Goodno, C. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 2620-2624.
Goodno, C., & Swenson, C. (1975) *Biochemistry 14*, 873-878.
Gordon, A., & Ford, R. (1972) *The Chemist's Companion*, Wiley, New York.
Griffiths, A., & Trayer, I. (1989) *FEBS Lett. 242*, 275.
Harrington, W., & Ueno, H. (1987) *Biopolymers 26*, S81-S98.
Hvidt, S., Rodgers, M., & Harrington, W. (1985) *Biopolymers 24*, 1647-1662.
Hynes, T., Block, S., White, B., & Spudich, J. (1987) *Cell 48*, 953-963.
Jacobson, A., & Henderson, J. (1973) *Can. J. Biochem. 51*, 71-86.
Kamath, U., & Shriver, J. (1989) *J. Biol. Chem. 264*, 5586-5592.
Kauzman, W. (1959) *Adv. Protein Chem. 14*, 1-63.
Kay, L., Pascone, J., Sykes, B., & Shriver, J. (1987) *J. Biol. Chem. 262*, 1984-1988.
King, L., & Lehrer, S. (1989) *Biochemistry 28*, 3498.
Margossian, S., & Lowey, S. (1982) *Methods Enzymol. 85*, 55-71.
Mornet, D., Pantel, P., Bertrand, R., Audemard, E., & Kassab, R. (1981) *FEBS Lett. 123*, 54-58.
Mornet, D., Ue, K., & Morales, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 736-739.
Muhlrad, A., & Morales, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1003-1007.
Okamoto, Y., & Sekine, T. (1987) *J. Biol. Chem. 262*, 7951-7954.
Privalov, P. (1979) *Adv. Protein Chem. 33*, 167-241.
Privalov, P. (1982) *Adv. Protein Chem. 35*, 1-104.
Privalov, P., Griko, Y., Venyaminov, S., & Kutyshenko, V. (1986) *J. Mol. Biol. 190*, 487-498.
Setton, A., & Muhlrad, A. (1984) *Arch. Biochem. Biophys. 235*, 411-417.
Shriver, J., & Sykes, B. (1981) *Biochemistry 20*, 2004-2012.
Sturtevant, J. (1987) *Annu. Rev. Phys. Chem. 38*, 463-488.
Sutoh, K., Sutoh, K., Karr, T., & Harrington, W. (1978) *J. Mol. Biol. 126*, 1-22.
Swenson, C., & Ritchie, P. (1980) *Biochemistry 19*, 5371-5375.
Takahashi, K., Casey, J., & Sturtevant, J. (1981) *Biochemistry 20*, 4693-4697.
Ueno, H., Rodgers, M., & Harrington, W. (1983) *J. Mol. Biol. 168*, 207-228.
Vibert, P., & Cohen, C. (1988) *J. Muscle Res. Cell Motil. 9*, 296-305.
Wagner, P. (1984) *Biochemistry 23*, 5950-5956.
Weeds, A., & Taylor, R. (1975) *Nature 257*, 54-56.
Weeds, A., & Pope, B. (1977) *J. Mol. Biol. 111*, 129-157.
Zehfus, M. (1987) *Proteins: Struct., Funct., Genet. 3*, 90-110.

pH Dependence of the Urea and Guanidine Hydrochloride Denaturation of Ribonuclease A and Ribonuclease T1†

C. Nick Pace,* Douglas V. Laurents, and James A. Thomson
Biochemistry Department, Texas A&M University, College Station, Texas 77843
Received September 5, 1989; Revised Manuscript Received November 6, 1989

ABSTRACT: To investigate the pH dependence of the conformational stability of ribonucleases A and T1, urea and guanidine hydrochloride denaturation curves have been determined over the pH range 2-10. The maximum conformational stability of both proteins is about 9 kcal/mol and occurs near pH 4.5 for ribonuclease T1 and between pH 7 and 9 for ribonuclease A. The pH dependence suggests that electrostatic interactions among the charged groups make a relatively small contribution to the conformational stability of these proteins. The dependence of ΔG on urea concentration increases from about 1200 cal mol^{-1} M^{-1} at high pH to about 2400 cal mol^{-1} M^{-1} at low pH for ribonuclease A. This suggests that the unfolded conformations of RNase A become more accessible to urea as the net charge on the molecule increases. For RNase T1, the dependence of ΔG on urea concentration is minimal near pH 6 and increases at both higher and lower pH. An analysis of information of this type for several proteins in terms of a model developed by Tanford [Tanford, C. (1964) *J. Am. Chem. Soc. 86*, 2050-2059] suggests that the unfolded states of proteins in urea and GdnHCl solutions may differ significantly in the extent of their interaction with denaturants. Thus, the conformations assumed by unfolded proteins may depend to at least some extent on the amino acid sequence of the protein.

Ribonuclease A (RNase A) from bovine pancreas is one of the most thoroughly studied enzymes, and the energetics (Hermans & Scheraga, 1961; Schwarz & Kirchhoff, 1988) and mechanism (White & Anfinsen, 1959; Wearne & Creighton, 1988; Udgaonkar & Baldwin, 1989) of folding have been studied in detail over the past 30 years. Ribonuclease T1 (RNase T1) from *Aspergillus oryzae* is the next best characterized ribonuclease and has proven to be an excellent model for investigating several aspects of protein folding (Thomson et al., 1989; Shirley et al., 1989; Pace & Laurents, 1989). These enzymes share a function, but they do not resemble one another in amino acid sequence or three-di-

†Supported by grants from the NIH (GM 37039), the Robert A. Welch Foundation (A 1060), and the Texas Agricultural Experiment Station.

0006-2960/90/0429-2564$02.50/0

mensional structure (Heinemann & Saenger, 1982). One interesting difference between the two enzymes is the content and properties of the basic amino acids. RNase A (124 residues) has 10 Lys, 4 Arg, and 4 His, but RNase T1 (104 residues) has only 1 Lys, 1 Arg, and 3 His. RNase T1 seems ill-equipped to tackle a large negatively charged substrate like RNA. Both enzymes have two His residues at the active site: 12 (pK_a = 5.8) and 119 (pK_a = 6.2) for RNase A [see Table II in Matthew and Richards (1982)] and 40 (pK_a = 7.8) and 92 (pK_a = 7.9) for RNase T1 (Inagaki et al., 1981). The much larger pK_a values for the His residues of RNase T1 result in additional positive charges at neutral pH that probably aid in the interaction with RNA. Because of this marked difference in basic amino acid content, RNase T1 is an acidic protein with an isoionic pH of about 4 (Iida & Ooi, 1969), but RNase A is a basic protein with an isoionic pH of 9.6 (Tanford & Hauenstein, 1956). Despite these differences, the pH optimum for the hydrolysis of RNA is 7–7.5 for both enzymes. This background suggested that RNase A and RNase T1 would be an interesting pair of proteins to choose to investigate the pH dependence of the conformational stability.

There has been considerable interest recently in developing theoretical models to calculate the contribution of electrostatic interactions to the conformational stability of globular proteins (Harvey, 1989). Unfortunately, there is very little experimental data available to test the theoretical models. The results reported here should be useful in testing theoretical models for the electrostatic interactions among the ionizable groups on globular proteins.

Tanford and co-workers presented evidence that proteins approach a randomly coiled conformation after unfolding by guanidine hydrochloride (GdnHCl) or urea but that thermal unfolding is less complete (Tanford, 1968, 1970). Both of these proposals have been debated over the years (Privalov et al., 1989). Recent results from Shortle and co-workers have rekindled interest in the unfolded state and its role in the conformational stability of globular proteins (Shortle, 1989). The results presented here provide further information on the interaction of proteins with urea and GdnHCl and on the unfolded states that they produce.

