanging states may be identified on the basis of chemical shift fingerprinting. In a ground-breaking NMR study of SH3 domain folding, CPMG data indicated that the native state interconverts with two weakly populated forms.[10] The chemical shifts of one of these forms correspond to those predicted for a random-coil peptide, providing clear evidence for a three-state protein folding reaction. Comparisons of CPMG-derived $\Delta\omega$ and Φ_{ex} parameters with ligand-dependent chemical shift changes have been used to characterize two-state[37] and multistate[11] protein binding mechanisms and to map complex enzymatic reaction pathways.[38] The work presented here illustrates the potential of combining NMR thermal melts with CPMG data to characterize millisecond to microsecond time scale protein motions, through chemical-shift fingerprinting of thermally denatured forms.

Nature of the Pre-unfolding Transition. The native form of PBX-HD undergoes an endothermic conformational transition over the entire temperature range studied. This conclusion is supported by the large anomalous heat capacity measured for this state by DSC, the unusually large changes in chemical shift observed for some residues as the temperature is raised, and the increasingly heterogeneous pattern of internal motions reported by transverse relaxation rates, R_2^0. Temperature-sensitive native-state ^{15}N chemical shifts and elevated R_2^0 values are concentrated in a region of the protein encompassing the C-terminal end of the third helix. Extension of the third helix upon DNA binding has been observed in PBX-HD[13] and several other homeodomains.[39−41] It is therefore likely that helix/coil dynamics in this region of the protein contribute to the pre-

(34) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. *J. Am. Chem. Soc.* **2001**, *123*, 2970–2978.
(35) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Wright, P. E.; Dyson, H. J. *J. Biomol. NMR* **2000**, *18*, 43–48.
(36) Munoz, V. *Int. J. Quantum Chem.* **2002**, *90*, 1522–1528.
(37) Demers, J. P.; Mittermaier, A. *J. Am. Chem. Soc.* **2009**, *131*, 4355–4367.
(38) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313*, 1638–1642.
(39) Tsao, D. H. H.; Gruschus, J. M.; Wang, L. H.; Nirenberg, M.; Ferretti, J. A. *Biochemistry* **1994**, *33*, 15053–15060.

unfolding transition detected by NMR and DSC. However, conformational and dynamical changes are not restricted to this region, as elevated R_2^0 values and large ^{15}N chemical shift temperature dependences are observed throughout the protein at 25 °C.

Nature of the Folding Energy Barrier. The PBX-HD folding rate is quite insensitive to temperature, with an apparent activation energy of about 3 kcal/mol. In contrast, enthalpic folding barriers for more slowly folding proteins range from about 15 to 40 kcal/mol.[42] Enthalpic barriers to protein folding are thought to relate to the desolvation of nonpolar side chains prior to hydrophobic core formation.[42] The low activation energy obtained for PBX-HD could therefore imply that nonpolar groups are only slightly more exposed to solvent in the denatured state than in the folding transition state. This is consistent with our observation that the denatured form contains residual structure and that the native form undergoes partial melting. If hydrophobic side chains cluster in the denatured state and are somewhat solvated in the native state, then only minimal desolvation would be required at the rate-limiting step of folding. Since the folding activation energy is close to zero, the energy barrier itself is primarily entropic. This agrees with statistical mechanical descriptions of protein folding in which an entropic "bottleneck" occurs at the point along the reaction coordinate where the decrease in the number of accessible conformations is incompletely compensated by the formation of favorable interactions.[43]

Discriminating between Downhill and Cooperative Folding. It can be difficult to establish whether the energy landscapes of fast-folding proteins contain significant barriers, since the time scale of conformational interconversion is short compared to the time scales of the experimental techniques used to monitor folding.[44] Several different strategies have been developed to overcome this problem. An analysis was recently proposed in which an enthalpy probability density function is fitted directly to DSC data.[4,29] We find that when this method is applied to PBX-HD, the extracted enthalpy distribution is unimodal; i.e., downhill folding is predicted. However, our results clearly show that this protein folds cooperatively over a free energy barrier. Similarly, this approach predicts that the Engrailed homeodomain, En-HD, undergoes downhill folding,[4] although extensive research has shown that it folds in two cooperative steps.[8,16,45,46] It appears that this method has a tendency to assign the downhill mechanism to proteins that fold cooperatively.

