CHAPTER 13

Measuring the conformational stability of a protein

C. NICK PACE, BRET A. SHIRLEY and JAMES A. THOMSON

1. INTRODUCTION

The stability of proteins, especially enzymes, has long been a practical concern (1), because this is usually the factor that most limits their usefulness. There are two very different aspects of protein stability. One is the chemical stability of the covalent structure, which involves covalent changes and is usually irreversible. The other is the conformational stability of the folded state, in the absence of covalent changes (2–10). The latter is the subject of this chapter, which will describe the simplest methods available for measuring how much more stable is the folded conformation of a protein than its unfolded conformations.

Measuring the conformational stability requires determining the equilibrium constant and the free energy change, ΔG, for the reaction:

$$\text{Folded (F)} \leftrightarrows \text{Unfolded (U)} \qquad (1)$$

We will refer to the value of ΔG at 25°C in the absence of a denaturant, $\Delta G(H_2O)$, as the conformational stability of a protein. Measurements of the conformational stabilities of proteins are needed for a variety of purposes and have become especially important now that we can construct proteins to order. Studies of proteins differing slightly in structure should allow us to gain a good understanding of the forces that determine the conformations of proteins and to optimize their stabilities.

We will describe how to determine and analyse thermal, urea and guanidinium chloride (GdmCl) denaturation curves. These are relatively simple experiments that can be done in almost any laboratory. We will show how to use this information to estimate $\Delta G(H_2O)$, and to measure differences in stability among proteins. These experiments sometimes reveal additional features of a protein such as the existence of domains or the presence of stable folding intermediates.

2. SELECTING A TECHNIQUE TO FOLLOW UNFOLDING

You must first decide which technique to use to follow unfolding. The techniques used most often are UV difference spectroscopy, fluorescence and circular dichroism (CD), which are described in Chapter 11. Other techniques used are biological activity measurements, optical rotatory dispersion (ORD), nuclear magnetic resonance (NMR), and viscosity and other hydrodynamic methods

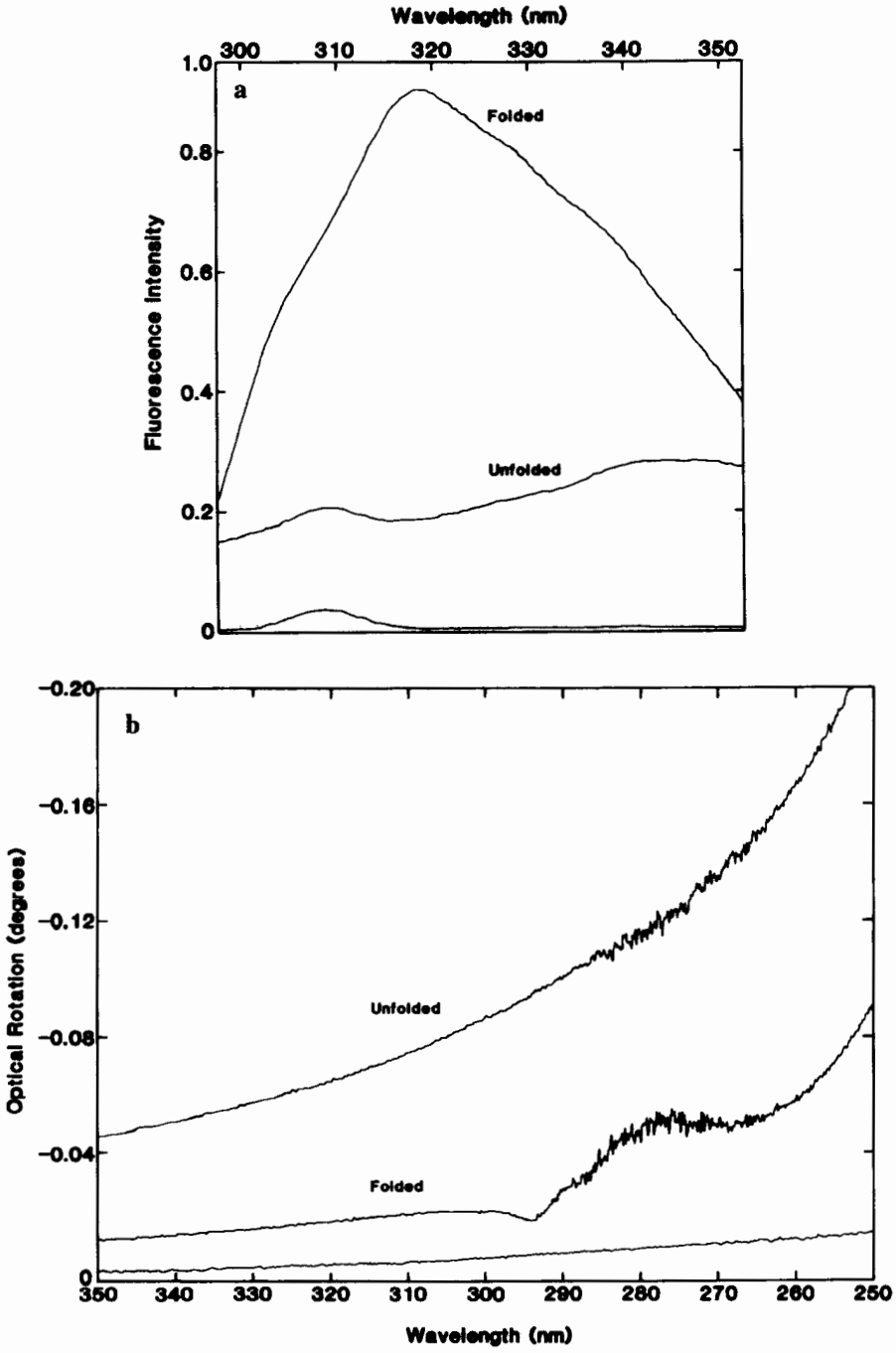

Figure 1. Spectra of folded and unfolded (8 M urea) RNase T1. (**a**) Intrinsic fluorescence emission spectra upon excitation at 278 nm. (**b**) ORD spectra. The RNase T1 concentration was 0.01 mg/ml for (**a**) and 1.2 mg/ml for (**b**). The unlabelled curves are of the solvent.

(Chapters 9 and 10). Only spectroscopic techniques will be discussed here; biological activity measurements present special problems (11). To decide on a technique, the spectra of the folded and unfolded conformations of the protein of interest should be determined. As examples, the fluorescence and ORD spectra of folded and unfolded ribonuclease T1 are shown in *Figure 1a* and *b*. Three features of the spectra are important in deciding on a technique to follow unfolding.

First, the magnitude of the response may be of crucial importance. The two ORD spectra in *Figure 1b* required 7 mg of RNase T1, but the two fluorescence spectra in *Figure 1a* required only 0.04 mg. This is generally true; when fluorescence can be used, it will require less protein than other techniques. UV difference spectroscopy, CD and ORD generally require greater amounts of material, although the amounts depend upon the wavelength chosen to follow unfolding. At present, NMR generally requires the greatest amount of protein. In return, however, NMR will generally yield considerably more detailed structural information than the other techniques, especially if the unfolding transition is complex. Thus, the technique you choose may be limited by the amount of protein that you have available for the experiments.