Experimental Procedures

Ribonuclease T1 was either purified from a crude *Aspergillus oryzae* extract obtained from Sigma (product P 4755) (Pace et al., 1987) or prepared from a gene expressed in *Escherichia coli* (Shirley et al., 1989). Ribonuclease A (type XII-A) was purchased from Sigma. The buffers used were the best grades available from Sigma (diglycine, MES, MOPS, HEPPSO) or Fisher (glycine, sodium formate, sodium acetate). "Ultrapure" urea and GdnHCl were purchased from Schwarz/Mann Biotech, and stock solutions were prepared as described previously (Pace, 1986).

Urea and GdnHCl denaturation curves were determined by measuring the intrinsic fluorescence of RNase T1 (278-nm excitation, 320-nm emission) or RNase A (278-nm excitation, 305-nm emission) solutions thermostated at 25 ± 0.1 °C in 1-cm cuvettes with a Perkin-Elmer MPF 44B spectrofluorometer. Denaturation was completely reversible for both proteins for all of the conditions reported here. The determination and analysis of urea and GdnHCl unfolding curves have been discussed in detail elsewhere (Pace et al., 1989). After the fluorescence measurements, about five solutions near the midpoint of the transition were selected, and their pH was measured at room temperature after a double buffer adjustment on a Radiometer Model 26 pH meter. The average of these measurements is the pH given in the tables and figures for individual urea or GdnHCl unfolding curves. Buffers that bind specifically to the folded or unfolded conformations of a protein should be avoided since they will alter the $(\text{denaturant})_{1/2}$ values (Thomson & Bigelow, 1986). For RNase T1, this was checked by measuring $(\text{urea})_{1/2}$ at the same pH for different buffers. In addition, for both RNase A and RNase T1, there is no evidence of discontinuities in the plots of $(\text{urea})_{1/2}$ vs pH (Figures 3A and 4A). Succinate and CHES were not used as buffers because of evidence that they bind to folded RNase T1. The buffers used for RNase A unfolding curves were the following: glycine (2.05–2.86 and 9.40–10.0), formate (3.27–4.36), acetate (4.53–5.36), MES (5.84–6.60), MOPS (6.79–7.60), HEPPSO (8.01), and diglycine (8.72). The buffers used for RNase T1 unfolding curves were as follows: glycine (1.67–2.22 and 9.48–9.99), formate (2.79–4.95), MES (5.56–6.62), MOPS (6.68–7.50), HEPPSO (7.58–8.29), and diglycine (8.53).

Results

For both RNase A and RNase T1, unfolding has been shown to closely approach a two-state mechanism (Freire & Biltonen, 1978; Thomson et al., 1989). This allows the calculation of the free energy of unfolding, ΔG, from an unfolding curve such as that shown in Figure 1A by using the equation:

$$\Delta G = -RT \ln K = -RT \ln [(y_N - y)/(y - y_D)] \quad (1)$$

where K is the equilibrium constant, y is the observed value of the parameter used to follow unfolding, and y_N and y_D are the values of y characteristic of the native and denatured conformations of the protein, respectively (Pace, 1975). The results reported here are based on urea and GdnHCl unfolding curves determined by using intrinsic fluorescence measurements to follow unfolding. One of the better urea unfolding curves for RNase A is shown in Figure 1A. Urea unfolding curves for RNase T1 have been published previously (Pace & Laurents, 1989). The dashed lines extrapolated from the pre- and posttransition regions were used to obtain y_N and y_D in the transition region. The equations shown for y_N and y_D were obtained by a least-squares analysis of data in the pre- and posttransition regions. ΔG calculated by using eq 1 was found to vary linearly with urea concentration, as shown in Figure 1B, and a least-squares analysis (linear regression was performed by using the proc reg subroutine in the SAS/STAT software available from the SAS Institute Inc., Cary, NC) was used to fit the data to the equation:

$$\Delta G = \Delta G(H_2O) - m(\text{urea}) \quad (2)$$

where $\Delta G(H_2O)$ is the value of ΔG in the absence of urea and m is a measure of the dependence of ΔG on urea concentration. The dashed line and the equation given on Figure 1B show the results of this method of analysis for the data shown in Figure 1A.

Santoro and Bolen (1988) have pointed out that this method of analyzing unfolding curves underestimates the error in the parameters determined because no error is assumed for the pre- and posttransition base lines. To get a more realistic assessment of the error, they recommend using nonlinear least squares to directly fit an unfolding curve such as that shown in Figure 1A to the equation:

$$y = \{(y_f + m_f[D]) + (y_u + m_u[D])(\exp[-(\Delta G(H_2O)/RT - m[D]/RT)])\}/\{1 + \exp[-(\Delta G(H_2O)/RT - m[D]/RT)]\} \quad (3)$$

where y_f and m_f, and y_u and m_u are the slope and intercept of the pre- and posttransition base lines, respectively, [D] is

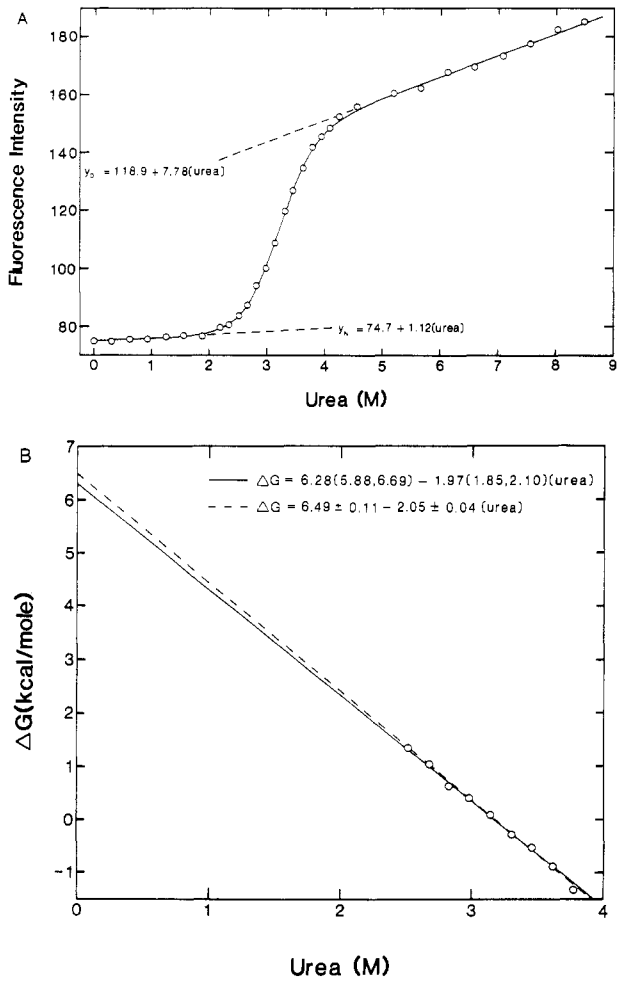

FIGURE 1: (A) Urea unfolding curve for RNase A at pH 3.55, 30 mM formate buffer, 25 °C. The fluorescence intensity at 305 nm was measured after excitation at 278 nm. The dashed lines and equations are based on a least-squares analysis of the pre- and posttransition base lines. The solid curve is based on the parameters derived from a nonlinear least-squares fit of the experimental data to eq 3. The parameters describing the pre- and posttransition base lines from eq 3 and their 67% confidence limits are as follows: y_f = 74.8 (74.0, 75.5), m_f = 0.96 (0.32, 1.61), y_u = 121.1 (119.2, 123.0), and m_u = 7.48 (7.2, 7.8). (B) ΔG as a function of urea concentration. ΔG was calculated from the data in (A) using eq 1 and then analyzed by a least-squares fit of eq 2 (dashed line), or the data in (A) were analyzed directly by a nonlinear least-squares fit of eq 3 (solid line). The standard errors are given for the parameters describing the dashed line, and the 67% confidence limits are given for the parameters describing the solid line.

the denaturant molarity, and $\Delta G(H_2O)$ and m were defined above. These six parameters were determined by using the nonlinear least-squares program described by Johnson and Frazier (1985). The solid lines shown in Figure 1A,B are based on the parameters determined in this way; 67% confidence limits for the values of $\Delta G(H_2O)$ and m are given with the equation that describes the solid line in Figure 1B. The error bars representing the m values in Figures 3B and 4B (below), and those representing the $\Delta G(H_2O)$ values in Figure 6A,B (below), also give the 67% confidence limits.