Other approaches involve examining whether different structural probes provide coincident denaturation profiles[3,5] or characterizing the coupling between different denaturation methods, such as temperature and chaotropic agents.[47] However, these tests are based on experiments performed at thermodynamic equilibrium. They probe the structural and energetic heterogeneity of the protein conformational ensemble, but they cannot directly address whether conformations are separated by energy barriers. Even kinetic measurements do not always provide a clear indication of folding cooperativity. Although simple two-state folding gives rise to monoexponential, probe-independent folding kinetics and downhill folding can produce stretched-exponential, probe-dependent kinetics, it has been shown that these criteria are insufficient to conclusively demonstrate the presence or absence of barriers in the energy landscape.[6] Far more detailed analyses of the effects of mutation, temperature, and solvent conditions on folding kinetics are required to draw conclusions regarding the underlying mechanisms.[46,48,49] Chevron plot analysis has been used to analyze folding cooperativity, since downhill folding is unlikely to produce kinetics with the strong linear dependence on denaturant concentration that is characteristic of two-state systems.[44] Native-state hydrogen exchange also allows two-state folding behavior to be distinguished in fast-folding proteins.[50] However, cooperativity would be more difficult to infer when the exchanging states are structurally labile, as is the case for PBX-HD. Our work represents a robust new approach for analyzing folding cooperativity that can help to shed light on unresolved questions in protein folding.

Conclusion

Elucidating the folding mechanisms of small, marginally stable proteins is of great interest because it tests our fundamental understanding of protein biophysics. We have used a combination of DSC, NMR thermal melt, and CPMG experiments to characterize the folding mechanism of the homeodomain from the human transcription factor PBX with atomic resolution. This protein folds rapidly, with $k_F = 4090\ s^{-1}$ at 25 °C, and exhibits a broad thermal denaturation profile that makes it a candidate for multistate or downhill folding mechanisms. However, the analysis shows that PBX-HD folds over a single, entropic energy barrier from a denatured state that contains residual structure to a native state that undergoes an endothermic conformational rearrangement. Discriminating between different folding models can be challenging for marginally stable, fast-folding proteins. As recently noted, "it is necessary to assign observed rate constants to specific processes" in order to conclusively identify folding mechanisms.[46] The global analysis of DSC, NMR thermal melt, and CPMG experiments provides a new avenue for this to be achieved.

Methods

Sample Preparation. Samples of PBX-HD comprising residues 233−294 of the human pre-B cell transcription factor[51] were expressed in *Escherichia coli* (BL21 DE3) harboring the Rosetta plasmid (Novagen) grown in minimal media.[52] The protein was purified using SP-sepharose ion-exchange and size-exclusion chromatography. NMR samples contained 1.0 mM ^{15}N-enriched protein, 10% ^2H_2O 1 mM EDTA, 10 μM DSS, and 20 mM sodium phosphate at pH 6.0. Control experiments were performed on 1.0 and 0.5 mM protein samples with 20 mM 2-(4-morpholino)ethanesulfonic acid (MES) at pH 6.0 in place of sodium phosphate. DSC samples contained between 40 and 140 μM protein and 20 mM sodium phosphate at pH 6.0.

