

## **The Linear Extrapolation Method of Analyzing Solvent Denaturation Curves**

**Running Title:** The linear extrapolation method

**Keywords:** Linear extrapolation method,  $m$  value, urea denaturation

C. Nick Pace\* and Kevin L. Shaw

Departments of Medical Biochemistry and Genetics, Biochemistry & Biophysics, and Center for  
Macromolecular Design  
Texas A&M University  
College Station, Texas 77843-1114

\*Corresponding Author: C. Nick Pace Email: [nickpace@tamu.edu](mailto:nickpace@tamu.edu) Phone: 409-845-1788; Fax:  
409-847-9481

**Pages:** 22, **Tables:** 2, **Figures:** 4

**Summary:** The two most common methods of measuring the conformational stability of a protein are differential scanning calorimetry and an analysis of solvent denaturation curves using the linear extrapolation method. In this article we trace the history of the linear extrapolation method, review how the method is used to measure protein stability, and then discuss some of the other important uses.

### **Historical Perspective**

The ability of urea to denature proteins has been known for 100 years<sup>1</sup>. Urea is also effective in denaturing more complex biological systems: “A dead frog placed in saturated urea solution becomes translucent and falls to pieces in a few hours.”<sup>2</sup>. (It was not reported if frog denaturation was reversible.) The even greater effectiveness of guanidine hydrochloride (GdnHCl) as a protein denaturant was first reported by Greenstein<sup>3</sup>. In this article the emphasis will be on urea which we favor for most applications.

Tanford<sup>4</sup> was the first to deal quantitatively with the unfolding of proteins by urea. The schematic in Figure 1 is from Tanford’s paper; it shows that when a protein unfolds many nonpolar side chains and peptide groups that were buried in the folded protein are exposed to solvent in the unfolded protein. In the lower half of the figure are measured values of the free energy of transfer,  $\Delta G_{tr}$ , of a leucine side chain and a peptide group from water to the solvents shown. It is easy to see why proteins unfold in urea and GdnHCl solutions: the free energy of both leucine side chains and peptide groups is lower in the presence of urea and GdnHCl than they are in water. (It turns out that this is true of all of the constituent groups of a protein.) In ethanol, the protein unfolds to expose the nonpolar side chains, but then refolds into more helical structures to shield the peptide groups from the ethanol. It is also easy to see why the four

compounds listed at the bottom stabilize proteins: transferring leucine side chains and peptide groups from water into these solvents is unfavorable so that the stability of a protein is increased when placed in these solvents<sup>5</sup>.

A typical urea denaturation curve is shown in Figure 2A. When a protein unfolds by a two-state mechanism, the equilibrium constant,  $K$ , can be calculated from the experimental data using:

$$K = [(y)_N - (y)] / [(y) - (y)_D] \quad (1)$$

where  $(y)$  is the observed value of the parameter used to follow unfolding, and  $(y)_N$  and  $(y)_D$  are the values  $(y)$  would have for the native state and the denatured state under the same conditions where  $(y)$  was measured. In the original analyses of urea denaturation curves<sup>6, 7</sup>,  $\log K$  was found to vary linearly as a function of  $\log [\text{urea}]$  and the slope of the plot was denoted by  $n$ , and the midpoint of the curve by  $(\text{urea})_{1/2}$  (where  $\log K = 0$ ). These parameters could then be used to calculate the dependence of the standard free energy of denaturation,  $\Delta G^\circ (= -RT \ln K)$ , on urea concentration with <sup>4</sup>:

$$d(\Delta G^\circ)/d(\text{urea}) = RTn / (\text{urea})_{1/2} \quad (2)$$

This equation was used by Alexander and Pace<sup>6</sup> to estimate the differences in stability among three genetic variants of a protein for the first time, and was the forerunner of the linear extrapolation method.

The linear extrapolation method (LEM) was first used to analyze urea and GdnHCl denaturation curves by Greene and Pace<sup>8</sup>. When  $\Delta G^\circ$  is calculated as a function of urea concentration using data such as those shown in Figure 2A,  $\Delta G^\circ$  is found to vary linearly with

urea concentration as shown in Figure 2B. Based on results such as these for several proteins, this equation was proposed for analyzing the data:

$$\Delta G = \Delta G(\text{H}_2\text{O}) - m[\text{urea}] \quad (3)$$

where  $\Delta G(\text{H}_2\text{O})$  is an estimate of the conformational stability of a protein that assumes that the linear dependence continues to 0 M denaturant, and  $m$  is a measure of the dependence of  $\Delta G$  on urea concentration, i.e., the slope of the plot shown in Figure 2B. The same approach has recently been proposed for measuring the stability of RNA molecules<sup>9</sup>.

In the early days,  $(y)_N$  and  $(y)_D$  were obtained by extrapolating the pre- and posttransition baselines into the transition region, and then using Equation 1 to calculate  $K$  and then  $\Delta G^\circ$ . However, Santoro and Bolen<sup>10</sup> had a better idea and proposed that nonlinear least squares be used to directly fit data such as those shown in Figure 2A. With their approach, six parameters are used to fit the data: a slope and an intercept for the pretransition and posttransition regions, and  $\Delta G(\text{H}_2\text{O})$  and  $m$  for the transition region leading to

$$y = \{(y_F + m_F[\text{urea}]) + (y_U + m_U[\text{urea}]) \cdot \exp - ((\Delta G(\text{H}_2\text{O}) - m \cdot [\text{urea}]) / RT)\} / (1 + \exp - (\Delta G(\text{H}_2\text{O}) - m \cdot [\text{urea}]) / RT) \quad (4)$$

where  $y_F$  and  $y_U$  are the intercepts and  $m_F$  and  $m_U$  the slopes of the pre and post-transition baselines, and  $\Delta G(\text{H}_2\text{O})$  and  $m$  are defined by Equation 3.

