Calorimetric study of a series of designed repeat proteins: Modular structure and modular folding

Aitziber L. Cortajarena[1,2]* and Lynne Regan[1,3]

[1]Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, Connecticut 06520
[2]IMDEA-Nanociencia and Centro Nacional de Biotecnología, Universidad Autonoma de Madrid, Cantoblanco, Madrid 28049, Spain
[3]Department of Chemistry, Yale University, New Haven, Connecticut 06520

Received 26 September 2010; Revised 12 November 2010; Accepted 15 November 2010
DOI: 10.1002/pro.564
Published online 3 December 2010 proteinscience.org

Abstract: Repeat proteins comprise tandem arrays of a small structural motif. Their structure is defined and stabilized by interactions between residues that are close in the primary sequence. Several studies have investigated whether their structural modularity translates into modular thermodynamic properties. Tetratricopeptide repeat proteins (TPRs) are a class in which the repeated unit is a 34 amino acid helix-turn-helix motif. In this work, we use differential scanning calorimetry (DSC) to study the equilibrium stability of a series of TPR proteins with different numbers of an identical consensus repeat, from 2 to 20, CTPRa2 to CTPRa20. The DSC data provides direct evidence that the folding/unfolding transition of CTPR proteins does not fit a two-state folding model. Our results confirm and expand earlier studies on TPR proteins, which showed that apparent two-state unfolding curves are better fit by linear statistical mechanics models: 1D Ising models in which each repeat is treated as an independent folding unit.

Keywords: repeat protein; tetratricopeptide repeat (TPR); protein design; differential scanning calorimetry (DSC); protein folding; Ising model; thermodynamic analysis; thermal denaturation

Introduction

Repeat proteins contain tandem arrays of a small structural motif, which differ in length (typically 18-47 amino acids) and in structure (alpha, beta, or alpha/beta) depending on the particular family of repeat protein. Their architecture is sustained by contacts between residues close in the primary sequence, with no long-range interactions. Because of their regular modular structure, repeat proteins represent a simple model system in which to study the physical basis of protein folding and stability.[1–3] As an exemplar of this class, we focus on the tetratricopeptide repeat (TPR) protein family,[4,5] in which the repeated unit is a 34 amino acid sequence that adopts a helix-turn-helix structure (Fig. 1).[6,7] We have previously described the construction and characterization of a set of TPR proteins comprised of different numbers of tandem repeats of an identical consensus TPR unit—CTPRa2 to CTPRa20.[8–10] We have studied the folding and stability of these CTPR proteins using a variety of biophysical methods.[8,9,11,12] Individual equilibrium chemical or thermal denaturation curves display a cooperative single transition, indistinguishable from the characteristic denaturation curves of small globular proteins, which exhibit a two-state folding/unfolding transition.[8,9,11] However, the denaturation curves of a family of proteins of different lengths are better described, and the experimental data are well-fit, by a 1D-Ising-model.[11,13,14] In this application of a 1D Ising model, CTPR proteins are treated as linear

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

*Correspondence to: Aitziber L. Cortajarena, IMDEA-Nanociencia and Centro Nacional de Biotecnología, c/Darwin 3, Universidad Autonoma de Madrid, Cantoblanco, Madrid 28049, Spain.
E-mail: aitziber.lopezcortajarena@imdea.org

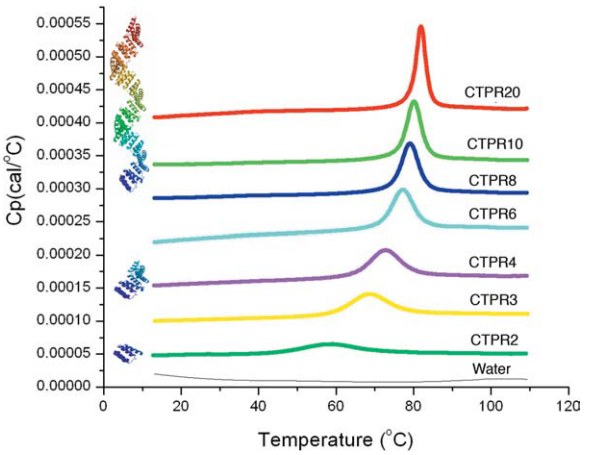

Figure 1. DSC thermograms of the proteins CTPRa2 to CTPRa20. The excess heat capacity is plotted versus temperature for CTPRa2 (green), CTPRa3 (yellow), CTPRa4 (magenta), CTPRa6 (cyan), CTPRa8 (blue), CTPRa10 (light green), and CTPRa20 (red). The curves are baseline-subtracted using a progressive baseline function and normalized to the molar protein concentration. Ribbon representations of the structures of three CTPR proteins within the series (CTPRa2, CTPRa4, and CTPRa20) are shown. Thermograms were recorded at a protein concentration of 0.5 mg/mL in 150 m*M* NaCl, 50 m*M* phosphate pH 6.8 buffer using a Micro Cal VP-Capillary DSC System (Micro Cal Llc, Northampton, MA), at a scan rate of 60°C/h.