GdnHCl unfolding curves for RNase A and RNase T1 were determined at pH 3, 7, and 10. The curves for RNase T1 at pH 7 and 10 are shown in Figure 2. Note that the intrinsic fluorescence increases between 0 and 1 M GdnHCl in the curve determined at pH 7. For reasons explained below, we think this results from the binding of a guanidinium ion to a cation binding site on folded RNase T1. To analyze this unfolding curve, we eliminated the points below 1 M GdnHCl.

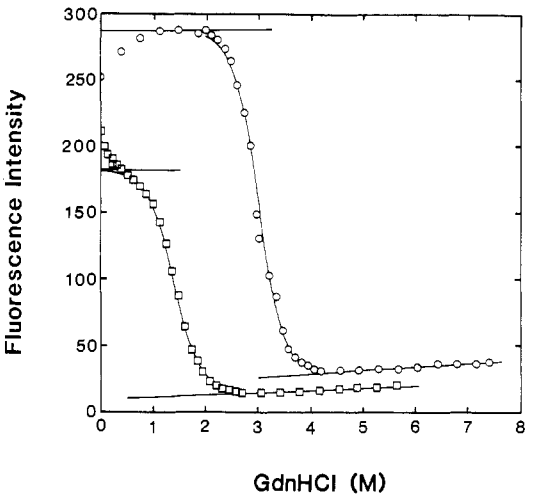

FIGURE 2: GdnHCl unfolding curves for RNase T1 at 25 °C in either 30 mM MOPS buffer, pH 6.98 (○), or 30 mM glycine buffer, pH 9.91 (□). The fluorescence intensity at 320 nm was measured after excitation at 278 nm. The solid lines in the transition region are based on the parameters given in Table I.

Table I: Parameters Characterizing the GdnHCl Unfolding of RNase A and RNase T1 at pH 3, 7, and 10[a]

pH	m[b] (cal mol^{-1} M^{-1})	$\Delta G(H_2O)$[b] (kcal/mol)	$(GdnHCl)_{1/2}$[b] (M)
		RNase A	
3.02	2910 (32)	5.29 (0.06)	1.82
	2885 (221)	5.20 (0.42)	1.80
		5.2 (urea)	
6.98	2930 (52)	8.73 (0.16)	2.98
	3090 (223)	9.24 (0.64)	2.99
		9.2 (urea)	
9.89	3285 (69)	8.41 (0.18)	2.56
	3270 (290)	8.40 (0.78)	2.57
		8.6 (urea)	
		RNase T1	
2.92	2865 (56)	8.02 (0.16)	2.80
	2890 (118)	8.04 (0.32)	2.78
		6.2 (urea)	
6.98	2540 (56)	7.62 (0.17)	3.00
	2580 (113)	7.70 (0.32)	2.98
		5.7 (urea)	
9.91	2520 (25)	3.40 (0.04)	1.35
	2580 (38)	3.52 (0.05)	1.36
		2.9 (urea)	

[a] At each pH, the top line gives the results of a least-squares analysis of plots of ΔG vs [denaturant] (eq 2) with the standard errors given in parentheses, and the second line gives the results obtained by using nonlinear least squares to fit the entire unfolding curve directly to eq 3 with the 67% confidence limits given in parentheses. The $\Delta G(H_2O)$ values given in the third line are results from urea unfolding from Figure 6. [b] Parameters of eq 2. $(GdnHCl)_{1/2}$ is the midpoint of the GdnHCl unfolding curve [$=\Delta G(H_2O)/m$].

For the curve at pH 10, a marked decrease in the intrinsic fluorescence is observed at low GdnHCl concentrations. Again, this may be due to guanidinium ion binding, but now the effect on the intrinsic fluorescence is different because the neighboring Tyr residues are partially ionized. To analyze this curve, we used a value of 182 for the fluorescence intensity of the folded protein, as shown on the figure, and again eliminated the points at low GdnHCl concentrations. This value led to the best fit of the data using the nonlinear least-squares program. The other GdnHCl unfolding curves were analyzed in exactly the same way as the urea unfolding curves described above, and all of the results are summarized in Table I. For comparison, we have given both the results of the linear least-squares fit to eq 2 and the results of the nonlinear least-squares fit to eq 3 in the table.

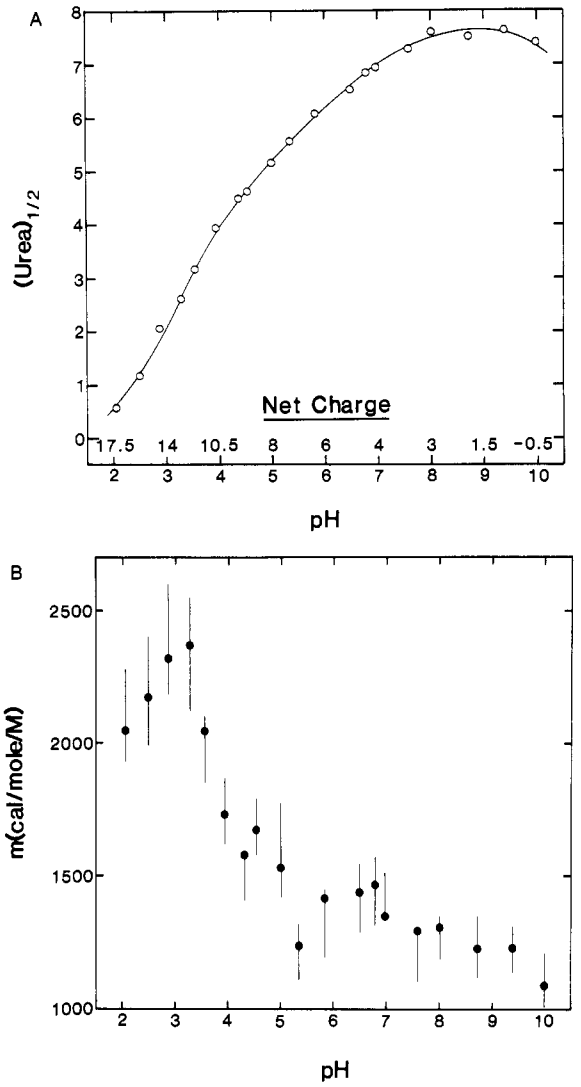

FIGURE 3: $(Urea)_{1/2}$ and m as a function of pH for the unfolding of RNase A. The circles represent the values obtained by analyzing the data using eq 2 as described in the text. (A) The solid line has no theoretical significance. The net charges given are based on the titration curve of RNase A reported by Tanford and Hauenstein (1956). (B) The error bars given for each m value represent the 67% confidence limits determined by analyzing the data using eq 3 as described in the text.