Second, it is necessary to pick a technique and a wavelength for which the spectra of the folded and unfolded conformations differ significantly. With fluorescence of RNase T1, we chose 320 nm where the intrinsic fluorescence of folded RNase T1 is 5- to 6-fold greater than that of the unfolded protein. With ORD, we chose 295 nm where the rotation of unfolded RNase T1 is 7-fold greater than that of the folded protein, and the signal-to-noise ratio is higher than at lower wavelengths. In general, pick the wavelength where the properties of the folded and unfolded conformations differ most.

The third factor that must be considered is the signal-to-noise ratio; the greater its value, the more accurate the measurements you will be able to make.

The spectral changes observed upon unfolding often depend upon different features of protein structure. For example, difference spectroscopy responds to changes in the environment of the tryptophan and tyrosine residues, and hence to changes in tertiary structure, while CD measurements below 250 nm depend mainly on changes in the secondary structure of a protein. This may also be a consideration in determining the technique you should use to follow unfolding. These topics are discussed in more detail in Chapter 11.

Finally, two practical matters should be mentioned. Fluorescence measurements should not be used for following thermal unfolding, because the pre- and post-transition baselines are steep and very temperature sensitive. Difference spectroscopy is less convenient for following urea and GdmCl unfolding because twice as many solutions must be prepared, and the pre- and post-transition baselines are steeper than with most of the other techniques.

3. DETERMINING UNFOLDING CURVES

Figure 2a and *b* show typical urea and thermal unfolding curves. In this case, optical rotation was used to follow unfolding, but we will refer to whatever

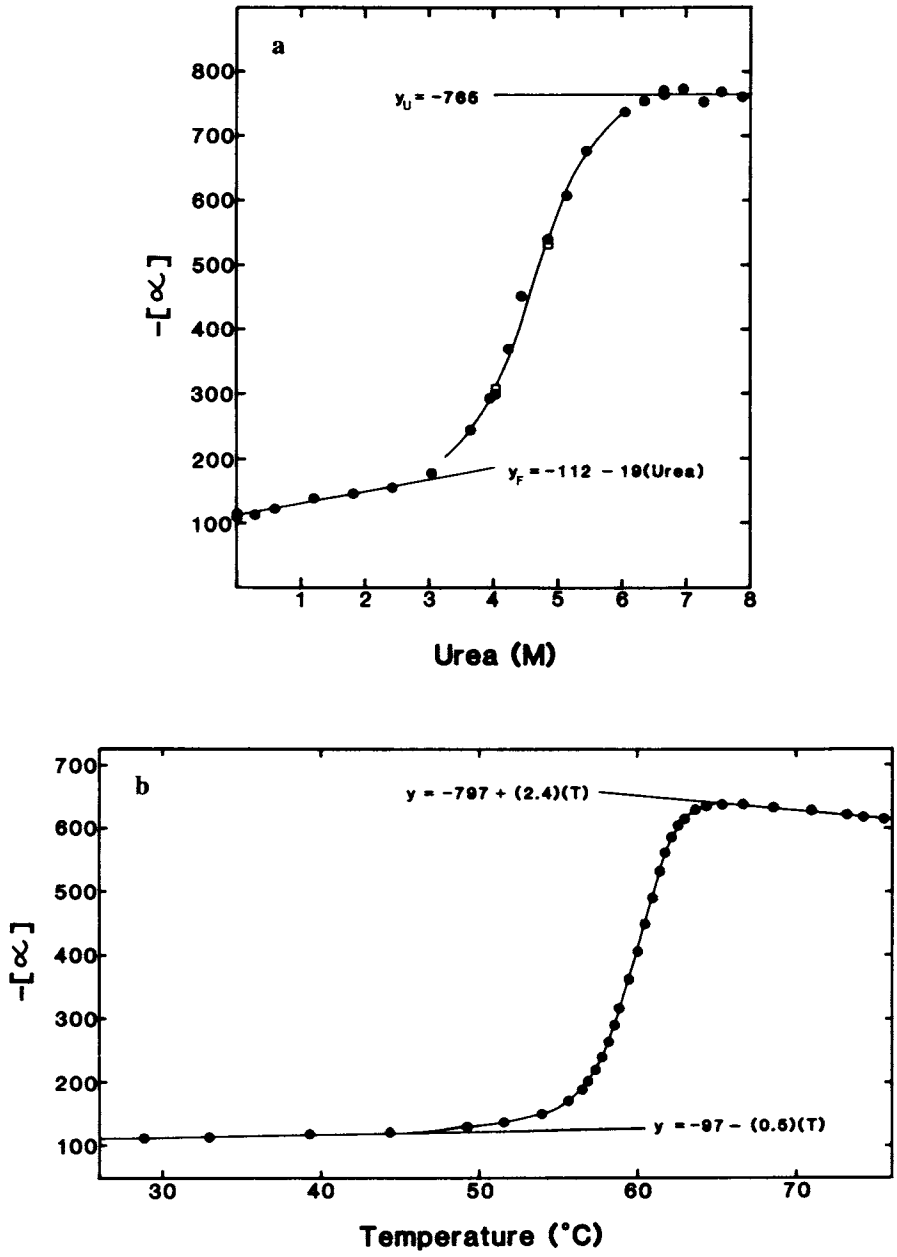

Figure 2. Urea- and heat-induced unfolding transitions of RNase T1, measured by ORD ([α] is the specific optical rotation at 295 nm). The equations for y_F and y_U give the value of [α] for the folded and unfolded states, respectively, where (urea) is the molar concentration of urea and (T) is the temperature in °C. (**a**) Unfolding as a function of urea concentration at 25°C in 30 mM Mops buffer at pH 7. The open squares indicate samples that were unfolded by exposure to 7 M urea, then diluted to the final urea concentration. (**b**) Heat-induced unfolding in 0.25 M NaCl and 0.1 M formate buffer, pH 5.0.

physical parameter is used as y for the discussion that follows. The curves can be conveniently divided into three regions.

(i) The pre-transition region, which shows how y for the folded protein, y_F, depends upon the denaturant.
(ii) The transition region, which shows how y varies as unfolding occurs.
(iii) The post-transition region, which shows how y for the unfolded protein, y_U, varies with denaturant.

All of these regions are important for analysing unfolding curves. As a minimum, we recommend determining four points in the pre- and post-transition regions, and five points in the transition region. Of course, the more points determined the better defined the curve. In general, points at the corners, between regions, are not used in the analysis of the results.

3.1 **Equilibrium and reversibility**

Since we are dealing with thermodynamic measurements, it is essential that the unfolding reaction has reached equilibrium before measurements are made, and that the unfolding reaction is reversible. The time required to reach equilibrium can vary from seconds to days, depending upon the protein and the conditions. For example, unfolding of RNase T1 reaches equilibrium in minutes at 30°C, but requires hours at 20°C. For unfolding, the time to reach equilibrium generally decreases as denaturant is increased. Thus, the time to reach equilibrium will be shorter in the post-transition region than it is near the midpoint of the transition. To ensure that equilibrium is reached, y is measured as a function of time to establish the time required to reach equilibrium.