The linear extrapolation method is now widely used for estimating the conformational stability of proteins and for measuring the difference in stability between proteins differing slightly in structure. (The original how-to-do-it paper on solvent denaturation<sup>11</sup> has been cited  $\approx$

900 times.) In addition, the  $m$  value has taken on a life of its own<sup>12</sup>. These topics will be discussed further below.

## **DG(H<sub>2</sub>O)**

The conformational stability of RNase Sa at 25°C and pH 5 is 7.0 kcal/mol. This corresponds to 1 unfolded molecule for each 135,000 folded molecules ( $K = 7.4 \times 10^{-6}$ ). To study the equilibrium between the folded and unfolded conformations by conventional techniques requires destabilizing the protein so that both conformations are present at measurable concentrations. With urea denaturation this is done by increasing the urea concentration. It is clear from Figure 2A that the unfolding equilibrium can be studied only near 7 M urea, and that a long extrapolation is needed to estimate  $\Delta G^\circ$  in the absence of urea,  $\Delta G(\text{H}_2\text{O})$ . The same is true for thermal denaturation. The unfolding equilibrium can only be studied near 55°C (Figure 3A). Thermal denaturation curves can be analyzed to obtain the midpoint of the thermal denaturation curve,  $T_m$ , and the enthalpy change at  $T_m$ ,  $\Delta H_m$ . These two parameters plus the heat capacity change for folding,  $\Delta C_p$ , can then be used in the Gibbs-Helmholtz equation<sup>13</sup>

$$\Delta G(T) = \Delta H_m(1 - T/T_m) + \Delta C_p[T - T_m - T \ln(T/T_m)] \quad (5)$$

to calculate  $\Delta G$  at any temperature  $T$ ,  $\Delta G(T)$ . The temperature dependence of  $\Delta G$  and the value of  $\Delta G(25^\circ\text{C})$  for the denaturation of RNase Sa are shown in Figure 3B. Note that the value of  $\Delta G(25^\circ\text{C})$  from the thermal denaturation experiment agrees with the  $\Delta G(\text{H}_2\text{O})$  value from the urea denaturation experiments within experimental error, which is about  $\pm 0.3$  kcal/mol (see Eftink and Ionescu[Eftink, 1997 #48] for an excellent discussion of the problems most frequently encountered in analyzing solvent and thermal denaturation curves).

Previously, we used RNase T1 to show that estimates of  $\Delta G$  (T) based on DSC experiments were in good agreement with estimates of  $\Delta G$  (H<sub>2</sub>O) from urea denaturation<sup>14, 15</sup>. In a more recent study using RNase A, we compared  $\Delta G$  (T) values from DSC with  $\Delta G$  (H<sub>2</sub>O) values determined using urea denaturation and the LEM<sup>16</sup>. The agreement is remarkably good (Table I). Most recently, we have compared conformational stabilities estimated from hydrogen exchange rates measured under native state conditions with those from thermal or solvent denaturation<sup>17</sup>. Again, the conformational stabilities are in good agreement.

There is considerable experimental evidence that proteins are more extensively unfolded in solutions of urea than they are in water (See, for example, Qu et al.<sup>18</sup>). Thus, it is surprising that conformational stabilities determined under conditions where the denatured state ensembles differ would be the same. This suggests that the denatured state ensemble that exists under physiological conditions is thermodynamically equivalent to the ensembles that exist after thermal or urea denaturation, even though they do not appear to be structurally equivalent.

The excellent agreement between conformational stabilities based on DSC results and the LEM suggests that the LEM is a reliable method for estimating the conformational stability of a protein. Two other methods have been used to analyze urea denaturation curves. They both give estimates of  $\Delta G$  (H<sub>2</sub>O) that are larger than those based on the LEM. The differences are not large with urea, but they are with GdnHCl. One is an extrapolation based on Tanford's model<sup>4</sup> that is explained in the next section. This method uses  $\Delta G_{tr}$  values for transfer of the constituent groups of a protein from water to urea solutions. Unfortunately, there is uncertainty in the  $\Delta G_{tr}$  value to use for a peptide group<sup>19</sup>, and this introduces uncertainty into the  $\Delta G$ (H<sub>2</sub>O) values obtained by

this method<sup>20</sup>. Experiments are underway to determine a more reliable value of  $\Delta G_{tr}$  for a peptide group (Auton & Bolen, personal communication). Makhatadze<sup>21</sup> has shown that the data on which this method is based are consistent with the LEM when urea but not GdnHCl is used as the denaturant.

The other method that can be used to determine  $\Delta G (H_2O)$  is the denaturant binding model used first by Aune and Tanford<sup>22</sup> and recently discussed by Wu and Wang<sup>23</sup>. The plots of  $\log K$  vs.  $\log [\text{urea}]$  mentioned above show that more urea molecules are bound by the denatured state than by the native state. Also, urea increases the solubility of all of the constituent parts of a protein, and it is possible to account for the enhanced solubility in terms of urea binding, even for leucine side chains (See Table XI in Tanford<sup>24</sup>). Thus, it might seem reasonable to try and analyze urea denaturation curves in terms of the denaturant binding model. Generally, all of the urea binding sites are considered identical and independent and  $K = 0.1$  is used for the binding constant<sup>11</sup>. However, this is surely not stoichiometric binding and it is unlikely that the binding sites are identical and independent. The enhanced solubility of the model compound data requires “binding constants” in the range 0.05 to 0.3<sup>21, 24, 25</sup>. Schellman has argued convincingly that a solvent exchange model is more reasonable than a stoichiometric binding model, and this model is consistent with the LEM (See Schellman & Gassner<sup>26</sup>, which gives references to earlier work).