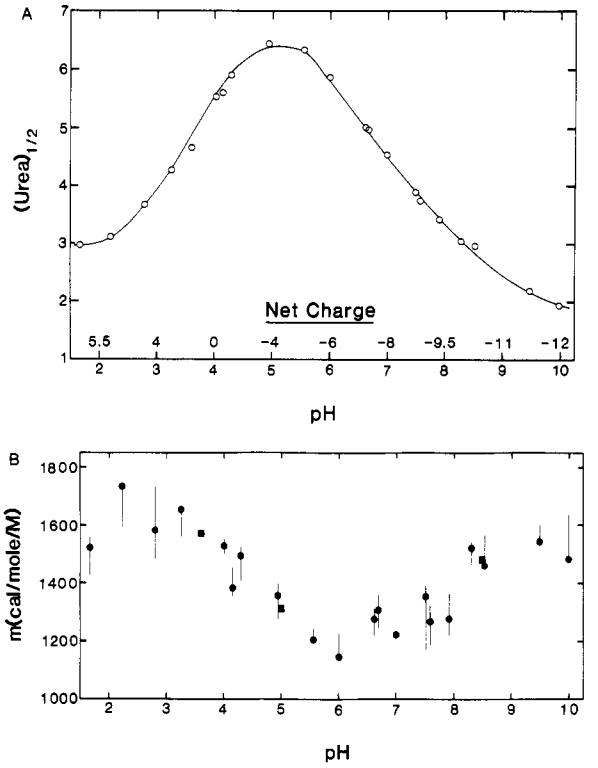

FIGURE 4: $(Urea)_{1/2}$ and m as a function of pH for the unfolding of RNase T1. The circles represent the values obtained by analyzing the data using eq 2 as described in the text. (A) The solid line has no theoretical significance. The net charges given are based on the titration curve of RNase T1 reported by Iida and Ooi (1969). (B) The error bars given for each m value represent the 67% confidence limits determined by analyzing data using eq 3 as described in the text. The circle without an error bar is from Thomson et al. (1989), and the squares without error bars are from Grimsley (1988).

The estimate of $\Delta G(H_2O)$ derived from the analysis of urea unfolding at each pH (from Figure 6) is also given in Table I for comparison with the values from the analysis of GdnHCl unfolding. Note that the $\Delta G(H_2O)$ values from urea and GdnHCl are in excellent agreement for RNase A, as previously noted by Ahmad and Bigelow (1982). However, for RNase T1, the estimates of $\Delta G(H_2O)$ derived from the GdnHCl unfolding experiments are greater than those from the urea unfolding experiments. We think this is the result of a specific binding site for a guanidinium ion on folded RNase T1. A binding constant of about 11 M^{-1} would be sufficient to account for the greater value of $\Delta G(H_2O)$ observed in the GdnHCl unfolding experiments. We showed previously that the stabilization of RNase T1 by salts results in part from the existence of a relatively nonspecific cation binding site on the folded conformation of the protein (Pace & Grimsley, 1988). This has since been confirmed by X-ray crystallography, where a Ca^{2+} ion was found to bind at a site where the most important interactions are with the carboxyl oxygens of Asp-15 (Kostrewa et al., 1989). The increase in the intrinsic fluorescence at low GdnHCl concentrations at pH 7 noted in Figure 2 is also observed when other cations bind. Thus, it seems likely that guanidinium ions also bind at the cation binding site.

Discussion

Two parameters can be used to characterize the transition region of a urea unfolding curve: $(urea)_{1/2}$ is a measure of the midpoint of the transition region, and m is a measure of the steepness of the transition region. These parameters are plotted as a function of pH for RNase A in Figure 3, and for RNase T1 in Figure 4. An estimate of the net charge is given at each pH from titration studies of RNase A (Tanford & Hauenstein, 1956) and RNase T1 (Iida & Ooi, 1969). It is clear that $(urea)_{1/2}$ can be measured more accurately, $\approx$0.05 M, than the m value. RNase A is most stable to urea near pH 9, which is slightly below pH 9.6 where the net charge is zero. RNase T1 is most stable to urea near pH 5, which is about 1 pH unit higher than the pH where the net charge is zero. For both proteins, the pH dependence of T_m, the midpoint of the thermal unfolding curve, is qualitatively similar to the pH dependence we observe for $(urea)_{1/2}$ (Hermans & Scheraga, 1961; Oobatake et al., 1979). In contrast, T4 lysozyme is most stable to thermal unfolding near pH 5.5, but it is a basic protein with an isoelectric point near 10 (Becktel & Baase, 1987). Still, T4 lysozyme is likely to be the exception, and the familiar assertion that proteins will be most stable near their isoelectric pH is probably reasonable.

We noticed previously that the m value for the unfolding of RNase A by urea measured at pH 2.8 was considerably larger, 2325 cal mol^{-1} M^{-1} (Pace & Laurents, 1989), than m values measured near neutral pH: 1100 cal mol^{-1} M^{-1} (Greene

& Pace, 1974); 1130 cal mol^{-1} M^{-1} (Pace & Marshall, 1980); 1140 cal mol^{-1} M^{-1} (Ahmad & Bigelow, 1982). This surprising observation prompted the research reported here. It is clear from Figure 3B that the m value for RNase A does indeed increase substantially as the pH is lowered. With RNase T1, on the other hand, m is minimal near pH 6 and increases at both higher and lower pH (Figure 4B). We will now consider what might give rise to the observed dependence of the m value on pH.

One possibility is that the variation of the m values reflects deviations from a two-state folding mechanism with pH. The presence of stable, partially unfolded intermediates at equilibrium would generally lower the m value compared to the value expected for a two-state mechanism (Pace, 1975). This would require that both RNase A and RNase T1 deviate significantly from a two-state mechanism near pH 7, but there is good evidence that this is not the case for either protein (Freire & Biltonen, 1978; Salahuddin & Tanford, 1970; Thomson et al., 1989). Thus, we consider this possibility very unlikely.

From solubility studies on amino acids and related model compounds, the free energies of transfer from water to various concentrations of urea and GdnHCl have been calculated for the component parts of a protein (Tanford, 1970). Qualitatively, it is easy to see why proteins unfold in the presence of urea and GdnHCl: all of the constituent parts of a protein are more soluble in urea and GdnHCl solutions than they are in water. Tanford (1964, 1970) has shown that these data can be used to account quantitatively for the unfolding of proteins by urea and GdnHCl. Tanford's model has been used to estimate $\Delta G(H_2O)$ (Pace, 1986), and for other purposes (Tanford, 1970; Pace, 1975; Greene & Pace, 1974; Ahmad & Bigelow, 1986). We will use this model here to discuss the observed pH dependence of m. The basic relationship is

$$\Delta G = \Delta G(H_2O) + \sum_i (\Delta\alpha_i) n_i \Delta g_{tr,i} \qquad (4)$$

where $\Delta g_{tr,i}$ is the free energy of transfer of a group of type i from water to denaturant, n_i is the total number of groups of type i in the protein, $\Delta\alpha_i$ is the average fractional change in the degree of exposure of groups of type i when the protein unfolds, and ΔG and $\Delta G(H_2O)$ were explained above. From the standpoint of the protein, then, the key factor which will determine the m value is the number and type of groups which are freshly exposed to solvent when the protein unfolds. The $\Delta g_{tr,i}$ values for a peptide group and 14 uncharged amino acid side chains have been measured [see Pace (1986)].[1] Assigning a $\Delta\alpha_i$ value for a given type of residue is difficult. Lee and Richards developed a method for estimating the accessibility of groups in a folded protein based on the three-dimensional structure (Richards, 1977). When this method is used, the fraction buried for the peptide groups and for the 14 uncharged side chains is 0.78 and 0.68 for RNase A and 0.70 and 0.65 for RNase T1, respectively. However, we can only guess at the accessibility of groups in the unfolded conformations of a protein because our understanding of the states that result after unfolding by urea or GdnHCl is limited. For this reason, we will use eq 4 in this form:

$$\Delta G = \Delta G(H_2O) + (\Delta\alpha) \sum_i n_i \Delta g_{tr,i} \qquad (5)$$

where $\Delta\alpha$ is the average change in degree of exposure for all of the peptide groups and the 14 side chains. The model compound data were used to calculate $\sum_i n_i \Delta g_{tr,i}$ as a function of urea and GdnHCl concentration for RNase T1, and the values are shown by the data points in Figure 5 (Δg_{tr} values were measured at only four denaturant concentrations). The solid curves were calculated by using parameters and equations given by Ahmad and Bigelow (1986) for representing the model compound data,[2] and they fit the four data points reasonably well. These curves show the predicted dependence of ΔG on denaturant concentration assuming that all of the peptide groups and 14 uncharged side chains are inaccessible to solvent in the folded protein and completely accessible to solvent in the unfolded protein. An example of the measured dependence of ΔG on denaturant concentration is shown in