To test the reversibility of unfolding, allow a solution to reach equilibrium in the post-transition region and then, by cooling or dilution, return the solution to the pre-transition region and measure y. The value of y measured after complete unfolding should be identical to that determined directly. In general, urea and GdmCl unfolding are more likely to be completely reversible than thermal unfolding. In fact, we have left RNase T1 in 6 M GdmCl solutions for 3 months and found that the protein will refold completely on dilution. The thermal unfolding of RNase T1 is not completely reversible, and the degree of irreversibility increases the longer the protein is exposed to unfolding conditions. Similar observations have been made with many other proteins by others. For this reason, we generally recommend that thermal denaturation curves be determined as quickly as possible.

Proteins that contain free sulphydryl groups present special problems. If the protein contains only free —SH groups and no disulphide bonds, then a reducing agent such as 10 mM dithiothreitol (DTT) can be added to ensure that no disulphide bonds form during the experiments. For proteins containing both free —SH groups and disulphide bonds, disulphide interchange can occur and this may lead to irreversibility. Disulphide interchange can be minimized by working at low pH (Chapter 7).

3.2 **Urea and GdmCl unfolding**

Each point shown in *Figure 2a* was determined on a separate solution. These solutions were prepared volumetrically, using the best available pipettes, by mixing a fixed volume of protein stock solution with the appropriate volumes of a buffer solution and a urea or GdmCl stock solution (see *Table 3*). The protein and buffer solutions are prepared by standard procedures. The urea or GdmCl stock solution must be prepared with considerable care; some suggestions are given below. After the solutions for measurement have been prepared, they are incubated until equilibrium is reached at the temperature chosen for determining the unfolding curve. After the measurements have been completed, it is good

Table 1. Information for preparing urea and GdmCl stock solutions

Property	*Urea*	*GdmCl*
Molecular weight	60.056	95.533
Solubility (25.0°C)	10.49 M	8.54 M
d/d_0[a]	$1 + 0.2658W + 0.0330W^2$	$1 + 0.2710W + 0.0330W^2$
Molarity[b]	$117.66(\Delta N) + 29.753(\Delta N)^2 + 185.56(\Delta N)^3$	$57.147(\Delta N) + 38.68(\Delta N)^2 - 91.60(\Delta N)^3$
Grams of denaturant per gram of water to prepare		
6 M	0.495	1.009
8 M	0.755	1.816
10 M	1.103	—

[a] W is the weight fraction denaturant in the solution, d is the density of the solution and d_0 is density of water (see ref. 17).

[b] ΔN is the difference in refractive index between the denaturant solution and water (or buffer) at the sodium D line. The equation for urea solutions is based on data from (18), and the equation for GdmCl solutions is from (13).

Table 2. Example of preparation of a urea stock solution

This describes the preparation of ~100 ml of ~10 M urea stock solution containing 30 mM Mops buffer, pH 7.0. We use a top loading balance with an accuracy of about ±0.02 g.

1. Add ≈60 g of urea to a tared beaker and weigh (59.91 g). Now add 0.694 g of Mops buffer (sodium salt), 1.8 ml of 1 M HCl and ≈52 ml of distilled water and weigh the solution again (114.65 g).
2. Allow the urea to dissolve and check the pH. If necessary, add a weighed amount of 1 M HCl to adjust to pH 7.0.
3. Prepare a 30 mM Mops buffer, pH 7.0.
4. Determine the refractive index of the urea stock solution (1.4173) and of the buffer (1.3343). Therefore, $\Delta N = 1.4173 - 1.3343 = 0.0830$.
5. Calculate the urea molarity from ΔN using the equation given in *Table 1*: M = 10.08.
6. Calculate the urea molarity based on the recorded weights. The density is calculated with the equation given in *Table 1*: weight fraction area (W) = 59.91/114.65 = 0.5226; therefore d/d_0 = 1.148. Therefore volume = 114.65/1.148 = 99.88 ml. Therefore urea molarity = 59.91/60.056/ .09988 = 9.99 M.
7. Since the molarities calculated in steps 5 and 6 differed by less than 1%, this solution was used to determine a urea unfolding curve.

practice to measure the pH of the solutions in the transition region, at least. [If the amount of protein is limited, instruments have been developed for following the unfolding of a protein simultaneously with several spectral techniques using a single protein solution (12).]

Both urea and GdmCl can be purchased commercially in highly purified forms (e.g. from Schwarz Mann Biotech). However, some lots of GdmCl are found to contain fluorescent impurities and some lots of urea contain significant amounts of metallic impurities. Methods are available for checking the purity of GdmCl and

Table 3. Determining a urea or GdmCl unfolding curve

1. Prepare three solutions: a denaturant stock solution as described in *Table 2*, a protein stock solution and a buffer solution.
2. Prepare the solutions on which measurements will be made volumetrically (e.g. with Pipetman pipettes) in clean, dry test tubes. Typical solutions used in determining a urea unfolding curve with fluorescence measurements are shown below. (Only two solutions are shown for the pre- and post-transition regions. In the actual experiment, a total of 32 solutions were prepared and their fluorescence measured. These results were used to determine the parameters given for wild-type RNase T1 in *Table 8*.)
3. Allow these solutions to equilibrate at the temperature chosen for the experiment until they reach equilibrium. (This is best determined in a separate experiment as described in the text.)
4. Carefully measure the experimental parameter being used to follow unfolding.
5. Plot these results to determine if any additional points are needed. If so, prepare the appropriate solutions and make the measurements just as for the original solutions.
6. Measure the pH of the solutions in the transition region.
7. Analyse the results as described in *Table 5*.

Urea[a] *(ml)*	*Buffer*[b] *(ml)*	*Protein*[c] *(ml)*	*Urea*[d] *(M)*	*I*[e]
0.45	2.55	0.2	1.41	135
0.90	2.10	0.2	2.83	136
1.20	1.80	0.2	3.77	115
1.25	1.75	0.2	3.93	113
1.30	1.70	0.2	4.09	105
1.35	1.65	0.2	4.25	99
1.40	1.60	0.2	4.40	87
1.45	1.55	0.2	4.56	81
1.50	1.50	0.2	4.72	69
1.55	1.45	0.2	4.87	62
1.60	1.40	0.2	5.03	53
1.65	1.35	0.2	5.19	47
1.70	1.30	0.2	5.34	40
2.05	0.95	0.2	6.45	21
2.50	0.50	0.2	7.87	22

[a] 10.07 M urea stock solution (30 mM Mops, pH 7.0).
[b] 30 mM Mops buffer, pH 7.0.
[c] 0.10 mg/ml RNase T1 stock solution (30 mM Mops, pH 7.0).
[d] Urea molarity = 10.07 ((ml urea)/(ml urea + ml protein + ml buffer)) = 10.07 (ml urea)/(3.2) = 3.1469 (ml urea).
[e] Fluorescence intensity (278 nm excitation and 320 nm emission) measured with a Perkin-Elmer MPF 44B spectrofluorometer.

for recrystallization when it is necessary (13). A procedure for purifying urea has also been described (14). GdmCl solutions are stable for months, but urea solutions slowly decompose to form cyanate and ammonium ions (15). The cyanate ions can react with amino groups on proteins (16). Consequently, a fresh urea stock solution should be prepared for each unfolding curve and used within 1 day.

Table 1 summarizes useful information for preparing urea and GdmCl stock solutions. We prepare urea stock solutions by weight, and then check the concentration by refractive index measurements using the equation given in *Table 1*. If the concentrations agree within 1%, we use the solution for determining an unfolding curve. The preparation of a typical urea stock solution is outlined in *Table 2*. Since GdmCl is quite hygroscopic, it is more difficult to prepare stock solutions by weight. Consequently, the molarity of GdmCl stock solutions is generally based on refractive index measurements and the equation given in *Table 1*. See *Table 3* for further information.