(40) Cox, M.; Vantilborg, P. J. A.; Delaat, W.; Boelens, R.; Vanleeuwen, H. C.; Vandervliet, P. C.; Kaptein, R. *J. Biomol. NMR* **1995**, *6*, 23–32.
(41) Carra, J. H.; Privalov, P. L. *Biochemistry* **1997**, *36*, 526–535.
(42) Liu, Z. R.; Chan, H. S. *J. Mol. Biol.* **2005**, *349*, 872–889.
(43) Abkevich, V. I.; Gutin, A. M.; Shakhnovich, E. I. *J. Chem. Phys.* **1994**, *101*, 6052–6062.
(44) Huang, F.; Sato, S.; Sharpe, T. D.; Ying, L. M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 123–127.
(45) Mayor, U.; Johnson, C. M.; Daggett, V.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13518–13522.
(46) Religa, T. L.; Johnson, C. M.; Vu, D. M.; Brewer, S. H.; Dyer, R. B.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9272–9277.
(47) Oliva, F. Y.; Munoz, V. *J. Am. Chem. Soc.* **2004**, *126*, 8596–8597.
(48) Gruebele, M. *Proteins* **2008**, *70*, 1099–1102.
(49) Gruebele, M. *C. R. Biol.* **2005**, *328*, 701–712.
(50) Meisner, W. K.; Sosnick, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15639–15644.
(51) Green, N. C.; Rambaldi, I.; Teakles, J.; Featherstone, M. S. *J. Biol. Chem.* **1998**, *273*, 13273–13279.
(52) Marley, J.; Lu, M.; Bracken, C. *J. Biomol. NMR* **2001**, *20*, 71–75.

NMR CPMG Relaxation Dispersion. ^{1}H-decoupled ^{15}N CPMG experiments[17] with power compensation schemes[37] were performed at 10, 15, 20, and 25 °C on Varian INOVA spectrometers equipped with cold-probes operating at 500 and 800 MHz ^{1}H Larmor frequencies. The experiments employed 13−15 kHz ^{1}H decoupling fields during the relaxation delay. In the reference experiment, which lacks the relaxation delay, compensatory ^{1}H decoupling was applied following the acquisition period.[17] Sample heating was assessed by monitoring the shift of the water resonance produced by several minutes of irradiation with the ^{1}H-decoupled ^{15}N CPMG pulse sequence, performed without engaging the D_2O lock. Shifts were typically very small, on the order of 1−5 ppb, which corresponds to a rise in temperature of about 0.1−0.5 °C.[53] No change in the water resonance frequency was obtained with the ^{1}H decoupling power reduced to zero. This slight heating was compensated by lowering the temperature setting on the spectrometer by a corresponding amount. Although the decoupling power can vary by about 10% among experiments with different ν_{CPMG} values,[17] we expect that this has almost no impact on the CPMG data, since spectra with different ν_{CPMG} values are collected in an interleaved fashion and overall heat effects are negligible. Peak intensities were quantified, and transverse relaxation rates, R_2, and their associated uncertainties, σ_{R_2}, were calculated as described previously.[37]

Global folding/unfolding exchange rates, k_{ex}^{global}, and per-residue Φ_{ex} parameters were calculated using a two-step protocol. In the first step, k_{ex}^{global} was estimated using only data for residues with dispersion magnitudes greater than 1 s^{-1} (9, 9, 11, and 16 residues at 10, 15, 20, and 25 °C, respectively). A grid-search of k_{ex} was performed for each residue at each temperature, fixing k_{ex} at values between 100 and 10 000 s^{-1} in increments of 10 s^{-1}, and adjusting all other parameters to minimize the χ^2 function

$$\chi^2(k_{ex}) = \sum \frac{(R_2^{exp} - R_2^{calc}(k_{ex}))^2}{\sigma_{R_2}^2} \quad (2)$$

where $R_2^{calc}(k_{ex})$ is computed using eq 1 and the sum runs over all R_2 data points for a single residue at a single temperature. Every considered residue but one (Gly44) exhibited a $\chi^2(k_{ex})$ minimum at a similar value of k_{ex}, suggesting that dynamics are collective. Aggregate χ^2 versus k_{ex} profiles were then calculated according to

$$\chi_{total}^2(k_{ex}) = \sum \chi^2(k_{ex}) \quad (3)$$

where the sum runs over data for all selected residues at a single temperature (as described above, with the exception of Gly44). The global value of k_{ex} was then selected such that

$$\chi_{total}^2(k_{ex}^{global}) = \min(\chi_{total}^2(k_{ex})) \quad (4)$$

Uncertainties in k_{ex}^{global} values were estimated using a Monte Carlo approach in which randomized $\chi_{total}^2(k_{ex})$ profiles were calculated by summing $\chi^2(k_{ex})$ profiles for subsets of residues selected randomly according to a bootstrap procedure.[54] The standard deviations of the k_{ex} estimates thus obtained were taken as the uncertainties in global k_{ex} values.