$\Delta G (H_2O)$  values from GdnHCl denaturation studies are less reliable, as pointed out most recently by Makhatadze<sup>21</sup>. This conclusion was reached earlier by Schellman and Gassner<sup>26</sup> who said “Finally this study as well as a number of others indicates that urea has a number of

advantages over guanidinium chloride as far as thermodynamic interactions are concerned. The free energy and enthalpy functions of both proteins and model compounds are more linear; the solutions more ideal; extrapolations to zero concentration are more certain; and least-squares analysis of the data is more stable.” One problem is that the ionic strength cannot be controlled with GdnHCl, and this has been shown to effect the results in several studies (See, for example, Ibarra-Molero et al.<sup>27</sup>, Monera et al.<sup>28</sup>, Santoro & Bolen<sup>29</sup>).

In summary, the agreement between  $\Delta G$  (T) values from DSC and thermal denaturation studies and  $\Delta G$  (H<sub>2</sub>O) values from urea denaturation studies analyzed by the LEM suggests that either method can be used to reliably measure the conformational stability of a protein.

### *m*

The thermodynamic cycle in Figure 4 shows that the dependence of  $\Delta G$  on denaturant concentration is determined by the groups in a protein that are exposed to solvent in the denatured state, D, but not in the native state, N<sup>4</sup>. In support of this, Myers et al.<sup>30</sup> showed that there is a good correlation between *m* values and the change in accessible surface area on unfolding. Since  $\Delta g_{tr,i}$  values are available for most of the constituent groups in a protein, the equation at the bottom of Figure 4 can be used to calculate the dependence of  $\Delta G$  on denaturant concentration. The *m* value is an experimental measure of the dependence of  $\Delta G$  on denaturant concentration. Consequently, for proteins that unfold by a two-state mechanism,  $\Delta\alpha$  can be estimated by comparing the calculated and measured *m* values. Thus,  $\Delta\alpha$  is the fraction of buried groups that must become exposed to solvent on unfolding to account for the measured *m* value<sup>31</sup>. In the original LEM paper<sup>8</sup>, this approach was used to estimate  $\Delta\alpha$  for four proteins. It was

clear that differences among the  $\Delta\alpha$  values for individual proteins revealed differences in the accessibility of the denatured state and that these differences depended to some extent on the number and location of the disulfide bonds in the protein.

Interest in  $m$  values was stimulated by a series of papers from the Shortle laboratory<sup>32</sup>,<sup>33</sup>. (A review of these studies was titled: “Staphylococcal nuclease: a showcase of  $m$  value effects.”<sup>12</sup>) Their studies of mutants of staphylococcal nuclease (S. Nuclease) showed a wide range of  $m$  values. Those with  $m$  values 5% or more greater than wild type were designated  $m^+$  mutants ( $\approx 25\%$ ) and those with  $m$  values 5% or more less than wild type were designated  $m^-$  mutants ( $\approx 50\%$ ). Their interpretation was that  $m^+$  mutants unfolded to a greater extent than wild type and visa versa for  $m^-$  mutants. These papers were extremely important because they focused attention on the denatured state which had largely been ignored in discussions of protein stability. (Another review from the Shortle lab<sup>34</sup> was titled: “The denatured state (the other half of the folding equation) and its role in protein stability”.) However, this interpretation is only straightforward if all of the mutants unfold by a two-state mechanism. Shortle and Meeker<sup>33</sup> chose to assume a two-state mechanism to analyze their results even though the  $m^-$  mutants did not appear to unfold by a two-state mechanism. It is now clear that some of the  $m^-$  mutants unfold by a three-state mechanism<sup>35-38</sup>. The presence of an intermediate state will generally lower the  $m$  value<sup>20</sup>. One important consequence of this is that the estimates of  $\Delta G(\text{H}_2\text{O})$  will be too low when a two-state mechanism is assumed and the mechanism is, in fact, three-state. In the case of the  $m^-$  mutant V66L+G88V+G79S, a two-state analysis indicates that this mutant is 3.0 kcal/mol LESS stable than wild type<sup>33</sup>, but a three-state analysis indicates that that this mutant is

0.8 kcal/mol MORE stable than wild type<sup>36</sup>. This is also true for other  $m^-$  mutants<sup>39</sup>. This means that the  $\Delta(\Delta G)$  values for the  $m^-$  mutants are not correct. Since 50% of the S. Nuclease mutants are  $m^-$ , this calls into question the interpretation of many of the  $\Delta(\Delta G)$  values determined for S. nuclease.

In Table II, we have compiled  $m$  and  $\Delta\alpha$  values for a selection of proteins. In all cases, the folding of the proteins is thought to closely approach a two-state mechanism. The value of  $\Delta\alpha$  expected for unfolding a protein depends on the model assumed for the denatured state. If the accessibility of the denatured state is modeled as a tripeptide using the Lee and Richards<sup>40</sup> approach, a value of  $\Delta\alpha \approx 0.7$  would be expected. However, if a compact denatured state based on fragments with native-state conformations is assumed, a  $\Delta\alpha \approx 0.53$  would be expected<sup>41</sup>. Note in Table II that most of the  $\Delta\alpha$  values are considerably less than 0.53. This suggests that the urea-denatured states of proteins are less completely unfolded than in the hypothetical state that Creamer et al.<sup>41</sup> thought might be a lower limit for the compactness of the denatured state at the midpoint of a thermal denaturation curve in water. However, because of uncertainty in the  $\Delta g_r$  value for a peptide group, interpreting the absolute values of  $\Delta\alpha$  is hazardous. In contrast, differences among the  $\Delta\alpha$  values for individual proteins or for the same protein under different conditions are more trustworthy.