3.3 **Thermal unfolding**

The time spent initially to ensure that the temperature of your cell can be measured accurately and maintained will be repaid many times over. Many instruments provide devices that allow water from a constant temperature bath to be circulated to maintain the temperature of the cell. If not, a variety of cells can be purchased that allow water from a constant temperature bath to be circulated around the solution on which measurements are to be made. It is *essential* that great care be taken in securing all of the tubing connections. Water at 90°C spewing from a loose tube can be a major disaster. It is generally a good idea to insulate the tubing leading from the water bath to the instrument. How accurately the temperature must be maintained depends mainly on the value of the enthalpy change, ΔH, for the unfolding of your protein and on whether you wish to determine the heat capacity change, ΔC_p, for unfolding (see Section 4.2). The greater the magnitude of ΔH, the more the equilibrium constant will fluctuate with temperature. In general, maintaining the temperature within ± 0.05°C is adequate.

One way to monitor the temperature is to insert a thermistor directly into the sample cell, and then seal the cell to minimize evaporation. We use a probe and a telethermometer manufactured by Yellow Springs Instrument Co. for this purpose. Since the scale on our instrument is small, we attach a voltmeter (Micronta Digital Multimeter) and use this to read the output of the telethermometer. The system is calibrated using a National Bureau of Standards certified thermometer. With this approach, it is essential that the temperature is homogeneous throughout the cell, because the thermistor is generally in the top of the cell and measurements are made on solution near the centre of the cell. If possible, the solution should be stirred. An alternate procedure is to calibrate the cell temperature against the bath temperature with your thermometer or thermistor in a separate experiment. Then it is necessary only to record the bath temperature

Table 4. Determining a thermal unfolding curve

1. Select a buffer with a small ΔH of ionization so that the variation of the pK of the buffer and, hence, the pH of the solution with temperature will be minimized. Near pH 7, Mops is excellent. In general, ΔH is small for carboxyl groups, and large for amino groups.
2. Make certain that the instrument to be used to follow unfolding is thoroughly warmed-up, and that the system for maintaining and measuring the temperature of the experimental cell is operating properly.
3. Prepare the protein solution, and measure the pH and protein concentration. If the solution is not completely clear, filter it with, for example, a 0.65 μm Millipore filter.
4. Carefully measure the experimental parameter, y, to be used to follow unfolding for a blank solution at the lowest temperature that will be used in the unfolding curve.
5. Replace the blank solution with your protein solution and measure y at sufficient temperatures in the pre-transition region to determine y_F.
6. Measure y at temperatures in the transition region. These measurements should be made as quickly as possible, but it is essential that the system reaches thermal and chemical equilibrium at each new temperature before the measurement is made. It may be best to determine the time required to reach equilibrium in a separate experiment designed for this purpose. If only T_m and ΔH_m are needed, it is only necessary to make measurements at about six temperatures: three above and three below T_m. If you are attempting to determine ΔC_p, measurements should be made with great care at more temperatures in the transition region.
7. Measure y at sufficient temperatures in the post-transition region to determine y_U.
8. At the highest temperature used, replace the protein solution with a blank solution and re-measure y. If y for the blank has not changed significantly, your experiment is probably successful. If your instrument is not very stable, it may be necessary to monitor the blank more carefully during the course of the experiment. These blank measurements can be very important, especially with single beam instruments.
9. Re-measure the pH of the protein solution after the experiment.
10. Analyse the results as described in *Table 7*. *Figure 2b* shows a typical thermal unfolding curve and *Table 7* shows typical data derived from a thermal unfolding curve.

while measurements are in progress and use the calibration curve to determine the cell temperature.

Only a single solution is needed to determine a thermal unfolding curve. Consequently, less protein is required than for urea and GdmCl unfolding experiments. For many techniques, it is a good idea to filter (e.g. with a 0.65 μm millipore filter) the protein solution before the experiment. This will remove any dust and often improves the signal-to-noise ratio. Since it may require several hours to determine a thermal unfolding curve, it is important that a very stable measuring instrument be used. Also, it is essential that blanks be run before and after the experiment to ensure that the instrument has not drifted significantly during the course of the experiment. After the solution has been cooled and the reversibility of unfolding checked, it is good practice to re-measure the pH of the solution. See *Table 4* for further information.

4. ANALYSING UNFOLDING CURVES

The unfolding of many small globular proteins has been found to closely approach a two-state folding mechanism, such as that shown in Equation 1. (See *Table 5* for

Table 5. Investigating the unfolding mechanism

1. If the unfolding curve shows more than one step, unfolding is more complex than a two-state mechanism and the analysis of the data is more complicated.[a] This behaviour is frequently observed with multi-domain proteins where the domains unfold independently.[b]
2. If a single-step unfolding curve is observed, it does not prove that unfolding closely approaches a two-state mechanism.[c]
3. Further insight can be gained by using different techniques to follow unfolding. Non-coincidence of plots of f_U as a function of denaturant determined using different techniques proves that significant concentrations of stable intermediates are present at equilibrium and hence that a two-state mechanism cannot be assumed in analysing the data.[d] Unfortunately, coincidence of the plots is consistent with, but does not prove, that unfolding follows a two-state mechanism.[e]
4. The best evidence that unfolding can be adequately approximated by a two-state mechanism is to show that ΔH determined by the van't Hoff relationship (Equation 6) is identical to ΔH determined calorimetrically. When $\Delta H(vH) < \Delta H(cal)$, it is clear evidence that significant concentrations of intermediates are present at equilibrium.[f]
5. The mechanism of folding of globular proteins is an area of great interest to biochemists and has been reviewed recently by Kim and Baldwin.[g] This review and Chapter 9 should be consulted for more information on the techniques used to investigate the mechanism of folding of proteins in detail. In particular, these articles consider the detection of kinetic intermediates, whereas, for the experimental techniques discussed here, our only interest is in the presence of significant concentrations of intermediates present at equilibrium.

[a] See ref. 19.
[b] See refs 12, 20 and 21.
[c] See refs 22, 23.
[d] See refs 11, 12.
[e] See ref. 24.
[f] See refs 6, 24–26.
[g] See ref. 27.

a summary of methods that can be used to gain more information about the folding mechanism.) We will assume a two-state folding mechanism for the discussion here. Consequently, for any of the points shown in *Figure 2a* or *b*, only the folded and unfolded conformations are present at significant concentrations, and $f_F + f_U = 1$, where f_F and f_U represent the fraction protein present in the folded and unfolded conformations, respectively. Thus, the observed value of y at any point will be $y = y_F f_F + y_U f_U$, where y_F and y_U represent the values of y characteristic of the folded and unfolded states, respectively, under the conditions where y is being measured. Combining these equations yields

$$f_U = (y_F - y)/(y_F - y_U) \tag{2}$$

The equilibrium constant, K, and the free energy change, ΔG, can be calculated using

$$K = f_U/(1 - f_U) = f_U/f_F = (y_F - y)/(y - y_U) \tag{3}$$

and

$$\Delta G = -RT \ln K = -RT \ln [(y_F - y)/(y - y_U)] \tag{4}$$

where R is the gas constant (1.987 calories/deg/mol) and T is the absolute temperature. Values of y_F and y_U in the transition region are obtained by extrapolating from the pre- and post-transition regions, as illustrated in *Figure 2a* and *b*. A least-squares analysis was used to determine the equations given for y_F and y_U. In some cases, important information can be obtained from these baselines. For example, when UV difference spectroscopy is used to follow unfolding, the dependence of y_F on urea concentration can be used to estimate how many of the tyrosine and tryptophan residues are exposed to solvent in the folded conformation (28).