In the second step, values of Φ_{ex} were obtained for all residues by minimizing eq 2, fixing $k_{ex} = k_{ex}^{global}$. This global analysis allows Φ_{ex} values to be determined for every residue, including those with small or nonexistent dispersions. This contrasts with fits on a per-residue basis, which require sizable dispersion magnitudes in order for Φ_{ex} to be accurately measured. For example, a flat dispersion profile implies either that k_{ex} is very slow (<1 s^{-1}) or very fast (>10^4 s^{-1}) or that $\Phi_{ex} \approx 0$. Analysis of a single flat dispersion profile cannot distinguish among these three possibilities. However, the exchange rate is fixed at k_{ex}^{global} in the global fits; therefore, nuclei with flat dispersions can be confidently assigned $\Phi_{ex} \approx 0$.

The expected values of R_2^0 for a globular, 61-residue protein were estimated on the basis of a simple model[53] in which the molecule is assumed to be spherical with a specific volume of 0.73 cm^3 g^{-1} and a hydration layer of between 1.6 and 3.2 Å. Rotational diffusion in solution was assumed to follow the Stokes−Einstein−Debye relationship, where the viscosity of water at 10 °C is 1.3 cP.[55] The transverse relaxation rate was calculated according to the Lipari−Szabo model-free formalism[56] with an order parameter, S^2, of 0.9. At 10 °C, the expected R_2^0 values (roughly 7−9 s^{-1}) are lower than those observed experimentally (12−15 s^{-1}). This may be due, in part, to transient intermolecular interactions that slightly elevate the effective rotational correlation time and lead to enhanced transverse relaxation rates throughout the molecule.[57] Reducing the protein concentration by a factor of 2 uniformly decreases R_2^0 values by about 1 s^{-1} but produces no other changes in the shapes or magnitudes of the dispersions. Values of R_2^0 are about 1 s^{-1} greater in 20 mM phosphate buffer than in 20 mM MES. This may be due to the greater ability of $H_2PO_4^-$ and HPO_4^{2-} anions to shield electrostatic repulsion between positively charged molecules of PBX-HD; the ionic strength of 20 mM sodium phosphate is 22 mM, while that of 20 mM MES is 8 mM, at pH 6.0.

NMR Thermal Melts. ^{15}N/^{1}H HSQC spectra[58] were collected every 2.5 °C from 10 to 60 °C and at 65 °C and referenced to an internal DSS standard.[59] Of the 58 peaks present at 10 °C, 47 disappeared at intermediate temperatures, likely due to broadening caused by ultrarapid microsecond time scale motions in the native and/or denatured states and folding/unfolding dynamics. Of the 22 ^{15}N and ^{1}H trajectories that could be tracked over the entire thermal melt (T13, E14, L16, E31, E32, A33, K34, A38, K39, W51, I57), 21 were clearly sigmoidal, and one was approximately linear (T13 ^{1}H). ^{15}N and ^{1}H chemical shifts, δ, and uncertainties, σ_δ, were determined using the NmrPipe/NmrDraw suite of programs.[60] As discussed below, the first stage of global fitting employed averaged ^{15}N and ^{1}H trajectories in order to simplify the calculations. Denaturation is associated with an increase in chemical shift values for some residues and a decrease in chemical shift for others. In order to avoid cancellation of positive and negative $\Delta\delta$ values in the average, trajectories were adjusted to make all $\Delta\delta$ values positive; trajectories showing a decrease in chemical shift upon denaturation were reflected about the mean value according to

$$\delta_i(T) = 2\overline{\delta_i'(T)} - \delta_i'(T) \quad (5)$$

where δ_i' refers to raw ^{1}H or ^{15}N chemical shifts for residue i, and the overbar indicates the average over all temperatures. Error-weighted average chemical shifts and associated uncertainties were calculated according to the expressions,