polymers in which each repeat or helix is considered as the equivalent of an Ising spin, which can exist in one of two states: folded or unfolded. The Ising model predicts that a significant population of partially folded species will be populated at equilibrium close to the transition midpoint, and that such partially unfolded forms will have mostly the N- and/or C-terminal repeats unfolded.[11,15] By contrast, in two-state folding/unfolding only fully folded and fully unfolded species are populated.

Evidence that the TPR proteins do 'fray from the ends' comes from NMR hydrogen exchange (HX) studies of CTPR2 and CTPR3: The centermost helices are significantly more protected from exchange than the outer helices.[16] In addition, native state HX (NHX) experiments on CTPR2 and CTPR3, corroborate and expand the results of the initial HX experiments. Each helix has a different subglobal stability, which match those predicted by the Ising model.[15]

Here we present a differential scanning calorimetry (DSC) study of the thermodynamics of the unfolding transition of the complete series of CTPRa proteins (2–20 repeats).[17–19] The DSC data clearly show that these proteins exhibit non–two state unfolding behavior.

Results and Discussion

We performed a DSC study of all the proteins in the series: CTPRa2 to CTPRa20. The DSC thermograms show the following three characteristics (See Fig. 1 and Table I):

1. The greater the number of repeats, the higher the value of T_m. This behavior is consistent with the thermal denaturation data obtained by circular dichroism (CD).[9,11]
2. The greater the number of repeats, the sharper the denaturation peak, that is, the width at half height of the denaturation peak ($\Delta T_{1/2}$) decreases as the number of repeats increases. This behavior is in agreement with the observation from the CD data that the unfolding curves become sharper (the *m*-value increases) as the number of identical repeats within the CTPR array increases.[9,11]
3. The greater the number of repeats, the greater the area under the endothermic peak (*i.e.*, the unfolding enthalpy).

To properly analyze such denaturation curves, it is essential to establish that the transition is reversible. To this end, the protein samples were rescanned under identical conditions after reaching the temperature of denaturation. The first and successive rescans for each CTPR protein are superimposable, clearly indicating that the unfolding transitions are reversible (Supporting Information Fig. 1A,B).

Calorimetric test for two-state behavior: van't Hoff versus calorimetric enthalpy

Whether or not an unfolding transition is two-state can be readily tested in a DSC experiment. The DCS thermogram can be fit to the van't Hoff equation for a two-state process, and the van't Hoff enthalpy

Table I. *Thermodynamic Parameters for CTPRs Unfolding by DSC and CD*

	DSC[a]				CD[b]
	T_m (°C)	$\Delta T_{1/2}$	ΔH_{cal} (kcal mol^{-1})	ΔC_p (kcal mol^{-1} deg^{-1})	*Apparent* ΔH_{VH} (kcal mol^{-1})
CTPRa2	57.9	15.9	47.1	0.7	46.5
CTPRa3	68.3	11.0	78.2	0.9	63.2
CTPRa4	72.7	8.9	102.6	0.9	76.3
CTPRa6	77.1	6.2	191.9	1.9	115.4
CTPRa8	78.9	5.1	257.0	2.7	118.2
CTPRa10	80.1	4.3	334.9	2.8	n.d.[c]

[a]The thermodynamic data has been calculated from the DSC thermograms using a model free analysis in microcal origin. A progress baseline was used to setup the baseline before using the integration function to calculate ΔH_{cal}, T_m, and $\Delta T_{1/1}$. A step baseline was applied to calculate approximate ΔC_p values. The *apparent* ΔH_{VH} was calculated from the CD data assuming a two state model with linear extrapolated baselines for the native and unfolded states.
[b]The *apparent* ΔH_{VH} is the slope of the linear fit to the Van't Hoff plot (In *K* vs. 1/*T*).
[c]n.d. not determined.