The calculation of f_U, K and ΔG from results such as those shown in *Figure 2a* and *b* is illustrated in *Tables 6* and *7*. Values of K can be measured most accurately near the midpoints of solvent or thermal denaturation curves, and the error becomes substantial for values outside the range 0.1–10. Consequently, we generally only use ΔG values with the range ± 1.5 kcal/mol.

4.1 **Urea and GdmCl unfolding**

The data shown in *Figure 2a* have been analysed as described in *Table 6* and the results are shown in *Figures 3* and *4*. In *Figure 3*, f_U is shown as a function of urea concentration. When comparing results from studies of related proteins that have similar pre-and post-transition behavior it is often useful to show just f_U as a function of denaturant in the transition region rather than a complete unfolding curve.

Table 6. Analysis of a urea unfolding curve

1. The data points from the transition region of *Figure 2a* are listed below. Values of f_U, K and ΔG were calculated using Equations 2, 3 and 4, respectively, and the y_F and y_U values given by the equations for the pre- and post-transition regions in *Figure 2a*.

[Urea][a] (M)	y^b	f_U	K	*ΔG (cal/mol)*
3.64	244	0.109	0.123	1242
3.94	293	0.186	0.229	873
4.03	302	0.199	0.248	825
4.24	370	0.312	0.453	469
4.54	452	0.449	0.815	121
4.85	536	0.594	1.462	−225
5.15	607	0.716	2.525	−549
5.45	677	0.840	5.261	−983

2. A least-squares analysis of this data yields the following for Equation 5: $\Delta G = 5680 - (1218)$ [urea]. This equation describes the solid line shown in *Figure 4*. Thus, $\Delta G(H_2O) = 5.68$ kcal/mol, $m = 1218$ cal/mol/M, and $[\text{urea}]_{1/2} = 4.66$ M. All of these parameters should be reported to characterize the results from a urea denaturation curve.

[a] Solutions with a total volume of 0.85 ml were prepared volumetrically, e.g. 3.64 M = 0.3 ml 10.30 M urea + 0.35 ml buffer + 0.2 ml 4.5 mg/ml RNase T1.

[b] The value of $-y$ is the specific rotation measured at 295 nm.

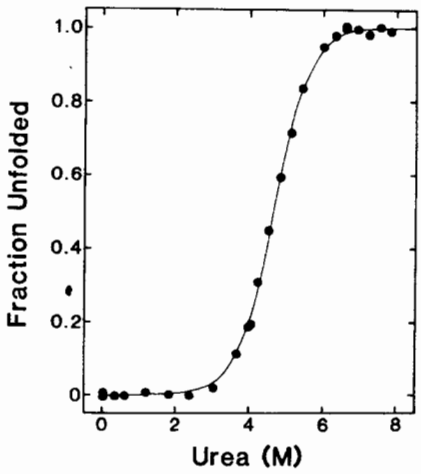

Figure 3. Fraction of RNase T1 unfolded, f_U, as a function of urea concentration. The values of f_U were calculated from the data of *Figure 2a* using Equation 2 of the text.

It can be seen in *Figure 4* that ΔG varies linearly with denaturant concentration in the limited region where ΔG can be measured. Similar results have been obtained with many other proteins (29, 30). The simplest method of estimating the conformational stability in the absence of urea, $\Delta G(H_2O)$, is to assume that this linear dependence continues to zero concentration and to use a least-squares analysis to fit the data to the following equation:

$$\Delta G = \Delta G(H_2O) - m\,[\text{denaturant}] \qquad (5)$$

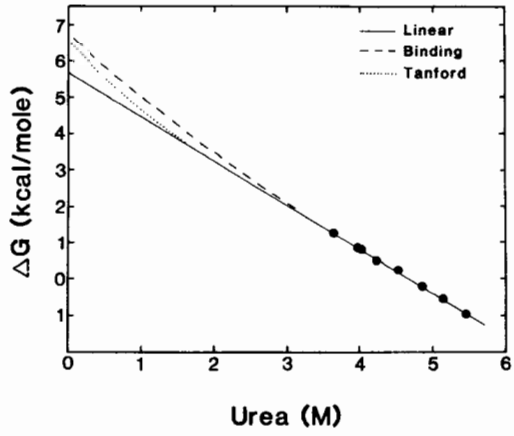

Figure 4. ΔG for RNase T1 unfolding as a function of urea concentration. The experimental values of ΔG (closed circles) were calculated from the data in *Figure 2a* using Equation 4 of the text. The solid line represents Equation 5 with $\Delta G(H_2O) = 5.7$ kcal/mol and $m = 1220$ cal/mol/M. See reference 19 for a discussion of the other two methods of extrapolating the experimental data.

where m is a measure of the dependence of ΔG on denaturant concentration. Note also that the denaturant concentration at the midpoint of the unfolding curve, $[\text{denaturant}]_{1/2} = \Delta G(H_2O)/m$. We strongly recommend that values of $\Delta G(H_2O)$, m and $[\text{denaturant}]_{1/2}$ be given in any study of the unfolding of a protein by urea or GdmCl.

As the dashed lines in *Figure 4* suggest, other extrapolation methods can be used to estimate $\Delta G(H_2O)$. At present, there is no good reason for using these more complicated procedures, and they will not be discussed here [see reference (19) for more information]. In general, the estimates resulting from the other methods do not differ substantially when urea is the denaturant, but they may with GdmCl. Schellman (31) has suggested that a linear dependence of ΔG on denaturant concentration may be reasonable based on theoretical grounds.

4.2 **Thermal unfolding**

We will now consider how to analyse a thermal unfolding curve, *Figure 2b*, in order to determine the conformational stability, $\Delta G(H_2O)$ [see Becktel and Schellman (32) for an excellent discussion of this subject]. This requires extrapolating ΔG measurements from the narrow temperature range where unfolding occurs to an ambient temperature, T, such as 25°C. *Table 7* illustrates the calculation of K and ΔG from results such as those in *Figure 2b*. The familiar method used to obtain the enthalpy change, ΔH, from these measurements is with the van't Hoff equation:

$$d(\ln K)/d(1/T) = -\Delta H/R \tag{6}$$

The van't Hoff plots ($\ln K$ versus $1/T$) of protein unfolding transitions are found to be non-linear. This indicates that ΔH varies with temperature, which is expected when the heat capacities of the products and reactants differ:

$$d(\Delta H)/d(T) = C_p(U) - C_p(F) = \Delta C_p \tag{7}$$

where $C_p(U)$ and $C_p(F)$ are the heat capacities of the unfolded and folded conformations, respectively, and ΔC_p is the change in heat capacity that accompanies protein unfolding. In cases such as this, both ΔC_p and ΔH are required to calculate ΔG as a function of temperature. ΔH is needed at only a single temperature, and the best temperature to use is T_m, the midpoint of the thermal unfolding curve, where $\Delta G(T_m) = 0 = \Delta H_m - T_m \Delta S_m$. Now, with these parameters, the equation used to calculate ΔG at any temperature T, $\Delta G(T)$, can be given:

$$\Delta G(T) = \Delta H_m(1 - T/T_m) - \Delta C_p[(T_m - T) + T \ln (T/T_m)] \tag{8}$$

Thus, we need T_m, ΔH_m and ΔC_p in order to calculate $\Delta G(T)$. The simplest method to determine T_m and ΔH_m is to plot ΔG as a function of temperature, and then use $T_m = T$ at $\Delta G = 0$, and $\Delta H_m = (T_m \text{ in K}) \times (\text{slope at } T_m)$. Calculations based on this approach and on the van't Hoff equation are illustrated in *Table 7*.