For the proteins without disulfide bonds, the  $\Delta\alpha$  values in Table II range from a low of 0.26 for thioredoxin to a high of 0.51 for calbindin. Thus, if these proteins unfold by a two-state mechanism, calbindin unfolds to a much greater extent than thioredoxin. Barnase also unfolds more completely than most of the other proteins in Table II. In support of this, we have shown

the tryptophan and tyrosine side chains of urea-denatured barnase are more accessible to a perturbant than those of several other proteins that have lower  $\Delta\alpha$  values<sup>42</sup>. The results in Table II suggest that urea-denatured proteins are not completely unfolded and that they unfold to different extents.

We were surprised to find that for RNases A and T1<sup>31</sup> and barnase<sup>42</sup>, the  $m$  value increases markedly as the pH is lowered from 7 to 3. This indicates that the denatured states interact more extensively with urea as the pH is lowered, and we suggested that the denatured state ensemble expands at low pH due to electrostatic repulsion among the positive charges. In support of this interpretation, Privalov et al.<sup>43</sup> have shown that the intrinsic viscosity of unfolded proteins increases at low pH. We have recently obtained results with RNase Sa that provide further support for this idea<sup>44</sup>.

RNase Sa is an acidic protein with a pI = 3.5 that contains no lysine residues. We have prepared a triple mutant that we call 3K (D1K, D17K, E41K) with a pI = 6.4, and a quintuple mutant that we call 5K (D1K, D17K, D25K, E41K, E74K) with a pI > 9. At pH 3, the estimated net charges are +8 for wild-type RNase Sa, + 11 for 3K, and +13 for 5K. The  $m$  values for the three proteins at pH 7 and pH 3 are given in Table II. For all three proteins, the  $m$  value is considerably greater at pH 3 than at pH 7, and the  $\Delta\alpha$  values increase from 0.32 for wild-type RNase Sa to 0.36 for 3K to 0.50 for 5K. This is consistent with an increase in accessibility to denaturant caused by an expansion of the denatured state due to electrostatic repulsion among the positive charges. As the pH increases from 3 to 7, the carboxyl groups are titrated and both negative and positive charges are present. Now attractive charge-charge interactions are possible, and the decrease in the  $m$  values suggests that the denatured state ensemble is more compact

because of attractive Coulombic interactions. In support of this, the  $\Delta\alpha$  value at pH 7 is lowest for 3K where the number of positive and negative charges is almost equal than for wild type RNase Sa which has an excess of negative charges or 5K which has an excess of positive charges. This and other evidence led us to conclude that long-range electrostatic interactions are important in determining the denatured state ensemble<sup>44</sup>.

We suggest that when the hydrophobic and hydrogen bonding interactions that stabilize the folded state are disrupted, the unfolded polypeptide chain rearranges to compact conformations that improve long-range electrostatic interactions. These charge-charge interactions in the denatured state will reduce the net contribution of electrostatic interactions to protein stability but will help determine the denatured state ensemble. If this idea is correct, then long-range charge-charge interactions in the denatured state may play an important role in the mechanism of proteins folding.

### **Concluding Remarks**

The LEM has proven to be a reliable method for measuring the conformational stability of a protein or the difference in stability between two proteins that differ slightly in structure. In addition, studies of the  $m$  values of proteins, especially *S. Nuclease*, have focused attention on the denatured state and it is now clear that denatured proteins do not approach a randomly-coiled conformation nearly as closely as the early studies by Tanford<sup>45</sup> suggested.

## References

1. Spiro K. Ueber die beeinflussung der eiweisscoagulation durch stickstoffhaltige substanzen. *Z. Physiol. Chem.* 1900;30:182-199.
2. Ramsden W. Some new properties of urea. *Journal of Physiology* 1902:xxiii-xxvii.
3. Greenstein JP. Sulfhydryl groups in proteins. *Journal of Biological Chemistry* 1938:501-513.
4. Tanford C. Isothermal unfolding of globular proteins in aqueous urea solutions. *J. Am. Chem. Soc.* 1964;86:2050-2059.
5. Wang A, Bolen DW. A naturally occurring protective system in urea-rich cells: mechanism of osmolyte protection of proteins against urea denaturation. *Biochemistry* 1997;36:9101-8.
6. Alexander SS, Pace CN. A comparison of the denaturation of bovine  $\alpha$ -lactoglobulins  $\alpha$  and  $\beta$  and goat  $\alpha$ -lactoglobulin. *Biochemistry* 1971;10:2738-43.
7. Pace NC, Tanford C. Thermodynamics of the unfolding of  $\beta$ -lactoglobulin A in aqueous urea solutions between 5 and 55 degrees. *Biochemistry* 1968;7:198-208.
8. Greene RF, Jr., Pace CN. Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme,  $\alpha$ -chymotrypsin, and  $\beta$ -lactoglobulin. *Journal of Biological Chemistry* 1974;249:5388-93.
9. Shelton VM, Sosnick TR, Pan T. Applicability of urea in the thermodynamic analysis of secondary and tertiary RNA folding. *Biochemistry* 1999;38:16831-16839.

10. Santoro MM, Bolen DW. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl  $\alpha$ -chymotrypsin using different denaturants. *Biochemistry* 1988;27:8063-8.
11. Pace CN. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* 1986;131:266-80.
12. Shortle D. Staphylococcal nuclease: a showcase of m-value effects. *Advances in Protein Chemistry* 1995;46:217-47.
13. Becktel WJ, Scheillman JA. Protein stability curves. *Biopolymers* 1987;26:1859-1877.
14. Yu Y, Makhatadze GI, Pace CN, Privalov PL. Energetics of ribonuclease T1 structure. *Biochemistry* 1994;33:3312-9.
15. Hu CQ, Sturtevant JM, Thomson JA, Erickson RE, Pace CN. Thermodynamics of ribonuclease T1 denaturation. *Biochemistry* 1992;31:4876-82.
16. Pace CN, Grimsley GR, Thomas ST, Makhatadze GI. Heat capacity change for ribonuclease A folding. *Protein Science* 1999;8:1500-4.
17. Huyghues-Despointes BM, Scholtz JM, Pace CN. Protein conformational stabilities can be determined from hydrogen exchange rates. *Nature Structural Biology* 1999;6:910-2.
18. Qu Y, Bolen CL, Bolen DW. Osmolyte-driven contraction of a random coil protein. *Proceedings of the National Academy of Sciences of the United States of America* 1998;95:9268-73.
19. Liu Y, Bolen DW. The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry* 1995;34:12884-91.