(ΔH_{VH}) of the unfolding process calculated.[20] In addition, the area beneath the peak in the DSC endotherm can be used to calculate calorimetric enthalpy (ΔH_{cal}), which is independent of any model assumptions. If both ΔH_{VH} and ΔH_{cal} are the same, the denaturation may be considered to be a two state process. Extensive calorimetric studies on small globular domains have demonstrated that these proteins typically show ideal two-state behavior: van't Hoff enthalpy equals calorimetric enthalphy.[17,21] If the van't Hoff enthalpy is smaller than the calorimetric enthalpy it is likely that the unfolding process includes the formation of unfolding intermediates. We applied a simple two-state model (Microcal Origin) to fit the thermograms for the CTPR series (Fig. 2) and thus calculate ΔH_{VH} for each protein. The DSC thermograms of CTPRs longer than 3 repeats could not be well fit by such a two-state model. As the number of repeats increases the quality of the two-state model fit to the data progressively decreases (Fig. 2). This result is by itself a clear indication that CTPR unfolding does not follow a simple two-state model. Because the DSC data fit so poorly to a two state model, we did not pursue this method of calculating ΔH_{VH}. Instead, we estimated ΔH_{VH} from the thermal denaturation data measured by CD. All such denaturation curves for the CTPR series, even though they are better described by a 1D-Ising model, can also be fit to a two-state model, and an "apparent ΔH_{VH}" thus estimated. We compared such estimates of ΔH_{VH} with the model free ΔH_{cal} values measured by DSC. For all the CTPRs in the series (2–20 repeats) we observed that ΔH_{VH} is smaller than ΔH_{cal}, and that the difference between them increases as the number of repeats in the protein increases (Table I).

Origin of the asymmetry of the thermograms

It is also clear that the thermograms are not symmetric, and that the asymmetry becomes increasingly pronounced as the number of repeats in the protein increases (Fig. 2). A possible explanation for the asymmetry could be the contribution of kinetically determined irreversible processes. If such processes play a significant role, the shape of the thermograms should depend on the scan rate. To investigate this possibility, we acquired data using two different scan rates 60°C/h and 30°C/h. The DSC thermograms were unaffected by the change in the scan rate (Supporting Information Figure 1C), thus effectively ruling out this possible artifactual origin of the asymmetry.

The asymmetry observed in the DSC thermograms is evident for all proteins in the series, and becomes more pronounced for longer arrays. The DSC unfolding peaks are skewed at temperatures below the transition midpoint (T_m), where the transition is less sharp and deviates more from the two-state fit (Fig. 2). This observation can be explained by the existence of partially unfolded species before and around the transition mid-point but not after it.[2,15,22]