(53) Cavanagh, J.; Fairbrother, W.; Palmer, A.; Skelton, N. *Protein NMR spectroscopy principles and practice*; Academic Press, Inc.: San Diego, CA, 1996.
(54) Efron, B.; Tibshirani, R. *Stat. Sci.* **1986**, *1*, 54–77.
(55) Cho, C. H.; Urquidi, J.; Singh, S.; Robinson, G. W. *J. Phys. Chem. B* **1999**, *103*, 1991–1994.
(56) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.
(57) Fushman, D.; Cahill, S.; Cowburn, D. *J. Mol. Biol.* **1997**, *266*, 173–194.
(58) Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663–10665.
(59) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135–40.
(60) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.

$$\overline{\delta(T)} = \frac{\sum_{i=1}^{n} \delta_i(T)(\sigma_\delta(T)_i)^{-2}}{\sum_{i=1}^{n} (\sigma_\delta(T)_i)^{-2}} \quad (6)$$

$$\sigma_{\overline{\delta}}(T) = \frac{1}{\sqrt{\sum_{i=1}^{n} (\sigma_\delta)_i^{-2}}} \quad (7)$$

where the sum includes all sigmoidal trajectories that extend over the entire temperature range, $n = 10$ and 11 for ^{1}H and ^{15}N trajectories, respectively.

DSC Analysis. DSC experiments were carried out with an N-DSC III calorimeter from Calorimetry Sciences Corp. (Linden, UT) at a scan rate of 1.0 K/min and a pressure of 4 atm. A total of 46 scans were recorded using four different protein preparations. Thermal unfolding was completely reversible, and thermograms were unaffected by protein concentration over a range of 40−140 μM, indicating that the broad DSC unfolding traces are not influenced by protein aggregation.

Baseline-subtracted raw DSC data, $\Phi_i(T)$ (μW measured in 0.1 °C increments), were averaged over all repeated measurements according to

$$\overline{\Phi(T)} = \frac{\sum_{i=1}^{n} \Phi_i(T)}{n} \quad (8)$$

where the sum runs over all scans, $n = 46$. In order to estimate the standard errors of $\overline{\Phi(T)}$, the individual scans were normalized for protein concentration and corrected for slight vertical offsets inherent to the instrument,[25] yielding $\Phi_i^{\text{norm}}(T)$. Uncertainties were calculated as the standard deviation of $\Phi_i^{\text{norm}}(T)$ divided by $\sqrt{n}$ according to the expression

$$\sigma_\Phi(T) = \sqrt{\frac{\sum_{i=1}^{n} (\Phi_i^{\text{norm}}(T) - \overline{\Phi(T)})^2}{n(n-1)}} \quad (9)$$

Molar heat capacity values were calculated as[61]

$$C_p(T) = \left(\frac{\overline{\Phi(T)}\, dt/dT}{\overline{[\text{protein}]} V_{\text{cell}}} + C_p(T)_{\text{solv}}\, \text{PSV}\, \rho_{\text{solv}} \right) M \quad (10)$$

where dt/dT is the inverse scan rate (60 s K^{-1}), $\overline{[\text{protein}]}$ is the concentration of PBX-HD averaged across all samples (g/mL), V_{cell} is the active volume of the calorimeter cell, 0.300 mL, $C_p(T)_{\text{solv}}$ is the heat capacity of water, 1 cal g^{-1} K^{-1}, PSV is the partial specific volume of the protein, 0.726 mL g^{-1},[62] ρ_{solv} is the density of the solvent, 1 g mL^{-1}, and M is the molecular weight of PBX-HD, 7438 g mol^{-1}. Uncertainties in heat capacity values are thus

$$\sigma_{CP}(T) = \sigma_\Phi(T) \frac{dt}{dT} \overline{[\text{protein}]}^{-1} V_{\text{cell}}^{-1} M \quad (11)$$

For PBX-HD, we find that values of σ_{CP} are on the order of 10 cal mol^{-1} K^{-1}, or about 0.2%.