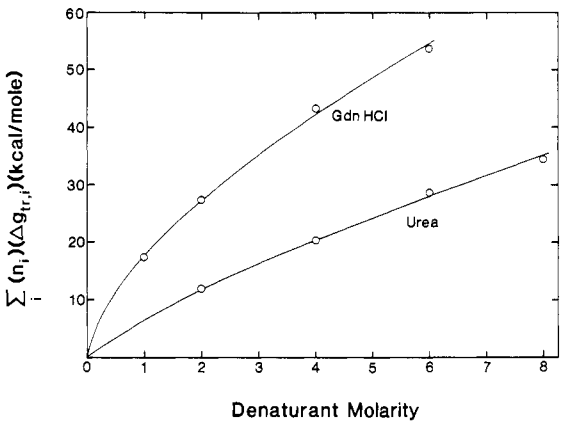

FIGURE 5: $\sum_i n_i \Delta g_{tr,i}$ (eq 5) as a function of denaturant molarity for RNase T1. The n_i values were taken from Dayhoff (1972). The values shown by the open circles were calculated by using the $\Delta g_{tr,i}$ values given by Pace (1986). The solid lines were calculated by using the $\Delta g_{tr,i}$ values from equations given by Ahmad and Bigelow (1986).

Table II: Change in Accessibility ($\Delta\alpha$) Calculated from the Measured m Values Using Tanford's Model

protein	M_r	disulfides	$\Delta\alpha$(urea)[a]	$\Delta\alpha$(GdnHCl)[a]
lysozyme	14 300	4	0.23	0.20
RNase A	13 700	4	0.35	0.35
RNase T1	11 100	2	0.33	0.35
RNase T1	11 100	0	0.38	
barnase	12 400	0	0.53	0.50
nuclease	16 800	0	0.44	0.43
apo Mb	16 900	0	0.46	
DHFR	17 700	0	0.32	
T4 lyso	18 600	0		0.39

[a] $\Delta\alpha$ was calculated from the measured m values and plots such as those shown in Figure 5, as described in the text.[1,2] The amino acid composition for each protein was from Dayhoff (1972). The measured m values were from the following references: lysozyme (Greene & Pace, 1974); RNase A and RNase T1 (this work, pH 7); RNase T1 (zero disulfide bonds) (Pace et al., 1988); barnase (Kellis et al., 1989); nuclease (Shortle & Meeker, 1986); apo Mb (Hughson & Baldwin, 1989); DHFR (Garvey & Matthews, 1989); T4 lysozyme (Chen & Schellman, 1989). See the references for the conditions used to measure m.

[1] Values of $\Delta g_{tr,i}$ for the side chains of Ser, Asp, Glu, Lys, and Arg were not determined (Nozaki & Tanford, 1963, 1970). The thought was that these side chains would make a small contribution either because the Δg_{tr} values would be low (Ser) or because the side chains are generally exposed in the folded conformation of the protein. The side chains of the charged amino acids are actually buried to about the same extent as those of the polar amino acids (Rose et al., 1985). If the charged amino acids are included in the calculations with Lys and Arg added as Val residues and Asp and Glu added as Ala residues, the values of $\Delta\alpha$ would be reduced by $\approx$0.03. This would not change any of the conclusions based on the $\Delta\alpha$ values.

[2] There is an error in Table I in the paper by Ahmad and Bigelow (1986) for the parameters characterizing Δg_{tr} to urea for a peptide group. The correct parameters are a = 0.030 and b = 0.727.

Table III: Change in Accessibility ($\Delta\alpha$) Calculated from the Measured m Values Using Tanford's Model

pH	Z^a	$\Delta\alpha$(urea)^b	$\Delta\alpha$(GdnHCl)^b
		RNase A	
3	14	0.45	0.28
7	4	0.35	0.35
10	0	0.28	0.36
		RNase T1	
3	4	0.41	0.38
7	−8	0.33	0.35
10	−12	0.32	0.25

[a] The approximate net charge was taken from the titration curves for RNase A (Tanford & Hauenstein, 1956) and RNase T1 (Iida & Ooi, 1969). [b] $\Delta\alpha$ was calculated from the measured m values and plots such as those shown in Figure 5, as described in the text.

Figure 1B. Thus, $\Delta\alpha$ can be estimated by dividing the measured m value by the slope of plots such as those shown in Figure 5. Note that the slope varies with denaturant concentration and should be measured at the $(\text{denaturant})_{1/2}$ value from the data used to determine the m value. Values of $\Delta\alpha$ calculated in this way for a number of the proteins under active study at present in the protein folding field are listed in Table II.

Results such as those shown for lysozyme and RNase A led us to conclude that the number and location of the disulfide bonds can significantly affect the accessibility of the unfolded conformations of a protein in urea and GdnHCl solutions (Greene & Pace, 1974). More direct proof based on solvent perturbation studies of proteins in 8 M urea with and without their disulfide bonds broken was reported earlier by Herskovits (1965). The results for RNase T1 provide further support for this idea. The $\Delta\alpha$ value increases from 0.33 to 0.38 when the two disulfide bonds are cleaved.

The results for the six proteins with no disulfide bonds are quite surprising. RNase T1 and barnase are related proteins with similar folded conformations (Hill et al., 1983). Nevertheless, $\Delta\alpha$ is substantially higher for barnase than for RNase T1. This indicates that the solvent accessibility is greater in unfolded barnase than in unfolded RNase T1. Even more striking, the $\Delta\alpha$ value for barnase is $\approx$60% greater than the $\Delta\alpha$ value for DHFR. The differences in the $\Delta\alpha$ values for the proteins with no disulfide bonds cannot be explained by differences in the accessibility of groups in the folded proteins and are unlikely to result from deviations from a two-state folding mechanism. For all of the proteins, $\Delta\alpha$ is less than the fraction of these groups buried in the folded protein estimated by Lee and Richard's method. This suggests that the constituent groups in a urea or GdnHCl unfolded protein are less accessible to solvent than they are in the model compounds used to measure the Δg_{tr} values. This unfolded accessibility must vary from protein to protein to account for the differences in the $\Delta\alpha$ values. We conclude, then, that the unfolded states of proteins in urea and GdnHCl solutions may differ significantly in their ability to interact with denaturants. This is difficult to reconcile with the idea that proteins approach a randomly coiled conformation after unfolding in urea and GdnHCl. Thus, the conformations assumed by an unfolded protein may depend, at least to some extent, on the amino acid sequence of the protein, as does the three-dimensional structure of the folded conformation.