The determination of ΔC_p is more difficult. The safest approach is to use a

Table 7. Analysis of a thermal unfolding curve

1. The data are from a thermal unfolding curve determined under the same conditions as the urea unfolding curve analysed in *Table 3*. Values of f_U, K and ΔG were calculated using Equations 2, 3 and 4, respectively. A least-squares analysis of the pre- and post-transition region gave $y_F = 87.2 + 0.66\,(T)$ and $y_U = 646$.

T(°C)	y[a]	f_U	K	ΔG (cal/mol)
16.2	98.1			
21.0	100.9			
25.6	103.7			
30.2	107.4			
45.4	221.3	0.197	0.245	890
46.3	263.9	0.277	0.383	610
47.2	313.9	0.371	0.589	337
48.1	367.6	0.472	0.894	72
49.0	422.2	0.575	1.353	−193
49.9	474.1	0.673	2.061	−464
50.8	518.5	0.757	3.123	−733
51.7	555.5	0.828	4.805	−1013
61.2	645.4			
65.5	646.3			
69.8	646.3			
73.8	645.4			

2. The slope of a plot of ΔG versus $T = -300.3$ cal/mol/deg, and $T = T_m = 48.3$°C at $\Delta G = 0$. Since $\Delta G = 0$ at T_m, $\Delta H_m = (T_m)(\Delta S_m)$. Therefore, $\Delta H_m = (48.3 + 273.2)(300.3) = 96.5$ kcal/mol. The slope of a van't Hoff plot (Equation 6) $= -48\,631$ K. Therefore, $\Delta H_m = -(1.987)(-48\,631) = 96.6$ kcal/mol. The values of ΔH obtained by taking the slope between individual points in the transition regions are: 99.9, 97.5, 94.7, 94.8, 96.7, 96.1 and 100.1 kcal/mol. Clearly, these data cannot be used to determine ΔC_p with Equation 7.
3. To estimate ΔG at 25°C, we use Equation 8. If we assume $\Delta C_p = 0$, the second term in the equation $= 0$, and $\Delta G(25°C) = 7.0$ kcal/mol. If we calculate ΔC_p using the rough rule of thumb mentioned in the text, ΔC_p = (12 cal/mol/deg/residue) × (104 residues) = 1250 cal/mol/deg. This leads to $\Delta G(25°C) = 7.0 - 1.1 = 5.9$ kcal/mol with Equation 8. Note that this is in excellent agreement with the value of $\Delta G(H_2O) = 5.7$ kcal/mol from the analysis of a urea unfolding curve in *Table 6*.

[a]The value of $-y$ is the specific rotation measured at 295 nm.

calorimeter to determine ΔC_p. In favourable cases, ΔC_p can be measured to about ±10% with a differential scanning microcalorimeter (Julian Sturtevant, personal communication). A non-calorimetric approach that has been used successfully with T4 lysozyme (32), the chymotrypsinogen family (33) and other proteins (25), is to measure T_m and ΔH_m as a function of pH. T_m usually varies with pH and this causes ΔH_m to vary also. The slope of a plot of ΔH_m versus T_m will yield an estimate of ΔC_p, as shown by Equation 7. This assumes that ΔH and ΔC_p do not vary significantly with pH, which appears to be the case normally (25, 32, 33). In favourable cases, ΔC_p can be determined more directly using data such as those shown in *Figure 2b*. Equations 6 and 7 show that this requires taking the second derivative of the experimental data; this is only possible if the experimental data

are exceptionally good. We had success with this approach in a favourable case (34), but we have not been able to determine a reasonable and reproducible value of ΔC_p for RNase T1, which is an unfavourable case. Brandts used this approach to determine ΔC_p for chymotrypsinogen in his pioneering studies in this area (35). All three of the methods described here give comparable values of ΔC_p for chymotrypsinogen (36).

For several proteins, estimates of ΔC_p in reasonable agreement with experimental values can be calculated using model compound data and the amino acid composition of the protein (37). As a rule of thumb, multiplying the number of residues in a protein by 12 cal/mol/deg will give a rough estimate of ΔC_p. We use this unsatisfactory approach in *Table 4*, to obtain a value of ΔC_p for RNase T1 folding to illustrate the use of Equation 8.

5. DETERMINING DIFFERENCES IN STABILITY

It is frequently of interest to determine differences in conformational stability among proteins that vary slightly in structure. The structural change might be a

Table 8. Determining differences in stability

1. Results from an analysis of urea unfolding curves for wild-type RNase T1 and two mutants differing by one residue in amino acid sequence are presented below.

Protein	$\Delta G(H_2O)$[a]	$\Delta(\Delta G)$[b]	m[c]	$[urea]_{1/2}$[d]	$\Delta[urea]_{1/2}$[e]	$\Delta(\Delta G)$[f]
Wild-type	5.59		1220	4.58		
Gln 25 Lys	6.39	0.80	1210	5.30	0.72	0.87
Glu 58 Ala	4.72	−0.87	1200	3.92	−0.66	−0.80

2. Results from an analysis of the thermal unfolding curves for the same proteins under the same conditions.

Protein	ΔS_m[g]	ΔH_m[h]	T_m[i]	$\Delta(T_m)$[j]	$\Delta(\Delta G)$[k]
Wild-type	295	95	49.2		
Gln 25 Lys	330	107	51.7	2.5	0.74
Glu 58 Ala	313	101	47.9	−1.3	−0.38

[a] From Equation 5, in kcal/mol.
[b] Difference between the $\Delta G(H_2O)$ values in kcal/mol.
[c] From Equation 5, in cal/mol/M.
[d] Midpoint of the urea unfolding curve in M.
[e] Difference between the $[urea]_{1/2}$ values.
[f] From $\Delta[urea]_{1/2} \times 1210$ (the average of the three m values) in kcal/mol.
[g] Slope of ΔG versus T at T_m in cal/mol/deg.
[h] $\Delta H_m = [T_m(K)] \times (\Delta S_m)$ in kcal/mol.
[i] Midpoint of thermal unfolding curve in °C.
[j] Difference between the T_m values.
[k] $\Delta(\Delta G) = [\Delta(T_m)] \times \Delta S_m = [\Delta(T_m)] \times (\Delta H_m/T_m)$, where ΔS_m and ΔH_m are values for the wild-type protein (see ref. 32).

single change in amino acid sequence achieved through site-directed mutagenesis, or a change in the structure of a side-chain resulting from chemical modification. In *Table 8*, we present results from urea and thermal unfolding studies of wild-type RNase T1 and two mutants that differ in amino acid sequence by one residue. Three different methods of calculating the differences in stability, $\Delta(\Delta G)$, are illustrated.