20. Pace CN. The stability of globular proteins. *CRC Critical Reviews in Biochemistry* 1975;3:1-43.
21. Makhatadze GI. Thermodynamics of protein interactions with urea and guanidinium hydrochloride. *Journal of Physical Chemistry B* 1999;103:4781-4785.
22. Aune KC, Tanford C. Thermodynamics of the denaturation of lysozyme by guanidine hydrochloride. II. Dependence on denaturant concentration at 25 degrees. *Biochemistry* 1969;8:4586-90.
23. Wu J, Wang Z. New evidence for the denaturant binding model. *Protein Science* 1999;8:2090-2097.
24. Tanford C. Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Advances in Protein Chemistry* 1970;24:1-95.
25. Liepinsh E, Otting G. Specificity of urea binding to proteins. *J. Am. Chem. Soc.* 1994;116:9670-9674.
26. Schellman JA, Gassner NC. The enthalpy of transfer of unfolded proteins into solutions of urea and guanidinium chloride. *Biophysical Chemistry* 1996;59:259-275.
27. Ibarra-Molero B, Loladze VV, Makhatadze GI, Sanchez-Ruiz JM. Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry* 1999;38:8138-49.
28. Monera OD, Kay CM, Hodges RS. Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Science* 1994;3:1984-91.

29. Santoro MM, Bolen DW. A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range. *Biochemistry* 1992;31:4901-4907.
30. Myers JK, Pace CN, Scholtz JM. Denaturant  $m$  values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding [published erratum appears in *Protein Sci* 1996 May;5(5):981]. *Protein Science* 1995;4:2138-48.
31. Pace CN, Laurents DV, Thomson JA. pH dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. *Biochemistry* 1990;29:2564-72.
32. Wrabl J, Shortle D. A model of the changes in denatured state structure underlying  $m$  value effects in staphylococcal nuclease. *Nature Structural Biology* 1999;6:876-83.
33. Shortle D, Meeker AK. Mutant forms of staphylococcal nuclease with altered patterns of guanidine hydrochloride and urea denaturation. *Proteins* 1986;1:81-9.
34. Shortle D. The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB Journal* 1996;10:27-34.
35. Gittis AG, Stites WE, Lattman EE. The phase transition between a compact denatured state and a random coil state in staphylococcal nuclease is first-order. *Journal of Molecular Biology* 1993;232:718-24.
36. Carra JH, Privalov PL. Energetics of denaturation and  $m$  values of staphylococcal nuclease mutants. *Biochemistry* 1995;34:2034-41.
37. Ionescu RM, Eftink MR. Global analysis of the acid-induced and urea-induced unfolding of staphylococcal nuclease and two of its variants. *Biochemistry* 1997;36:1129-40.

38. Baskakov IV, Bolen DW. Monitoring the sizes of denatured ensembles of staphylococcal nuclease proteins: implications regarding  $m$  values, intermediates, and thermodynamics. *Biochemistry* 1998;37:18010-7.
39. Carra JH, Anderson EA, Privalov PL. Three-state thermodynamic analysis of the denaturation of staphylococcal nuclease mutants. *Biochemistry* 1994;33:10842-50.
40. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. *Journal of Molecular Biology* 1971;55:379-400.
41. Creamer TP, Srinivasan R, Rose GD. Modeling unfolded states of proteins and peptides. II. Backbone solvent accessibility. *Biochemistry* 1997;36:2832-5.
42. Pace CN, Laurents DV, Erickson RE. Urea denaturation of barnase: pH dependence and characterization of the unfolded state. *Biochemistry* 1992;31:2728-34.
43. Privalov PL, Tiktopulo EI, Venyaminov S, Griko Yu V, Makhatadze GI, Khechinashvili NN. Heat capacity and conformation of proteins in the denatured state. *Journal of Molecular Biology* 1989;205:737-50.
44. Pace CN, Alston RW, Shaw KL. Charge-Charge Interactions Influence the Denatured State Ensemble and Contribute to Protein Stability. *Journal of Molecular Biology* 2000;*submitted for publication*.
45. Tanford C. Protein denaturation. *Advances in Protein Chemistry* 1968;23:121-282.
46. Pace CN, Hebert EJ, Shaw KL, et al. Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2 and Sa3. *Journal of Molecular Biology* 1998;279:271-86.

47. Pace CN, Laurents DV. A new method for determining the heat capacity change for protein folding. *Biochemistry* 1989;28:2520-5.

**Table I.** Comparison of  $\Delta G(T)$  values from DSC with  $\Delta G(H_2O)$  values from urea denaturation curves.

pH	T (°C)	$T_m^a$ (°C)	$\Delta H_m^a$ (kcal/mol)	$\Delta G(T)^a$ (kcal/mol)	$\Delta G(H_2O)^b$ (kcal/mol)
2.8	17.1	44.9	79.4	5.5	5.4
2.8	21.1	44.9	79.4	4.9	4.9
2.8	24.9	44.9	79.4	4.3	4.3
2.8	27.8	44.9	79.4	3.7	3.5
2.8	25.0	44.9	79.4	4.5	4.3
3.0	25.0	49.1	82.7	5.1	5.2
3.55	25.0	54.5	91.5	6.7	6.4
4.0	25.0	56.1	94.2	7.2	7.3
5.0	25.0	58.6	99.1	8.1	7.9
6.0	25.0	60.3	100.7	8.5	8.6
7.0	25.0	61.8	102.3	8.9	9.1

<sup>a</sup>The  $T_m$ ,  $\Delta H_m$ , and  $\Delta C_p = 1.15 \text{ kcal mol}^{-1} \text{ K}^{-1}$  values are from Pace et al.<sup>16</sup>. They were used in Equation 5 to calculate the  $\Delta G(T)$  values.