The same approach has been used to calculate $\Delta\alpha$ from the urea and GdnHCl data at pH 3, 7, and 10, and the values are given in Table III. The most interesting result is the large increase in $\Delta\alpha$(urea) with decreasing pH observed with RNase A. To investigate this further, we measured a difference spectrum between two RNase A samples in 9 M urea that differed only in pH. A typical Tyr difference spectrum was observed indicating that some Tyr residues were more exposed to solvent at low pH than at neutral pH. Similarly, Tsong (1975) showed that cytochrome c, also a basic protein ($pI >$ 10), retains residual structure in 9 M urea at neutral pH which can be unfolded by lowering the pH to 4. This unfolding results in an increase in the intrinsic viscosity from 15 to 21 mL/g. Our present explanation for the increase in the m value with decreasing pH for RNase A is that the unfolded conformations become more accessible to urea because of an expansion of the molecules resulting from repulsive electrostatic interactions among the positive charges.[3] However, this is not in line with the GdnHCl results for RNase A. The change in $\Delta\alpha$(GdnHCl) with pH is less than half as large as and in the opposite direction to $\Delta\alpha$(urea). Also, $\Delta\alpha$ estimates from urea and GdnHCl are generally in good agreement (Table II; Greene & Pace, 1974; Ahmad & Bigelow, 1986), but the estimates at pH 3 and 10 for RNase A differ substantially, as do the results for RNase T1 at pH 10. (Keep in mind that the GdnHCl results at pH 10 for RNase T1 are uncertain because we are not sure what gives rise to the unusual unfolding curve shown in Figure 2.) At present, we do not have a good explanation for these observations, but experiments are underway that may provide some answers. The much higher ionic strength of the GdnHCl solutions would be expected to suppress effects due to electrostatic interactions among charged groups on the proteins. Also, since GdnHCl is a salt and urea is not, the basic mechanisms by which they solubilize the constituent groups of a protein may differ. It is interesting, and perhaps reassuring, that even though the m values from urea unfolding vary twice as much with pH as those from GdnHCl, the $\Delta G(H_2O)$ values from the two denaturants are in excellent agreement at pH 3, 7, and 10 for RNase A (Table I).

As discussed under Results, we think folded RNase T1 binds a guanidinium ion, probably at the nonspecific cation binding site previously reported (Pace & Grimsley, 1988). This would account for the larger estimate of $\Delta G(H_2O)$ derived from GdnHCl than from urea unfolding experiments. Specific binding sites for the denaturant on either the folded or the unfolded conformations of a protein will also affect the measured m values. For example, the m value for the GdnHCl unfolding of RNase T1 would be expected to be lowered by about 170 cal mol^{-1} M^{-1} by the existence of a site on folded RNase T1 that binds a guanidinium ion with a binding constant of 11 M^{-1}. This possibility must be kept in mind as an explanation for differences in the m values between proteins differing slightly in amino acid sequence (Shortle & Meeker, 1986; Shirley et al., 1989), for the nonlinearity observed at low concentrations in plots of ΔG vs denaturant concentration (Shortle et al., 1989; Pace & Vanderberg, 1979), and also for the pH dependence of m values as reported here.

Tanford's model can also be used to estimate $\Delta G(H_2O)$. Note from eq 4 that $\Delta G(H_2O)$ is equal to the value of $\Delta\alpha\sum_i n_i\Delta g_{tr,i}$ evaluated at $(\text{denaturant})_{1/2}$, where $\Delta G = 0$. The values of $\Delta G(H_2O)$ from Tanford's model are always higher than those from linear extrapolation by an average of 1.6

[3] RNase A is a basic protein, and the positive charge of 19 at low pH is unusually high. In addition, the positive charges tend to be concentrated. Five of the positive charges are located in the first 26 residues, and 5 and 6 positive charges, respectively, are located in loops of 28 and 34 residues formed by the disulfide bonds. Expansion due to electrostatic repulsion among the charges in these loops might increase the accessibility as does breaking the disulfide bonds.

kcal/mol for urea and 4.0 kcal/mol for GdnHCl for the results presented in Tables II and III. The marked nonlinearity observed in the plots of $\sum_i n_i g_{tr,i}$ vs denaturant concentration (Figure 5) may be the main reason for this.[4] The observed dependence of ΔG on denaturant concentration is generally linear (Figure 2), but the measurements typically cover a narrow enough concentration range that nonlinearity would be difficult to observe. For RNase T1, for example, the m value of 1220 cal mol^{-1} M^{-1} at 4.48 M urea (pH 7 data) would be expected to be 1190 and 1285 cal mol^{-1} M^{-1} (urea)$_{1/2}$ values 1 M higher and lower than 4.48 M, based on the results in Figure 5. Likewise, the m value of 2560 cal mol^{-1} M^{-1} at 2.99 M GdnHCl would be expected to be 2305 and 2975 cal mol^{-1} M^{-1} (GdnHCl)$_{1/2}$ values 1 M higher and lower than 2.99 M. Only in a few cases have measurements been made at low concentrations of denaturant where the model compound data indicate that the nonlinearity will be greatest. For GdnHCl, linearity was observed with three unstable mutant forms of staph nuclease (Shortle et al., 1989), but nonlinearity was observed with myoglobin (Pace & Vanderburg, 1979). For urea, nonlinearity was observed with the three staph nuclease mutants, but the deviation was opposite in direction to that shown in Figure 5; i.e., the m value decreased at low concentrations. For myoglobin (Pace, unpublished observations), for RNase A (Brandts & Hunt, 1967), and for RNase T1 destabilized by breaking the disulfide bonds (Pace et al., 1988), no evidence of nonlinearity was observed at low urea concentrations.

This question is important because at present more estimates of $\Delta G(H_2O)$ are based on a linear extrapolation. Schellman (1978, 1987) has argued convincingly that a linear dependence is reasonable based on logical and theoretical considerations, but he has not considered the nonlinearity observed in plots such as those shown in Figure 5. The curvature observed at low concentrations of denaturant would be expected if the enhanced solubility results in part from denaturant binding (Pace & Vanderburg, 1979), but the curvature is observed even with side chains such as leucine where denaturant binding does not offer a plausible explanation for the solubility enhancement (Tanford, 1970). This is surely an area that deserves further study. Schellman[5] is currently testing extrapolations based on activities or concentration scales other than molarity to see if more reasonable results are obtained. At present, it still seems most reasonable to estimate $\Delta G(H_2O)$ as is done in Figure 1B using a molarity-based linear extrapolation.

The pH dependence of $\Delta G(H_2O)$ is shown for RNase A in Figure 6A and for RNase T1 in Figure 6B. The pH dependence of the conformational stability of a globular protein will depend mainly on two related factors: (1) electrostatic interactions involving the charged groups, and (2) differences between the pK values of the ionizable groups in the folded and unfolded conformations. The difference between the sum of the attractive and repulsive interactions in the folded and unfolded conformations of the protein will surely contribute to the conformational stability. In general, the interactions

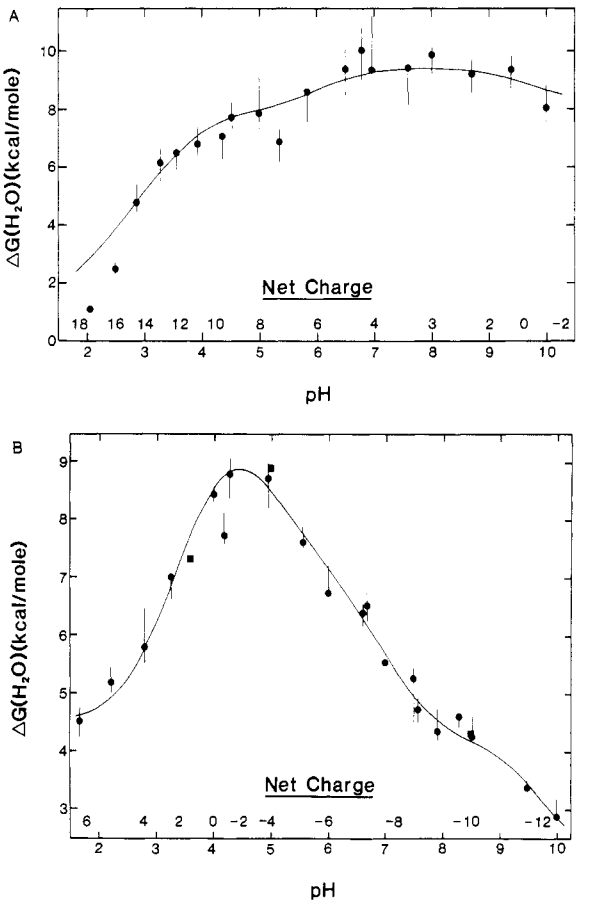

FIGURE 6: $\Delta G(H_2O)$ as a function of pH for the unfolding of RNase A (A) or RNase T1 (B). The circles, error bars, and net charge have the same significance as in Figures 3 and 4. The solid lines were calculated with eq 6 by using the following pK values for the folded conformation: RNase A, two Asp (pK's = 1.5), one Glu (pK = 4.8), two His (pK's = 5.4 and 6.2), and three Tyr (pK's = 9.9); RNase T1, three Asp (pK's = 2.7), two Glu (pK's = 5.5 and 5.7), two His (pK's = 7.2 and 7.8), five Tyr (pK's = 9.9). The following pK values were used for the unfolded conformation: Asp (pK = 4.1), Glu (pK = 4.5), His (pK = 6.4), and Tyr (pK = 9.6) from Nozaki and Tanford (1967a,b).

in the folded protein will be much larger because the charges are closer together, and the effective dielectric constant between the charges will be lower (Tanford, 1961; Imoto, 1983). These contributions are difficult to calculate, in part because of the problem of estimating the effective dielectric constant (Fersht & Sternberg, 1989).