The midpoints of urea, $[\text{urea}]_{1/2}$, and thermal, T_m, unfolding curves can be determined quite accurately and do not depend to a great extent on the unfolding mechanism (19). In contrast, measures of the steepness of urea, m, and thermal, $-\Delta S_m$, unfolding curves cannot be determined as accurately, and deviations from a two-state folding mechanism will generally lower these values. Consequently, with differences in stability determined by comparing the $\Delta G(H_2O)$ values, the errors can be quite large. However, when comparing completely different proteins or forms of a protein that differ markedly in stability, no other choice is available. This approach is illustrated by the first column of $\Delta(\Delta G)$ values in *Table 8*. In *Table 9*, we illustrate the dangers of trying to draw conclusions about the conformational stabilities of unrelated proteins based solely on the midpoints of their unfolding curves. Lysozyme and myoglobin have similar $\Delta G(H_2O)$ values, but a much higher concentration of GdmCl is needed to denature lysozyme because the m value is much smaller. Likewise, lysozyme and cytochrome c are unfolded at about the same temperature, even though lysozyme has a much larger value of $\Delta G(H_2O)$.

A second approach is to compare the proteins in the presence of urea by taking the difference between the $[\text{urea}]_{1/2}$ values and multiplying this by the average of the m values. This is illustrated by the $\Delta(\Delta G)$ values in the last column of *Table 8*. The rationale here was that the error in measuring the m values should generally be greater than any differences resulting from the effect of small changes in structure on the m value. Recent studies (38), however, have observed substantial differences in m values between proteins differing by one amino acid in sequence. This is claimed to reflect a substantial change in the interaction of the unfolded molecule with denaturant, rather than a deviation from a two-state mechanism. This question is being investigated further at present.

Becktel and Schellman recently considered how to analyse thermal denatura-

Table 9. Unfolding curve midpoints may not reflect conformational stabilities

Protein	$[GdmCl]_{1/2}$ *(M)*	$\Delta G(H_2O)$ *(kcal/mol)*
Myoglobin	1.7	12
Lysozyme	3.1	12
	T_m *(°C)*	$\Delta G(H_2O)$ *(kcal/mol)*
Cytochrome *c*	80	8
Lysozyme	80	12

tion curves to determine differences in stability (32). We have used their suggestions in calculating the $\Delta(\Delta G)$ values from thermal unfolding in *Table 8*. For one mutant, the three estimates of $\Delta(\Delta G)$ are in good agreement, which is reassuring. For the other mutant, $\Delta(\Delta G)$ from thermal unfolding is considerably less than the

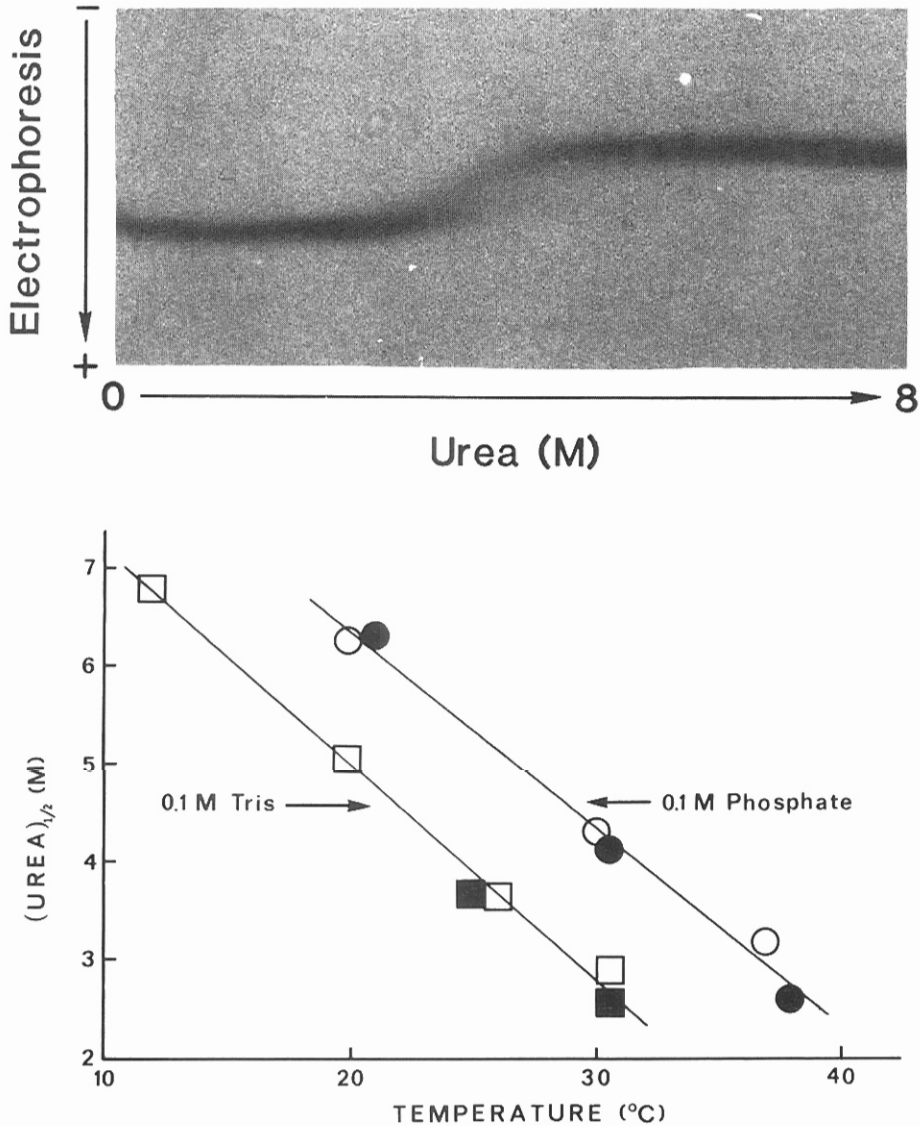

Figure 5. Unfolding of RNase T1 measured by urea gradient electrophoresis. (**a**) An example of one urea gradient gel (Chapter 10). (**b**) Comparison of the unfolding measured by urea gradient gels (open symbols) and by conventional urea unfolding curves (closed symbols), using two different buffers. The midpoints of the urea denaturation curves are given at various temperatures.

values from urea unfolding. We are not sure why. We think it is prudent to measure $\Delta(\Delta G)$ by both urea and thermal unfolding studies whenever possible.

An ingenious method for determining a urea denaturation curve in a single experiment using electrophoresis (39) is described in Chapter 10. *Figure 5a* shows a urea gradient gel for RNase T1. $[\text{Urea}]_{1/2}$ values can be estimated from these gels reasonably well using a ruler. In *Figure 5b*, we show that $[\text{urea}]_{1/2}$ values determined by urea gradient gel electrophoresis are in remarkably good agreement with values determined by conventional urea unfolding curves. These results show that RNase T1 is much more stable in a phosphate buffer than it is in a Tris buffer; they also show the marked dependence of the stability of RNase T1 on temperature. Urea gradient gels are useful for a variety of purposes and provide a quick means of assessing possible changes in stability. In addition, they can be used to obtain a rough estimate of $\Delta G(H_2O)$ (40).