DSC traces were fit using standard equations for two-state and multistate folding,[25] as well as a variable-barrier model appropriate for both cooperative and downhill folding mechanisms, as described in the Supporting Information.[29] In order to compare the native DSC baseline for PBX-HD to those of typical globular proteins, the average specific heat capacity and corresponding temperature dependence were calculated for barnase, bovine pancreatic trypsin inhibitor, cytchrome *c*, lysozyme, myoglobin, ribonuclease A, Tendamistat, SH3 domain, and ubiquitin, using tabulated C_p values at 5, 25, and 40 or 50 °C.[26]

Global Fitting. Global fitting of the DSC and NMR thermal melt data was performed using a two-step protocol. In the first step, the sigmoidal chemical shift trajectories were averaged, yielding one ^{15}N and one ^{1}H thermal melting profile, in order to simplify the following calculations. The NMR and DSC data sets were then fitted simultaneously by varying 18 adjustable parameters comprising the ΔH and ΔS of unfolding at a reference temperature, 8 parameters defining third-order polynomial native and denatured heat capacity baselines, and 8 parameters defining linear temperature dependences for native and denatured ^{15}N and ^{1}H chemical shifts. In the second step, the trajectory for each peak was analyzed separately using the thermodynamic parameters extracted in step 1, in order to obtain native and denatured ^{15}N chemical shifts as linear functions of temperature on a per-residue basis. In both steps, the positions of the minor peaks were fitted to the linear denatured-state chemical shift profiles, while the positions of the major peaks were fitted to the population-weighted averages of the native- and denatured-state chemical shifts. The analysis was first performed using the averaged DSC thermogram, given by eq 10. It was then repeated for each of traces individually. Reported errors are the standard deviations of the resultant 46 sets of fitted values (in the case of p_D values, three outlying traces were excluded from the standard deviation calculation).

Global fitting of DSC and NMR thermal melt data must account for the fact that PBX-HD contains two proline residues (Pro27 and Pro29), both of which adopt the trans-isomer in the native state[13] and are free to adopt the cis configuration in the denatured state.[14] Thermal unfolding therefore involves five states: trans/trans native and trans/trans, cis/trans, trans/cis, and cis/cis denatured states. However, studies of short peptides suggest that the Tyr28-Pro29 peptide bond populates the cis-isomer to a much greater extent than does the Asp26-Pro27 bond; therefore, folding is predominantly a three-state reaction involving the native and the trans/trans and trans/cis denatured forms. Proline cis−trans isomerization occurs slowly on the NMR chemical shift time scale, and the major peaks in NMR spectra are due to rapid exchange between the native state and only the trans/trans denatured state. The chemical shifts of the major peaks are thus given by

$$\delta_{\text{calc}}^{^{15}\text{N},^1\text{H}} = \frac{1}{1 + K_{\text{tt}}} \delta_{\text{N}}^{^{15}\text{N},^1\text{H}} + \frac{K_{\text{tt}}}{1 + K_{\text{tt}}} \delta_{\text{D}}^{^{15}\text{N},^1\text{H}} \quad (12)$$

where K_{tt} is the equilibrium constant for the native and trans/trans denatured states. The native, $\delta_{\text{N}}^{^{15}\text{N},^1\text{H}}$, and denatured, $\delta_{\text{D}}^{^{15}\text{N},^1\text{H}}$, chemical shift profiles were modeled as straight lines,

$$\delta_{\text{N}}^{^{15}\text{N}}(T) = A_{\text{N0}}^{^{15}\text{N}} + A_{\text{N1}}^{^{15}\text{N}}\, T$$

$$\delta_{\text{N}}^{^{1}\text{H}}(T) = A_{\text{N0}}^{^{1}\text{H}} + A_{\text{N1}}^{^{1}\text{H}}\, T$$

$$\delta_{\text{D}}^{^{15}\text{N}}(T) = A_{\text{D0}}^{^{15}\text{N}} + A_{\text{D1}}^{^{15}\text{N}}\, T$$

$$\delta_{\text{D}}^{^{1}\text{H}}(T) = A_{\text{D0}}^{^{1}\text{H}} + A_{\text{D1}}^{^{1}\text{H}}\, T \quad (13)$$

The equilibrium constant K_{tt} is given by

$$K_{\text{tt}} = \exp\left\{ \frac{-\Delta H_{\text{tt}} + T\Delta S_{\text{tt}}}{RT} \right\} \quad (14)$$