The results in Figure 6 show that the maximum conformational stability for both proteins is about 9 kcal/mol and that the gain in conformational stability in going from the most unfavorable net charge is ≈8 kcal/mol for RNase A (from +18 at pH 2) and ≈6 kcal/mol for RNase T1 (from −12 at pH 10). Obviously, the gain in stability for bringing the charges from far apart to the arrangement that exists on the folded protein at the pH of maximum stability should be considerably less than these values of 8 and 6 kcal/mol. Thus, these results indicate that electrostatic interactions among the charged groups make a relatively small contribution to the conformational stability of RNases A and T1, probably much less than the contributions of hydrogen bonding and hydrophobic interactions. A similar conclusion was reached for RNase A on the basis of experimental studies (Hollecker & Creighton, 1982), and on the basis of theoretical calculations (Matthew & Richards, 1982). In fact, it is encouraging to note that the measured difference in stability between pH 2 and 7 for RNase A (≈8 kcal/mol) is almost identical with the

[4] Both Schellman (1987) and Matouschek et al. (1989) suggest that the free energies of transfer are approximately linear functions of denaturant concentration. The three figures in Tanford (1970) give this impression, but this is clearly not the case when all of the contributions are summed as shown in Figure 5. Since these calculations are dominated by the Δg_{tr} values for a peptide group, it might be worthwhile to repeat the Δg_{tr} measurements using other models for the peptide groups in a protein. It might also be worthwhile to make careful measurements as a function of denaturant concentration at low concentrations of denaturant for a variety of the model compounds.

[5] J. A. Schellman, personal communication.

value from the theoretical calculations.

An estimate of the number of protons released or bound on unfolding can be obtained from the slope of plots such as those shown in Figure 6 (Hermans & Scheraga, 1961; Tanford, 1970). For RNase A, the maximum slope occurs between pH 2 and 3 and indicates that about 2.2 protons are released on unfolding. For RNase T1, a maximum of about 1.7 protons are released on unfolding at acid pH, and a maximum of about 1.2 protons are taken up on unfolding above pH 4.5. This suggests that it might be possible to represent the pH dependence of unfolding in terms of a relatively small number of groups that have different pK's in the folded and unfolded conformations. Groups that bind protons more tightly (higher pK) in the folded conformation than in the unfolded conformations will stabilize the protein as the H^+ concentration increases. NMR titrations have shown that the pK's of the His residues are 7.9 (His-92), 7.8 (His-40), and 7.2 (His-27) in folded RNase T1 (Inagaki et al., 1981) and a pK of about 6.4 would be expected for the His residues in the unfolded protein (Nozaki & Tanford, 1967b). This alone would account for the 3 kcal/mol decrease in $\Delta G(H_2O)$ that occurs between pH 6 and 9 for RNase T1. In line with this reasoning, we have assumed that the pH dependence of $\Delta G(H_2O)$ results entirely from pK differences such as these in order to calculate the solid curves shown in Figure 6. The curves were calculated by using (Tanford, 1970)

$$\Delta G(H_2O) = \Delta G(H_2O)([H^+] = \infty) - RT\sum_i \ln \{[1 + K_{i,D}/(H^+)]/[1 + K_{i,N}/(H^+)]\} \quad (6)$$

where $\Delta G(H_2O)([H^+] = \infty)$ is $\Delta G(H_2O)$ for the fully protonated protein, $K_{i,D}$ and $K_{i,N}$ are the dissociation constants of group i in the denatured and native conformations, respectively, and (H^+) is the hydrogen ion activity. The set of pK's used to calculate the solid curves in Figure 6 are given in the figure legend. Note that the data below pH 3 for RNase A cannot be fit by using this approach. This suggests that electrostatic interactions among the charged groups must be taken into account at low pH for RNase A, which is not surprising given the large positive charge and the low ionic strength. The set of pK's given is not unique, but they will obviously account for most of the observed pH dependence of $\Delta G(H_2O)$.

In summary, we are sure of two things: (1) the experimental results are correct, and (2) we do not understand all of the experimental results. Perhaps the most important conclusion that should be drawn from the results is that we still have much to learn about the unfolded conformations of proteins and about the interactions of proteins with denaturants.

ACKNOWLEDGMENTS

We thank Don Pettigrew for help in computing and Lisa Lohman for help in preparing the figures.

Registry No. RNase A, 9001-99-4; RNase T_1, 9026-12-4; urea, 57-13-6; guanidine hydrochloride, 50-01-1.