6. CONCLUDING REMARKS

There is presently considerable interest in gaining a better understanding of the forces that contribute to the stability of globular proteins, so that more stable and useful proteins may be produced, especially through the use of genetic engineering techniques. The hope is that it will be possible to draw conclusions by comparing the thermodynamic stabilities of proteins that differ in sequence by a single amino acid. It has been emphasized that it will generally be necessary to determine high resolution crystal structures of the proteins being compared (41). It has not been emphasized how important it is that considerable care be exercised in determining the differences in conformational stability before attempting to draw conclusions, or before deciding to determine a crystal structure to investigate an intriguing finding.

With any study of protein conformation, it is important to determine whether protein instability is due to unfolding of the native conformation or chemical modification of the covalent structure of the protein. The two are often linked, as the rate of covalent modification usually depends upon the conformational state of the protein. In order to separate the two, it is important that covalent modifications not take place while measuring the conformational stability of the protein by the procedures used here.

In the past, we preferred results from urea and GdmCl unfolding over thermal unfolding studies for several reasons: the product of unfolding seemed to be better characterized, unfolding was more likely to closely approach a two-state mechanism and unfolding was more likely to be completely reversible. However, in the cases that have been carefully investigated, the two techniques seem to give estimates of $\Delta G(H_2O)$ that are in good agreement. After getting the equipment set up and learning the procedures, either type of unfolding curve can be determined and analysed in a single day. Consequently, the safest course is to determine both types of unfolding curves whenever possible. Also, we generally prefer urea over GdmCl because salt effects can be investigated, and sometimes reveal interesting information (42). However, GdmCl is a more potent denaturant and this is sometimes very useful.

Finally, we should emphasize that this is a 'how to do it chapter' and we ignored or only briefly mentioned some interesting and unresolved questions. See references (2–10) for more information on the topics discussed here.

7. ACKNOWLEDGEMENTS

We are grateful to Gaston Matthyssens, Patrick Stanssens and Chris Thoen of Plant Genetic Systems in Gent, Belgium for supplying us with the RNase T1 mutants used to obtain the data in *Table 8*; to Lisa Lohman for preparing the figures; and to Carus Lortz for performing some important tasks. We are also grateful to NIH (Grant GM 37039) and the Robert A. Welch Foundation (Grant A 1060) for financial support.

8. REFERENCES

1. O'Sullivan,C. and Tompson,F.W. (1890) *J. Chem. Soc.*, **57**, 834.
2. Kauzmann, W. (1959) *Adv. Protein Chem.*, **14**, 1.
3. Tanford, C. (1968) *Adv. Protein Chem.*, **23**, 121.
4. Tanford, C. (1970) *Adv. Protein Chem.*, **24**, 1.
5. Pace, C. N. (1974) *Crit. Rev. Biochem.*, **3**, 1.
6. Privalov, P. L. (1979) *Adv. Protein Chem.*, **33**, 167.
7. Pfeil, W. (1981) *Mol. Cell. Biochem.*, **40**, 3.
8. Creighton, T. E. (1978) *Progr. Biophys. Mol. Biol.*, **33**, 231.
9. Schellman, J. A. (1987) *Annu. Rev. Biophys. Biophys. Chem.*, **16**, 115.
10. Goldenberg, D. P. (1988) *Annu. Rev. Biophys. Biophys. Chem.*, **17**, 481.
11. Wong, K. P. and Tanford, C. (1973) *J. Biol. Chem.*, **248**, 8518.
12. Saito, Y. and Wada, A. (1983) *Biopolymers*, **22**, 2105.
13. Nozaki, Y. (1972) In *Methods in Enzymology*. Hirs, C. H. W. and Timasheff, S. N. (eds), Academic Press, New York, Vol. 26, p. 43.
14. Prakash, V., Loucheux, C., Scheufele, S., Gorbunoff, M. J. and Timasheff, S. N. (1981) *Arch. Biochem. Biophys.*, **210**, 455.
15. Hagel, P., Gerding, J. J. T., Fieggen, W. and Bloemendal, H. (1971) *Biochim. Biophys. Acta*, **243**, 366.
16. Stark, G. R. (1965) *Biochemistry*, **4**, 1030.
17. Kawahara, K. and Tanford, C. (1966) *J. Biol. Chem.*, **241**, 3228.
18. Warren, J. R. and Gordon, J. A. (1966) *J. Phys. Chem.*, **67**, 1524.
19. Pace, C. N. (1986) In *Methods in Enzymology*. Hirs, C. H. W. and Timasheff, S. N. (eds), Academic Press, New York, Vol. 131, p. 266.
20. Waheed, A., Qasim, M. A. and Salahuddin, A. (1977) *Eur. J. Biochem.*, **76**, 383.
21. Azuma, T., Hamaguchi, K. and Migita, S. (1972) *J. Biochem.*, **72**, 1457.
22. Matthews, C. R. and Crisanti, M. M. (1981) *Biochemistry*, **20**, 784.
23. Rowe, E. S. and Tanford, C. (1973) *Biochemistry*, **12**, 4822.
24. Lumry, R., Biltonen, R. and Brandts, J. F. (1966) *Biopolymers*, **4**, 917.
25. Privalov, P. L. and Khechinashvili, N. N. (1974) *J. Mol. Biol.*, **86**, 665.
26. Freire, E. and Biltonen, R. L. (1978) *Biopolymers*, **17**, 463.
27. Kim, P. S. and Baldwin, R. L. (1982) *Annu. Rev. Biochem.*, **51**, 459.
28. Herskovits, T. T. (1967) In *Methods in Enzymology*. Hirs, C. H. W. (ed.), Academic Press, New York, Vol. 11, p. 748.
29. Greene, R. F. and Pace, C. N. (1974) *J. Biol. Chem.*, **249**, 5388.
30. Ahmad, F. and Bigelow, C. C. (1982) *J. Biol. Chem.*, **257**, 12935.
31. Schellman, J. A. (1978) *Biopolymers*, **17**, 1305.
32. Becktel, W. J. and Schellman, J. A. (1987) *Biopolymers*, **26**, 1859.
33. Shiao, D. F., Lumry, R. and Fahey, J. (1971) *J. Am. Chem. Soc.*, **93**, 2024.
34. Pace, C. N. and Tanford, C. (1968) *Biochemistry*, **7**, 198.
35. Brandts, J. F. (1964) *J. Am. Chem. Soc.*, **86**, 4291.
36. Jackson, W. M. and Brandts, J. F. (1970) *Biochemistry*, **9**, 2294.

37. Edelhoch, H. and Osborne, J. C., Jr. (1976) *Adv. Protein Chem.*, **30**, 183.
38. Shortle, D. and Meeker, A. K. (1986) *Proteins*, **1**, 81.
39. Creighton, T. E. (1986) In *Methods in Enzymology*. Hirs, C. H. W. and Timasheff, S. N. (eds), Academic Press, New York, Vol. 131, p. 156.
40. Hollecker, M. and Creighton, T. E. (1982) *Biochim. Biophys. Acta*, **701**, 395.
41. Matthews, B. (1987) *Biochemistry*, **26**, 6885.
42. Pace, C. N. and Grimsley, G. G. (1988) *Biochemistry*, **27**, 3242.