(61) Privalov, G.; Kavina, V.; Freire, E.; Privalov, P. L. *Anal. Biochem.* **1995**, *232*, 79–85.
(62) Hackel, M.; Hinz, H. J.; Hedwig, G. R. *Biophys. Chem.* **1999**, *82*, 35–50.

$$\Delta H_{tt}(T) = \Delta H_0 + \int_{T_0}^{T} dT\, \Delta C_p \quad (15)$$

$$\Delta S_{tt}(T) = \Delta S_0 + \int_{T_0}^{T} dT \frac{\Delta C_p}{T} \quad (16)$$

where T_0 is an arbitrary reference temperature, 318 K in this case, and ΔH_{tt}, ΔS_{tt}, and ΔC_p are the molar differences in enthalpy, entropy, and heat capacity between the trans/trans denatured and native states, respectively. The heat capacities of the native and denatured states were modeled as third-order polynomials, such that

$$C_p^{N} = B_{N0} + B_{N1}T + B_{N2}T^2 + B_{N3}T^3$$

$$C_p^{D} = B_{D0} + B_{D1}T + B_{D2}T^2 + B_{D3}T^3 \quad (17)$$

$$\Delta C_p = C_p^{D} - C_p^{N}$$

The populations of the cis-proline isoforms do not affect the positions of the major peaks in the NMR thermal melts, since these depend only on the relative populations of the rapidly exchanging native and trans/trans denatured forms. However, DSC reflects the thermodynamics of both fast and slow processes; therefore, cis-proline denatured forms must be considered. Experiments performed at scanning rates of 0.5 and 1.0 K/min give essentially superimposable thermograms; therefore, we are confident that all forms of the protein are close to their equilibrium populations throughout the DSC experiment. In order to combine NMR and DSC data, the global analysis must correct for the fact that estimates of the unfolding equilibrium constant obtained by NMR will appear lower than those obtained by DSC, since $K_{NMR} = [D_{tt}]/[N]$, while $K_{DSC} = ([D_{tt}] + [D_{tc}] + [D_{ct}] + [D_{cc}])/[N]$ (using the nomenclature of eq 18, below, to indicate trans/cis isomers). We incorporated proline isomerization into the folding model by including information from thermodynamic studies of short peptides.[14] The cis isoform of Tyr-Pro peptide bonds is populated to a level of about 21%, and is 0.81 kcal mol^{-1} higher in Gibbs free energy and 0.65 kcal mol^{-1} higher in enthalpy than the trans configuration at 298 K.[14] The difference in entropy between the two states is therefore -0.54 cal mol^{-1} K^{-1}. We have modeled isomerization of the Asp26-Pro27 peptide bond on the basis of results for Ala-Pro bonds in short peptides.[14] In this case, the cis configuration is populated to a level of about 7.5%, and is thus 1.5 kcal mol^{-1} higher in Gibbs free energy than the trans form. Stabilization of aromatic proline cis-peptide bonds is thought to be primarily enthalpic;[14] therefore, we used the same entropy value for isomerization of both Tyr28-Pro29 and Asp26-Pro27. This leads to an enthalpy change of 1.33 kcal mol^{-1} for trans/cis isomerization of the Asp26-Pro27 bond. Finally, we assumed that the heat capacities of all denatured states are equal and that cis/trans isomerization is independent at the two sites. Relative to the native state, the enthalpies and entropies of the cis-proline isoforms are therefore

$$\Delta H_{tc} = \Delta H_{tt} + 0.65 \text{ kcal mol}^{-1}$$

$$\Delta S_{tc} = \Delta S_{tt} - 0.54 \text{ cal mol}^{-1} \text{ K}^{-1}$$

$$\Delta H_{ct} = \Delta H_{tt} + 1.33 \text{ kcal mol}^{-1}$$

$$\Delta S_{ct} = \Delta S_{tt} - 0.54 \text{ cal mol}^{-1} \text{ K}^{-1}$$

$$\Delta H_{cc} = \Delta H_{tc} + 1.33 \text{ kcal mol}^{-1}$$

$$\Delta S_{cc} = \Delta S_{tc} - 0.54 \text{ cal mol}^{-1} \text{ K}^{-1} \quad (18)$$