<sup>b</sup> The first four  $\Delta G(H_2O)$  values are from Pace & Laurents<sup>47</sup>, and the last five  $\Delta G(H_2O)$  values are interpolated from Figure 6 in Pace et al.<sup>31</sup>.

**Table II.** Change in Accessibility ( $\Delta\alpha$ ) on Urea Denaturation Calculated from the Measured  $m$  Values Using Tanford's Model<sup>f</sup>

Protein	Residues	Disulfides	(urea) <sub>1/2</sub>	$m$	$\Delta\alpha$
RNase Sa (0K)(pH7) <sup>b</sup>	96	1	6.44	0.99	0.32
RNase Sa (0K)(pH3) <sup>b</sup>	96	1	1.48	1.75	0.41
RNase Sa (3K)(pH7) <sup>b</sup>	96	1	4.07	0.99	0.29
RNase Sa (3K)(pH3) <sup>b</sup>	96	1	1.47	2.00	0.46
RNase Sa (5K)(pH7) <sup>b</sup>	96	1	5.31	1.14	0.36
RNase Sa (5K)(pH3) <sup>b</sup>	96	1	1.30	2.21	0.50
RNase T1 <sup>c</sup>	104	2	5.30	1.21	0.32
RNase A <sup>d</sup>	124	4	6.98	1.40	0.35
Lysozyme (Hen) <sup>c</sup>	129	4	6.80	1.29	0.29
SH3 Domain <sup>c</sup>	62	0	3.74	0.77	0.34
Calbindin <sup>c</sup>	75	0	5.30	1.13	0.51
HPr (E. Coli) <sup>e</sup>	88	0	4.32	1.14	0.45
Acyl Phosphatase <sup>f</sup>	99	0	4.54	1.25	0.39
$\lambda$ Repressor <sup>c</sup>	102	0	4.42	1.09	0.38
RNase T1 <sup>d</sup>	104	0	3.34	1.64	0.38
FKBP <sup>c</sup>	107	0	4.18	1.46	0.42
Thioredoxin <sup>c</sup>	108	0	6.70	1.30	0.26
Barnase <sup>c</sup>	110	0	4.49	1.94	0.50
Che Y <sup>c</sup>	129	0	3.51	1.60	0.36
FABP <sup>c</sup>	131	0	5.43	1.77	0.40
S. Nuclease <sup>g</sup>	149	0	2.56	2.36	0.44
ApoMb <sup>h</sup>	153	0	2.05	2.49	0.47
RNase H <sup>c</sup>	155	0	3.88	1.93	0.35
DHFR (E.Coli) <sup>c</sup>	159	0	3.10	1.90	0.32
T4 Lysozyme <sup>c</sup>	163	0	6.30	2.00	0.38

- a. Tanford's model is described in Tanford[Tanford, 1964 #26] and the details of the approach used to calculate  $\Delta\alpha$  are given in Pace et al.[Pace, 1990 #29].
- b. The (urea)<sub>1/2</sub> and  $m$  values are from Shaw[Shaw, 2000 #50].
- c. The original reference for these proteins is given in Myers et al.[Myers, 1995 #36].
- d. See Pace et al.[Pace, 1990 #29].
- e. See Nicholson and Scholtz[Nicholson, 1996 #51].
- f. See ???
- g. See Shortle[Shortle, 1995 #14].
- h. See Hughson and Baldwin[REF].

## Figure Legends

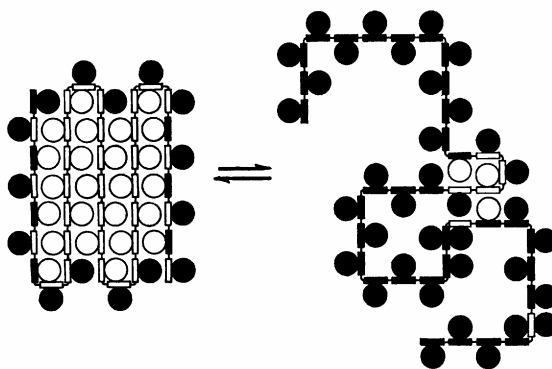
Figure 1. Schematic from Tanford<sup>4</sup> illustrating the changes in accessibility of peptide groups (rectangles) and side chains (circles) on protein unfolding. The closed symbols represent groups that are accessible to solvent and the open symbols represent groups that are not accessible to solvent. See Liu and Bolen[Liu, 1995 #10] for osmolyte and Pace et al.[Pace, 1990 #29] for urea  $\Delta G_{tr}$  values.

Figure 2. A. Urea denaturation curve for RNase Sa determined at 25°C, pH 5.0, 30 mM Sodium Acetate buffer as described by Pace et al.<sup>46</sup>. The solid line is based on an analysis of the data using Equation 4. B.  $\Delta G$  (o) as a function of urea molarity. The  $\Delta G$  values were calculated from the data in the transition region as described in the text. The  $\Delta G(H_2O)$  and  $m$  values can be determined by fitting the data in Part B to Equation 3, or by analyzing the data in Part A with Equation 4.