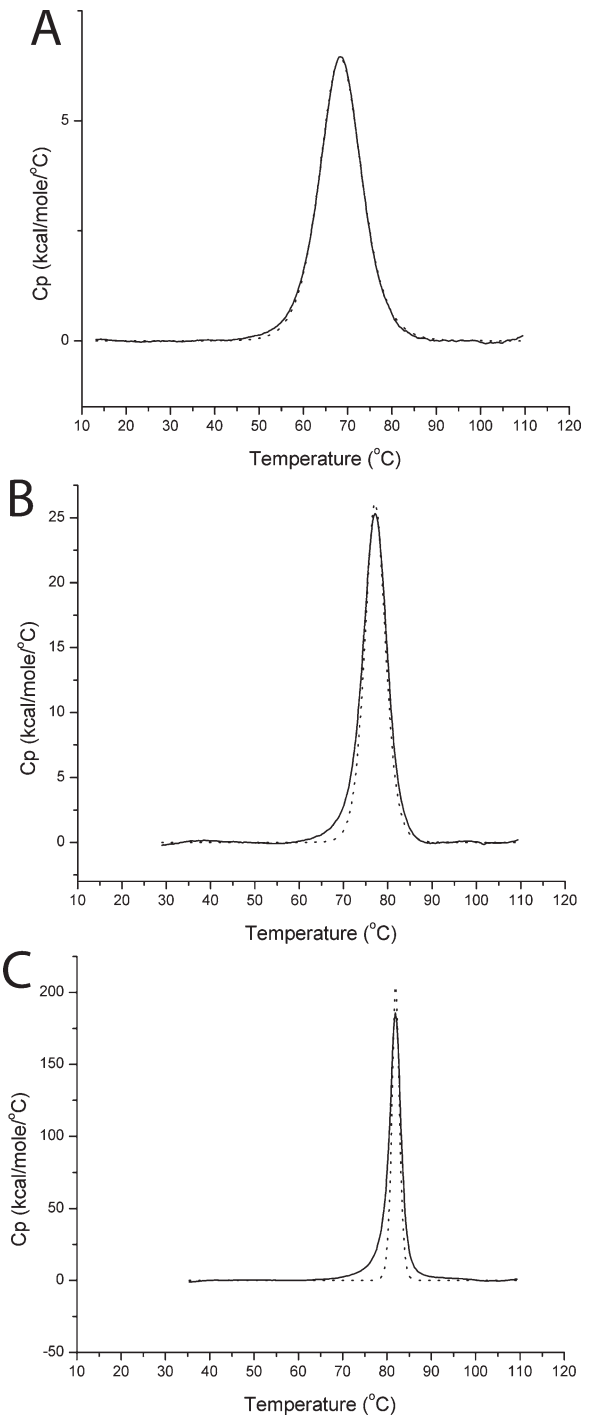

Figure 2. Attempting to fit the DSC thermograms of CTPRa3 (A), CTPRa6 (B) and CTPRa20 (C) to a two-state model. Solid black lines are the data, dotted lines show the best fit of the data to a two-state unfolding model (MicroCal Origin software package). It is clear that the quality of the fit to a two state model becomes worse as the number of tandem repeats in the protein increases. Chi-square values for the fit are $6.4 \times 10^{+3}$ for CTPRa3, $6.3 \times 10^{+5}$ for CTPRa6, and $1.2 \times 10^{+8}$ for CTPRa20.

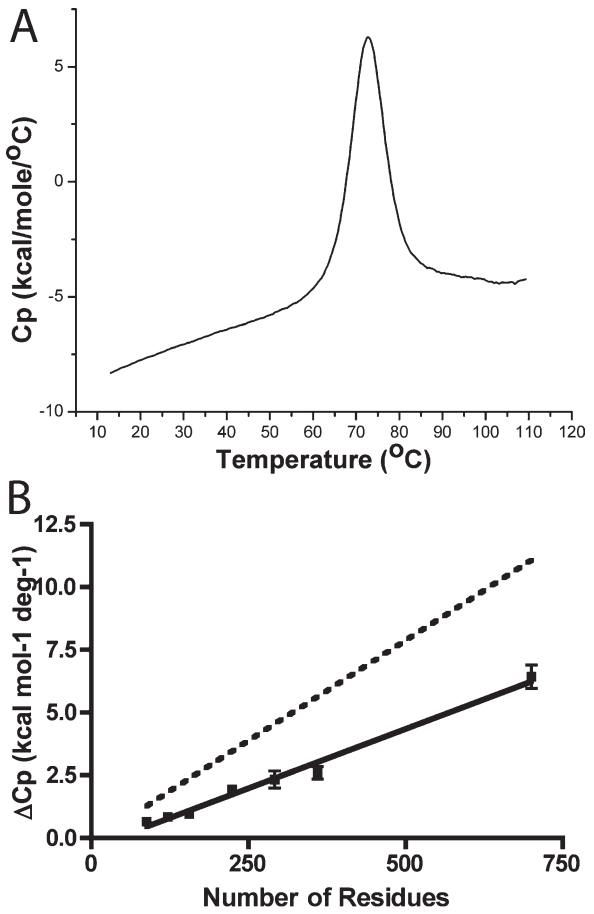

Figure 3. Heat capacity change for CTPR unfolding (A) Representative thermogram of CTPR to show the data quality for the ΔC_p calculations. Thermogram of CTPRa4 protein normalized to the protein concentration and from which the buffer-buffer thermogram baseline has been subtracted. (B) Linear relationship between ΔC_p and the number of residues in the protein. ΔC_p is plotted versus number of residues for the proteins CTPRa2 to CTPRa20 (solid symbols). Data are average of two independent experiments. Error bars are the standard deviation, when not visible are smaller than the symbols of the average ΔC_p value. The solid black line is linear fit of these data, giving a slope of 0.010 kcal mol^{-1} deg^{-1} $residues^{-1}$. The dotted line shows the data for ΔC_p versus number of residues for many globular proteins.[23] The slope of this line is 0.016 kcal mol^{-1} deg^{-1} $residues^{-1}$.