where the first and second letters of the subscript indicate the configurations of the Asp26-Pro27 and Tyr28-Pro29 peptide bond, respectively. The heat capacity profile is then calculated as

$$C_p^{calc} = C_p^{N} + \sum_{i=tt,ct,tc,cc} \frac{\partial P_i}{\partial T} \Delta H_i + \sum_{i=tt,ct,tc,cc} P_i \Delta C_p \quad (19)$$

where P_i is the relative population of state i, given by

$$P_i = \frac{K_i}{1 + \sum_{j=tt,ct,tc,cc} K_j} \quad (20)$$

$$K_i = \exp\left\{\frac{-\Delta H_i + T\Delta S_i}{RT}\right\} \quad (21)$$

The temperature derivatives of the populations are given by the expression

$$\frac{\partial P_i}{\partial T} = \frac{P_i}{RT^2}\left(\Delta H_i - \sum_{j=tt,ct,tc,cc} P_j \Delta H_j\right) \quad (22)$$

Thus, the parameters B_{N0}, B_{N1}, B_{N2}, B_{N3}, B_{D0}, B_{D1}, B_{D2}, and B_{D3} specify the heat capacities of the native and all denatured states via eq 17. These together with ΔH_0 and ΔS_0 define the enthalpy and entropy of the trans/trans denatured state, eqs 15 and 16, as well as those of the trans/cis, cis/trans, and cis/cis states, eq 18. This permits calculation of the DSC trace using eqs 19−22. Defining linear chemical shift baselines with A_{N0}^{15N}, A_{N1}^{15N}, A_{N0}^{1H}, A_{N1}^{1H}, A_{D0}^{15N}, A_{D1}^{15N}, A_{D0}^{1H}, and A_{D1}^{1H}, NMR thermal melt ^{15}N and ^{1}H trajectories may be calculated using eqs 12−14. The parameters $A_i{}^j$, B_i, ΔH_0, and ΔS_0 were varied to minimize the χ^2 function,

$$\chi^2 = \sum_T \frac{(C_p(T) - C_p^{calc}(T))^2}{(\sigma_{CP}(T))^2} + \sum_{T,major} \frac{\overline{(\delta^{15N}(T) - \delta_{calc}^{15N}(T))^2}}{(\sigma_\delta^{15N}(T))^2} + \sum_{T,major} \frac{\overline{(\delta^{1H}(T) - \delta_{calc}^{1H}(T))^2}}{(\sigma_\delta^{1H}(T))^2} + \sum_{T,minor} \frac{\overline{(\delta^{15N}(T) - \delta_{D}^{15N}(T))^2}}{(\sigma_\delta^{15N}(T))^2} + \sum_{T,minor} \frac{\overline{(\delta^{1H}(T) - \delta_{D}^{1H}(T))^2}}{(\sigma_\delta^{1H}(T))^2} \quad (23)$$

such that the positions of major peaks were fit to $\delta_{calc}^{15N,1H}(T)$ and the positions of the minor peaks were fit to $\delta_{D}^{15N,1H}(T)$. The fit had 166 degrees of freedom and produced a residual χ^2 value of 931, suggesting that the experimental uncertainties, eqs 11 and 7, were underestimated by about a factor of 2. In order to obtain native and denatured ^{15}N chemical shifts for each residue and to verify that all residues obey the same two-state mechanism, we used values of K_{tt} calculated with eqs 14−17 and globally optimized unfolding parameters to fit chemical shift values on a per-residue basis, according to eq 12. Nearly perfect agreement was obtained for all trajectories, as shown in Supporting Information Figure S1.

Acknowledgment. This research was supported by the National Science and Engineering Research Council (Canada). We thank Kalle Gehring for the gift of the PBX-HD expression plasmid, Lewis Kay for providing NMR pulse sequences, and both Lewis Kay and Julie Forman-Kay for critical reading of the manuscript.

Supporting Information Available: Plots of NMR thermal melt data for all residues included in the global analysis; fits of DSC data to a variable-barrier folding model and to a two-state model with linear heat capacity baselines; analysis of curvilinear heat capacity baselines. This material is available free of charge via the Internet at http://pubs.acs.org.

JA100815A