Figure 3. A. Thermal denaturation curve for RNase Sa determined at pH 5.0 in 30 mM Sodium Acetate buffer as described<sup>46</sup>. The solid line is based on a nonlinear least squares analysis of the data similar to that described for Equation 4, but for thermal denaturation curves<sup>46</sup>. B.  $\Delta G$  (o) as a function of temperature. The  $\Delta G$  values were calculated from data in the transition region as described in the text, and the data were analyzed to determine  $T_m$  and  $\Delta H_m$  as described<sup>46</sup>. The

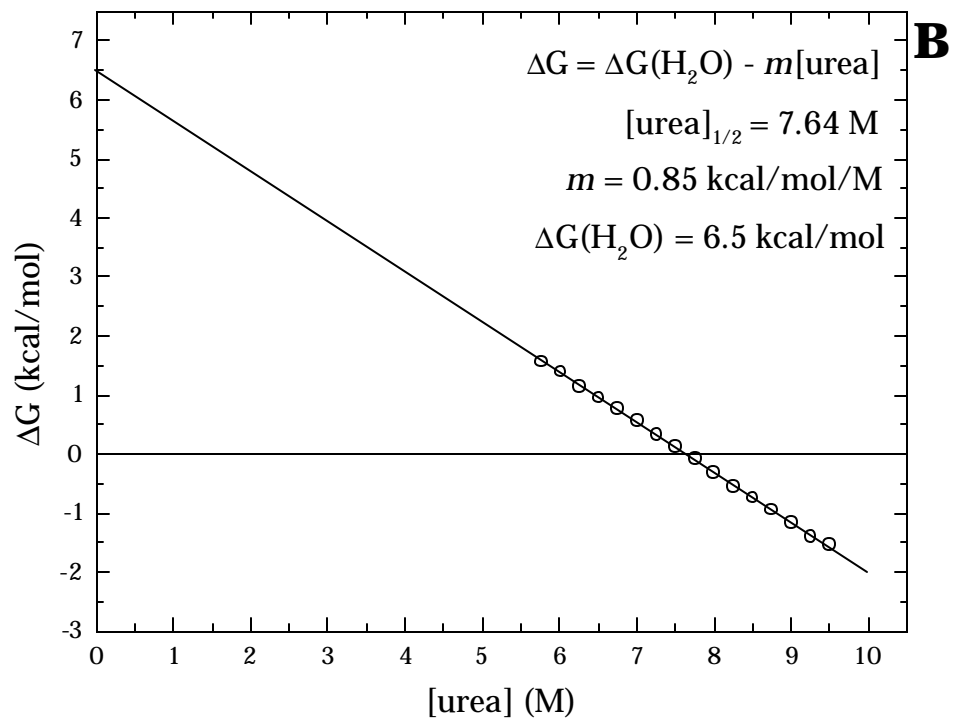
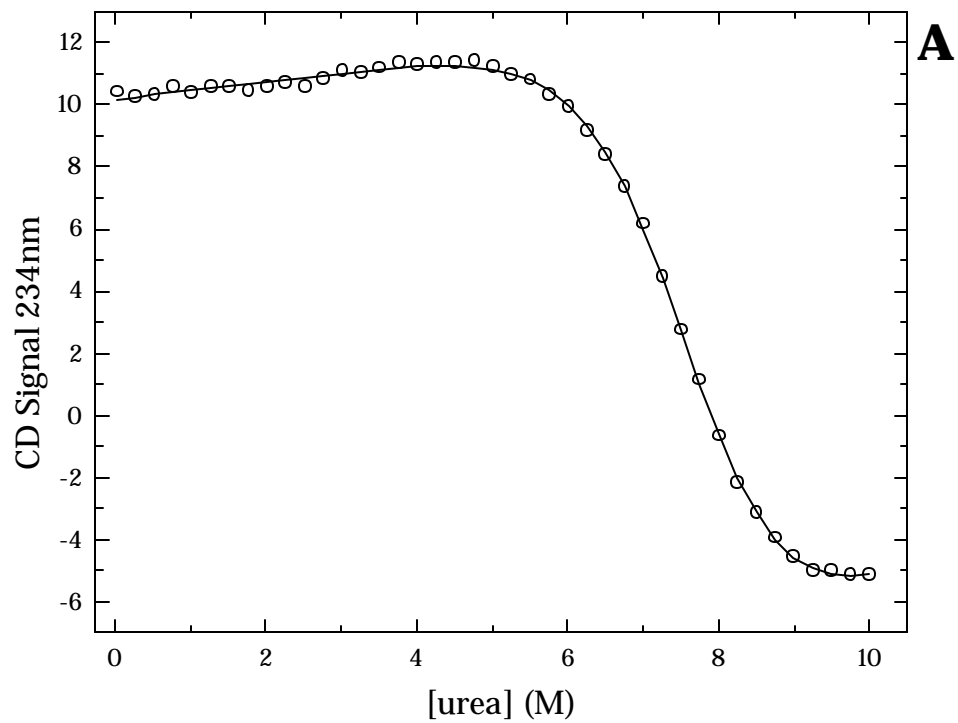
$\Delta C_p$  value is from Pace et al.[Pace, 1998 #47]. The  $\Delta G(25^\circ\text{C})$  value ( $\bullet$ ) was calculated using Equation 5 which is also shown in B.

Figure 4. A thermodynamic cycle connecting protein folding in water and urea. N and D denote the native and denatured states,  $\Delta G$  and  $\Delta G(\text{H}_2\text{O})$  are defined by Equation 3,  $n_i$  is the total number of groups of type  $i$  in the protein,  $\alpha_i$  is fraction of groups of type  $i$  that are exposed in D but not N, and  $\Delta g_{\text{tr},i}$  is the free energy of transfer of groups of type  $i$  from water to a given concentration of urea.  $\Delta\alpha$  is the average change in exposure of all of the peptide groups and uncharged side chains in the protein<sup>31</sup>.