Heat capacity change for unfolding

Another important thermodynamic parameter that can be extracted from the thermograms is the heat capacity change, ΔC_p, associated with the denaturation transition, that is, the difference in the heat capacity of the native and denatured states (Table I) (Fig. 3A). Direct calculation of the heat capacity from a single DSC scan depends on the quality of the DSC data and baseline stability. We consider the DSC data we present to be of sufficient quality (Fig. 3A) to allow the direct calculation of accurate ΔC_p values in this fashion. We present the standard error obtained from two independent experiments (Fig. 3B) to lend additional support to this view. We found that the value of ΔC_p increases linearly with the number of residues in the protein (Fig. 3B). It is widely accepted that the value of ΔC_p is linearly dependent on the change in solvent-accessible nonpolar surface area upon unfolding (ΔASA), which is directly proportional to the number of residues in a protein.[24] Such data has been assembled for a variety of globular proteins, and the slope of the plot of ΔC_p versus number of residues is 0.016 kcal mol^{-1} deg^{-1} $residues^{-1}$. Interestingly, the slope of a plot of ΔC_p versus number of residues for the set of CTPR proteins is 0.010 kcal mol^{-1} deg^{-1} $residues^{-1}$, significantly smaller than for globular proteins (Fig. 3B).[23,24] This observation implies that the surface area exposed upon unfolding of a CTPR protein is smaller than expected from the number of residues in the protein. It can be rationalized by two unusual features of CTPR proteins: (1) their unfolded state is more compact than predicted by the typical self avoiding random walk treatment of the denatured state,[25] and (2) their elongated native state results in a greater surface area to volume ratio than for a globular protein.[9] Both these effects will contribute to the observed deviation from the behavior of globular proteins. The change in accessible surface area in going from folded to compact unfolded is less than that in going from folded to extended unfolded. Similarly, the change in going from extended folded to unfolded is less than in going from compact folded to unfolded. We cannot estimate the possible contribution of a more compact unfolded state, because we do not have a detailed structural characterization of the thermally denatured unfolded state. We can, however, estimate the contribution of the extended folded state by calculating the ASA_N from the crystals structures of the CTPR proteins using the program GETAREA (http://curie.utmb.edu/getarea.html). We can then compare these calculated ASA_N to the ASA of globular proteins. We observe a small deviation for CTPR proteins with 20 or fewer repeats (Supporting Information Figure 2). We therefore conclude that the predominant cause of the deviation in the ΔC_p values is the nature of the unfolded state.

Conclusions

Here we present direct experimental evidence for the non–two state unfolding behavior of modular designed proteins (CTPRs). The DSC thermograms for CTPR proteins cannot be fit to a two state model, which clearly shows that the CTPR proteins do not unfold by a two-state mechanism. In addition, the estimated van't Hoff enthalpy does not equal the calorimetric enthalpy, the van't Hoff enthalpy is lower.

Interestingly, different conclusions were reached from DSC studies on a different class of repeat protein, the Notch ankyrin repeat domain (Nank).

Zweifel and Barrick studied the thermal denaturation of Nank1-7 (the first 7 ank repeats of the Notch Ank domain) and Nank1-6 (first 6 repeats), note that these repeats are not identical to each other.[26] For Nank1-7 a ratio of ΔH_{VH} to ΔH_{cal} of 0.99 was determined, consistent with two-state unfolding/folding.[26,27] The unfolding transition of Nank1-6 was much broader making quantitative analysis problematic, and preventing the definitive interpretation of the data as indicative of two-state or multistate unfolding. The reason for the difference in behavior of the TPR versus Ank proteins in the DSC experiments is not clear, because other studies on the Nank proteins clearly indicate non–two state "Ising-like" behavior.[28]

In conclusion, the DSC thermograms of CTPR proteins and thermodynamic parameters calculated from them, clearly show how the modularity of these TPR proteins confers on them features that are readily distinguishable from those of globular proteins. The modular structure of TPR proteins translates into modular thermodynamics.

Acknowledgments

The authors thank the members of the Regan Lab: Tijana Z. Grove, Robielyn Ilagan, Daniel Schlingman, and Alice Q. Zhou for their comments on the manuscript. They thank G. I. Makhatadze and Simon G. J. Mochrie for insightful suggestions. This work was supported, in part, by the Human Frontier Science Program and by NIH (Grant R01GM080515). DSC experiments were acquired using a Micro Cal VP-Capillary DSC System (Micro Cal Llc, Northampton, MA) at the Yale Chemical Instrumentation Center.