REFERENCES

Ahmad, A., & Bigelow, C. C. (1982) *J. Biol. Chem. 249*, 12935.
Ahmad, A., & Bigelow, C. C. (1986) *Biopolymers 25*, 1623.
Becktel, W. J., & Baase, W. A. (1987) *Biopolymers 26*, 619.
Brandts, J. F., & Hunt, L. (1967) *J. Am. Chem. Soc. 89*, 4826.
Chen, B., & Schellman, J. A. (1989) *Biochemistry 28*, 685.
Dayhoff, M. O., Ed. (1972) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington, D.C.
Fersht, A. R., & Sternberg, M. J. (1989) *Protein Eng. 2*, 527.
Freire, E., & Biltonen, R. L. (1978) *Biopolymers 17*, 463.
Garvey, E. P., & Matthews, C. R. (1989) *Biochemistry 28*, 2083.
Greene, R. F., Jr., & Pace, C. N. (1974) *J. Biol. Chem. 249*, 5388.
Grimsley, G. R. (1988) Ph.D. Dissertation, Texas A&M University, College Station, TX.
Harvey, S. C. (1989) *Proteins: Struct., Funct., Genet. 5*, 78.
Heinemann, U., & Saenger, W. (1982) *Nature 299*, 27.
Hermans, J., Jr., & Scheraga, H. A. (1961) *J. Am. Chem. Soc. 83*, 3283.
Herskovits, T. T. (1965) *J. Biol. Chem. 240*, 628.
Hill, C., Dodson, G., Heinemann, U., Saenger, W., Mitsui, Y., Nakamura, K., Borisov, S., Tischenko, G., Polyakov, K., & Pavlovsky, S. (1983) *Trends Biochem. Sci. 8*, 364.
Hollecker, M., & Creighton, T. E. (1982) *Biochim. Biophys. Acta 701*, 395.
Hughson, F. M., & Baldwin, R. L. (1989) *Biochemistry 28*, 4415.
Iida, S., & Ooi, T. (1969) *Biochemistry 8*, 3897.
Imoto, T. (1983) *Biophys. J. 44*, 293.
Inagaki, F., Kawano, Y., Shimada, I., Takahasi, K., & Miyazawa, T. (1981) *J. Biochem. (Tokyo) 89*, 1185.
Johnson, M. L., & Frazier, S. G. (1985) *Methods Enzymol. 117*, 301.
Kellis, J. T., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914.
Kostrewa, D., Choe, H.-W., Heinemann, U., & Saenger, W. (1989) *Biochemistry 28*, 7592.
Matouschek, A., Kellis, J. T., Serrano, L., & Fersht, A. R. (1989) *Nature 340*, 122.
Matthew, J. B., & Richards, F. M. (1982) *Biochemistry 21*, 4989.
Nozaki, Y., & Tanford, C. (1963) *J. Biol. Chem. 238*, 4074.
Nozaki, Y., & Tanford, C. (1967a) *J. Biol. Chem. 242*, 4731.
Nozaki, Y., & Tanford, C. (1967b) *Methods Enzymol. 11*, 715.
Nozaki, Y., & Tanford, C. (1970) *J. Biol. Chem. 245*, 16.
Oobatake, M., Takahashi, K., & Ooi, T. (1979) *J. Biochem. (Tokyo) 86*, 55.
Pace, C. N. (1975) *CRC Crit. Rev. Biochem. 3*, 1.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266.
Pace, C. N., & Vanderburg, K. E. (1979) *Biochemistry 18*, 288.
Pace, C. N., & Marshall, H. F., Jr. (1980) *Arch. Biochem. Biophys. 199*, 270.
Pace, C. N., & Grimsley, G. R. (1988) *Biochemistry 27*, 3242.
Pace, C. N., & Laurents, D. V. (1989) *Biochemistry 28*, 2520.
Pace, C. N., Grimsley, G. R., & Barnett, B. J. (1987) *Anal. Biochem. 167*, 418.
Pace, C. N., Shirley, B. A., & Thomson, J. A. (1989) in *Protein Structure: a practical approach* (Creighton, T. E., Ed.) pp 311–330, IRL Press, Oxford.
Privalov, P. L., Tiktopulo, E. I., Venyaminov, S. Yu., Griko, Yu. V., Makhatadze, G. I., & Khechinashvili, N. N. (1989) *J. Mol. Biol. 205*, 737.
Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151.
Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., & Zehfus, M. H. (1985) *Science 229*, 834.
Salahuddin, A., & Tanford, C. (1970) *Biochemistry 9*, 1342.
Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063.
Schellman, J. A. (1978) *Biopolymers 17*, 1305.
Schellman, J. A. (1987) *Biopolymers 26*, 549.

Schwarz, F. P., & Kirchhoff, W. H. (1988) *Thermochim. Acta 128*, 267.
Shirley, B. A., Stanssens, P., Steyaert, J., & Pace, C. N. (1989) *J. Biol. Chem. 264*, 11621.
Shortle, D. (1989) *J. Biol. Chem. 264*, 5315.
Shortle, D., & Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81.
Shortle, D., Meeker, A. K., & Gerring, S. L. (1989) *Arch. Biochem. Biophys. 272*, 103.
Tanford, C. (1961) *Physical Chemistry of Macromolecules*, Wiley, New York.
Tanford, C. (1964) *J. Am. Chem. Soc. 86*, 2050.
Tanford, C. (1968) *Adv. Protein Chem. 23*, 121.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1.
Tanford, C., & Hauenstein, J. D. (1956) *J. Am. Chem. Soc. 78*, 5287.
Thomson, J. A., & Bigelow, C. C. (1986) *Biochem. Cell Biol. 64*, 993.
Thomson, J. A., Shirley, B. A., Grimsley, G. R., & Pace, C. N. (1989) *J. Biol. Chem. 264*, 11614.
Tsong, T. Y. (1975) *Biochemistry 14*, 1542.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694.
Wearne, S. J., & Creighton, T. E. (1988) *Proteins: Struct., Funct., Genet. 4*, 251.
White, F. H., & Anfinson, C. B. (1959) *Ann. N.Y. Acad. Sci. 81*, 515.

Isolation and Functional Reconstitution of Soluble Melibiose Permease from *Escherichia coli*

Paul D. Roepe and H. Ronald Kaback*
Howard Hughes Medical Institute and Department of Physiology, Molecular Biology Institute, University of California, Los Angeles, California 90024-1574

Received August 23, 1989; Revised Manuscript Received December 26, 1989

ABSTRACT: By use of techniques described recently for *lac* permease [Roepe, P. D., & Kaback, H. R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6087], the melibiose permease from *Escherichia coli*, another polytopic integral plasma membrane protein, has been purified in a metastable soluble form after overexpression of the *melB* gene via the T7 RNA polymerase system. As demonstrated with *lac* permease, soluble melibiose permease is dissociated from the membrane with 5.0 M urea and appears to remain soluble in phosphate buffer at neutral pH after removal of urea by dialysis, although the protein aggregates in a time- and concentration-dependent fashion. Moreover, soluble melibiose permease behaves as a monomer during purification by size exclusion chromatography in the presence of urea. Circular dichroism of purified soluble melibiose permease reveals that the protein is highly helical in potassium phosphate buffer and that secondary structure is disrupted in 5.0 M urea. Finally, purified melibiose permease can be reconstituted into proteoliposomes, and the preparations catalyze membrane potential driven H^+/melibiose or Na^+/methyl 1-thio-β,D-galactopyranoside symport. The results provide further support for the notion that hydrophobic transmembrane proteins may be able to assume a nondenatured conformation in aqueous solution and extend the implication that the approach described may represent a general method for rapid isolation and reconstitution of this class of membrane proteins.

The facultative anaerobe *Escherichia coli* utilizes four types of membrane transport mechanisms [cf. Neidhart et al. (1987)]: (i) vectorial phosphorylation via the phosphoenolpyruvate–phosphotransferase system which catalyzes covalent modification of certain sugars during translocation; (ii) periplasmic binding protein mediated transport in which high-energy phosphate is used directly as a source of energy; (iii) primary active transport in which H^+ is extruded by the respiratory chain or the H^+-ATPase leading to the generation of an H^+ electrochemical gradient ($\Delta\bar{\mu}_{H^+}$[1]; interior negative and/or alkaline); and (iv) secondary active transport in which solute is accumulated against a concentration gradient in response to $\Delta\bar{\mu}_{H^+}$.

Accumulation of sugars such as lactose or melibiose by secondary active transport involves cotransport with H^+ or Na^+ in response to $\Delta\bar{\mu}_{H^+}$ (i.e., symport). The process is mediated by substrate-specific polytopic inner membrane proteins called porters, carriers, or permeases which utilize the free energy released from the downhill flow of H^+ or Na^+ to drive uphill accumulation of solute. One well-studied example of a secondary active transport system is the *lac* permease, which has been solubilized from the cytoplasmic membrane, purified to homogeneity, and shown to catalyze the coupled translocation of a single H^+ with a single β-galactoside molecule in monomeric form [cf. Kaback (1988, 1989) for reviews]. The *lacY* gene, which encodes the permease, has been sequenced (Büchel et al., 1980), and on the basis of circular dichroic studies with purified *lac* permease and the hydrophathy profile of the deduced amino acid sequence, it was suggested that the protein is composed of 12 hydrophobic transmembrane α-helices connected by loops containing most of the charged residues (Foster et al., 1983). Evidence supporting some of the general features of the model has been obtained [cf. Kaback (1988, 1989)], and more recently, a series of *lacY–phoA* fusions [cf. Manoil and Beckwith (1986)] has provided strong support for

* To whom correspondence should be addressed.

[1] Abbreviations: $\Delta\bar{\mu}_{H^+}$, the H^+ electrochemical gradient; CD, circular dichroism; TMG, methyl 1-thio-β,D-galactopyranoside; KP_i, potassium phosphate; DTT, dithiothreitol; NaP_i, sodium phosphate; $NaDodSO_4$, sodium dodecyl sulfate; Ch, choline.

0006-2960/90/0429-2572$02.50/0