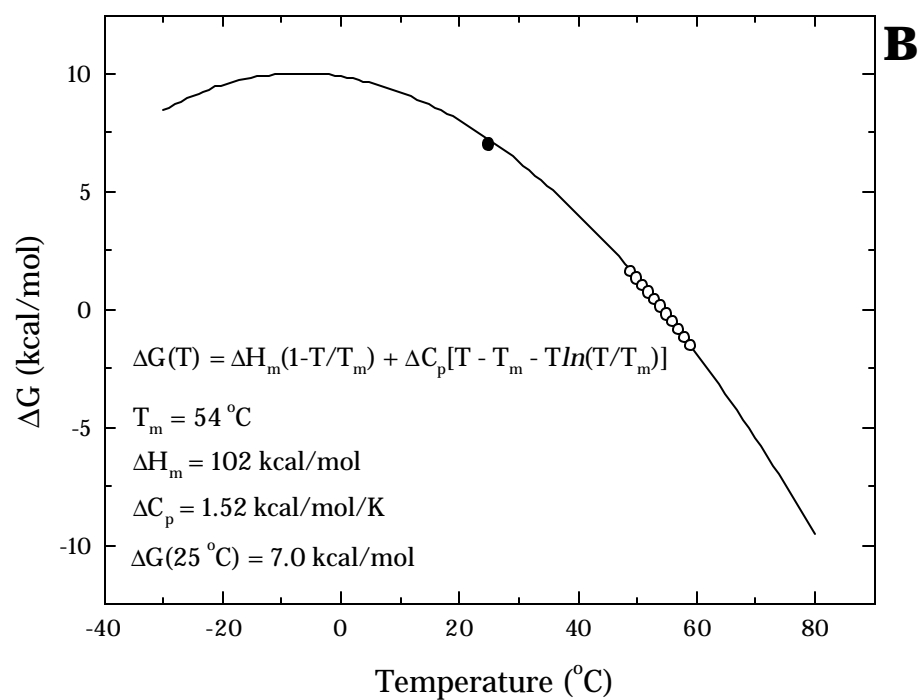
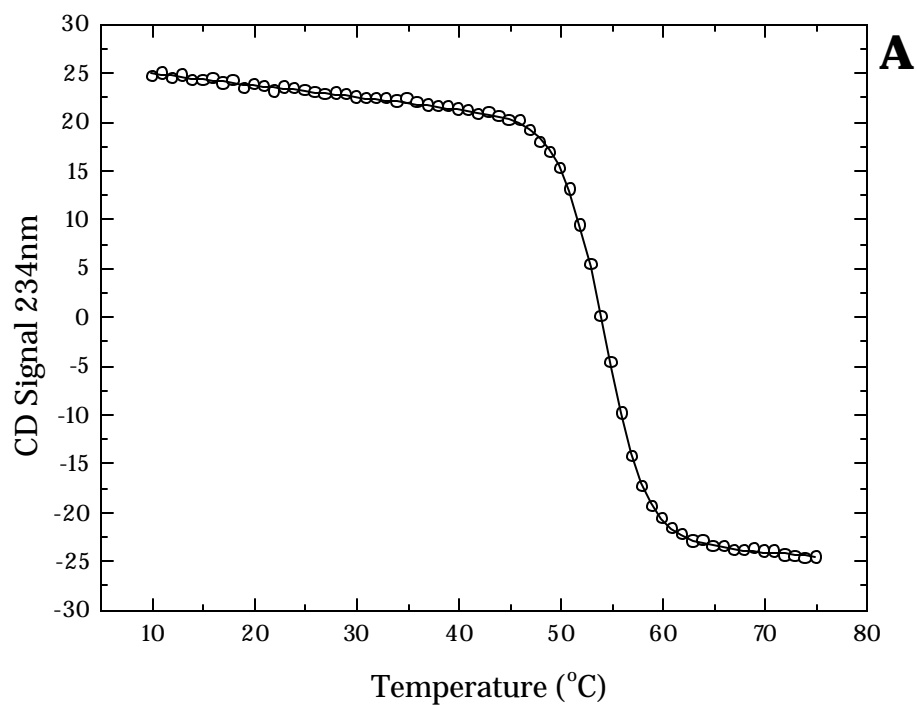


$\Delta G_{tr}(H_2O \rightarrow \text{Solvent})$  (cal/mol)

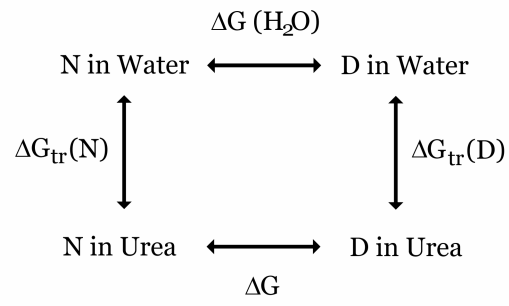
<u>Solvent</u>	<u>Peptide Group</u>	<u>Leu Side Chain</u>
Urea (2M)	-50	-105
GdnHCl (2M)	-85	-150
Ethanol (40%)	+270	-375
Glycerol (20%)	+55	+16
Sucrose (2M)	+52	+74
Sarcosine (2M)	+89	+77
TMAO (2M)	+177	+18



The linear extrapolation method – Figure 2



The linear extrapolation method – Figure 3



$$\begin{aligned}
 \Delta G + \Delta G_{\text{tr}}(\text{N}) &= \Delta G (\text{H}_2\text{O}) + \Delta G_{\text{tr}}(\text{D}) \\
 \Delta G - \Delta G (\text{H}_2\text{O}) &= \Delta G_{\text{tr}}(\text{D}) - \Delta G_{\text{tr}}(\text{N}) = \sum n_i \alpha_i \Delta g_{\text{tr},i} \\
 \Delta G &= \Delta G (\text{H}_2\text{O}) + \sum n_i \alpha_i \Delta g_{\text{tr},i} \\
 \Delta G &= \Delta G (\text{H}_2\text{O}) + (\Delta \alpha) \sum n_i \Delta g_{\text{tr},i}
 \end{aligned}$$