References

1. Main E, Lowe A, Mochrie S, Jackson S, Regan L (2005) A recurring theme in protein engineering: the design, stability and folding of repeat proteins. Curr Opin Struc Biol 15:464–471.
2. Cortajarena AL, Regan L (2010) The folding of repeat proteins. Oxford: Elsevier Science.
3. Kloss E, Courtemanchea N, Barrick D (2008) Repeat-protein folding: New insights into origins of cooperativity, stability, and topology. Arch Biochem Biophys 469:83–99.
4. Schultz J, Marshall-Carlson L, Carlson M (1990) The N-terminal TPR region is the functional domain of SSN6, a nuclear phosphoprotein of *Saccharomyces cerevisiae*. Mol Cell Biol 10:4744–4756.
5. Lamb JR, Tugendreich S, Hieter P (1995) Tetratrico peptide repeat interactions: to TPR or not to TPR? Trends Biochem Sci 20:257–259.
6. D'Andrea L, Regan L (2003) TPR proteins: the versatile helix. Trends Biochem Sci 28:655–662.
7. Main ERG, Jackson SE, Regan L (2003) The folding and design of repeat proteins: reaching a consensus. Curr Opin Struct Biol 13:482–489.
8. Main ERG, Xiong Y, Cocco MJ, D'Andrea L, Regan L (2003) Design of stable alpha-helical arrays from an idealized TPR motif. Structure 11:497–508.
9. Kajander T, Cortajarena AL, Mochrie SG, Regan L (2007) Structure and stability of a consensus TPR superhelix. Acta Cryst D63:800–811.
10. Kajander T, Cortajarena AL, Regan L (2006) Consensus design as a tool for engineering repeat proteins. Methods Mol Biol 340:151–170.
11. Kajander T, Cortajarena AL, Main ER, Mochrie SG, Regan L (2005) A new folding paradigm for repeat proteins. J Am Chem Soc 127:10188–10190.
12. Cheng CY, Jarymowycz VA, Cortajarena AL, Regan L, Stone MJ (2006) Repeat motions and backbone flexibility in designed proteins with different numbers of identical consensus tetratricopeptide repeats. Biochemistry 45:12175–12183.
13. Zimm BH, Bragg JK (1959) Theory of the phase transition between helix and random coil in polypeptide chains. J Chem Phys 31:526–535.
14. Ising E (1925) Beitrag zur Theorie des Ferromagnetismus. Z Phys 31:253–258.
15. Cortajarena AL, Mochrie SG, Regan L (2008) Mapping the energy landscape of repeat proteins using NMR-detected hydrogen exchange. J Mol Biol 379:617–626.
16. Main ER, Stott K, SEJ, Regan L (2005) Local and long-range stability in tandemly arrayed tetratricopeptide repeats. Proc Natl Acad Sci USA 102:5721–5726.
17. Freire E (1995) Differential scanning calorimetry. Methods Mol Biol 40:191–218.
18. Sanchez-Ruiz J (1995) Differential scanning calorimetry of proteins. Subcell Biochem 24:133–176.
19. Makhatadze G, Privalov P (1995) Energetics of protein structure. Adv Prot Chem 47:307–425.
20. Privalov PL, Khechinashvili NN (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. J Mol Biol 86:665–684.
21. Privalov PL (1979) Stability of proteins: small globular proteins. Adv Prot Chem 33:167–241.
22. Cziepluch C, Kordes E, Poirey R, Grewenig A, Rommelaere J, Jauniaux J (1998) Identification of a novel cellular TPR-containing protein, SGT, that interacts with the nonstructural protein NS1 of parvovirus H-1. J Virol 72:4149–4159.
23. Geierhaas CD, Nickson AA, Lindorff-Larsen K, Clarke J, Vendruscolo M (2007) BPPred: a Web-based computational tool for predicting biophysical parameters of proteins. Protein Sci 16:125–134.
24. Myers JK, Pace CN, Scholtz JM (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Protein Sci 4:2138–2148.
25. Cortajarena AL, Lois G, Sherman E, O'Hern CS, Regan L, Haran G (2008) Extensive non-native polyproline II secondary structure induces compaction of a protein's denatured state. J Mol Biol 382:203–212.
26. Zweifel ME, Barrick D (2001) Studies of the ankyrin repeats of the *Drosophila melanogaster* notch receptor. 2. Solution stability and cooperativity of unfolding. Biochemistry 40:14357–14367.
27. Bradley CM, Barrick D (2002) Limits of cooperativity in a structurally modular protein: response of the Notch ankyrin domain to analogous alanine substitutions in each repeat. J Mol Biol 324:373–386.
28. Mello CC, Barrick D (2004) An experimentally determined protein folding energy landscape. Proc Natl Acad Sci USA 101:14102–14107.