Plant Molecular Biology
Plasmids (2nd edition)
★ Platelets
Pollination Ecology
Postimplantation Mammalian Embryos
Preparative Centrifugation
Prostaglandins and Related Substances
Protein Blotting
Protein Engineering
★ Protein Function (2nd edition)
Protein Phosphorylation
Protein Purification Applications
Protein Purification Methods
Protein Sequencing
★ Protein Structure (2nd edition)
★ Protein Structure Prediction
Protein Targeting
Proteolytic Enzymes
★ Pulsed Field Gel Electrophoresis
Radioisotopes in Biology
Receptor Biochemistry
Receptor–Ligand Interactions
RNA Processing I and II
★ Subcellular Fractionation
Signal Transduction
Solid Phase Peptide Synthesis
Transcription Factors
Transcription and Translation
Tumour Immunobiology
Virology
Yeast

Protein Structure

A Practical Approach

Edited by

T. E. CREIGHTON

European Molecular Biology Laboratory, Heidelberg, Germany

2º edição
1997

IRL PRESS
at
OXFORD UNIVERSITY PRESS
Oxford New York Tokyo

12

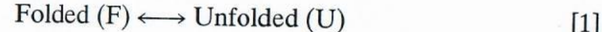
Measuring the conformational stability of a protein

C. NICK PACE and J. MARTIN SCHOLTZ

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 (2–8). 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:



$$\text{Folded (F)} \longleftrightarrow \text{Unfolded (U)} \quad [1]$$

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.[1] 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 are helping us gain a better understanding of the forces that determine the conformations of proteins and to optimize their stabilities (9,10).

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 $\Delta G(H_2O)$, to determine the stability curve for a protein (11), 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.

[1]Energies are given here in units of calories (cal); to convert to joules (J), multiply by 4.18 J/cal.

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 absorbance spectroscopy, fluorescence, and circular dichroism (CD), which are described in Chapter 11. Other techniques used are biological activity measurements, nuclear magnetic resonance (NMR), viscosity, and other hydrodynamic methods (Chapters 8 and 9). Only spectroscopic techniques will be discussed here; biological activity measurements present special problems (12). 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 UV absorbance spectra of folded and unfolded ribonuclease T1 (RNase T1) are shown in *Figures 1a* and *1b*. 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 absorbance spectra in *Figure 1b* required considerably more RNase T1 than the fluorescence spectra in *Figure 1a*. This is generally true; when fluorescence or far-UV CD can be used, less protein will be required than with other techniques. UV absorbance spectroscopy and near-UV CD generally require greater amounts of material, although the amounts depend on the wavelength chosen to follow unfolding. NMR generally requires the greatest amount of protein. In return, however, NMR will generally yield considerably more detailed structural information than the other techniques. Thus, the technique you choose may be limited by the amount of protein that you have available for the experiments.

Secondly, it is necessary to pick a technique and a wavelength for which the spectra of the folded and unfolded conformations differ significantly. With fluorescence, we chose 320 nm where the intrinsic fluorescence of folded RNase T1 is approximately sixfold greater than that of the unfolded protein (*Figure 2a*). With UV absorbance, we chose 286 nm and 292 nm which are maxima in the difference spectrum (*Figure 2b*). 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 will be.

The spectral changes observed upon unfolding often depend upon different features of protein structure. For example, fluorescence and UV absorbance spectroscopy respond to changes in the environment of the tryptophan (Trp) and tyrosine (Tyr) 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 measure-

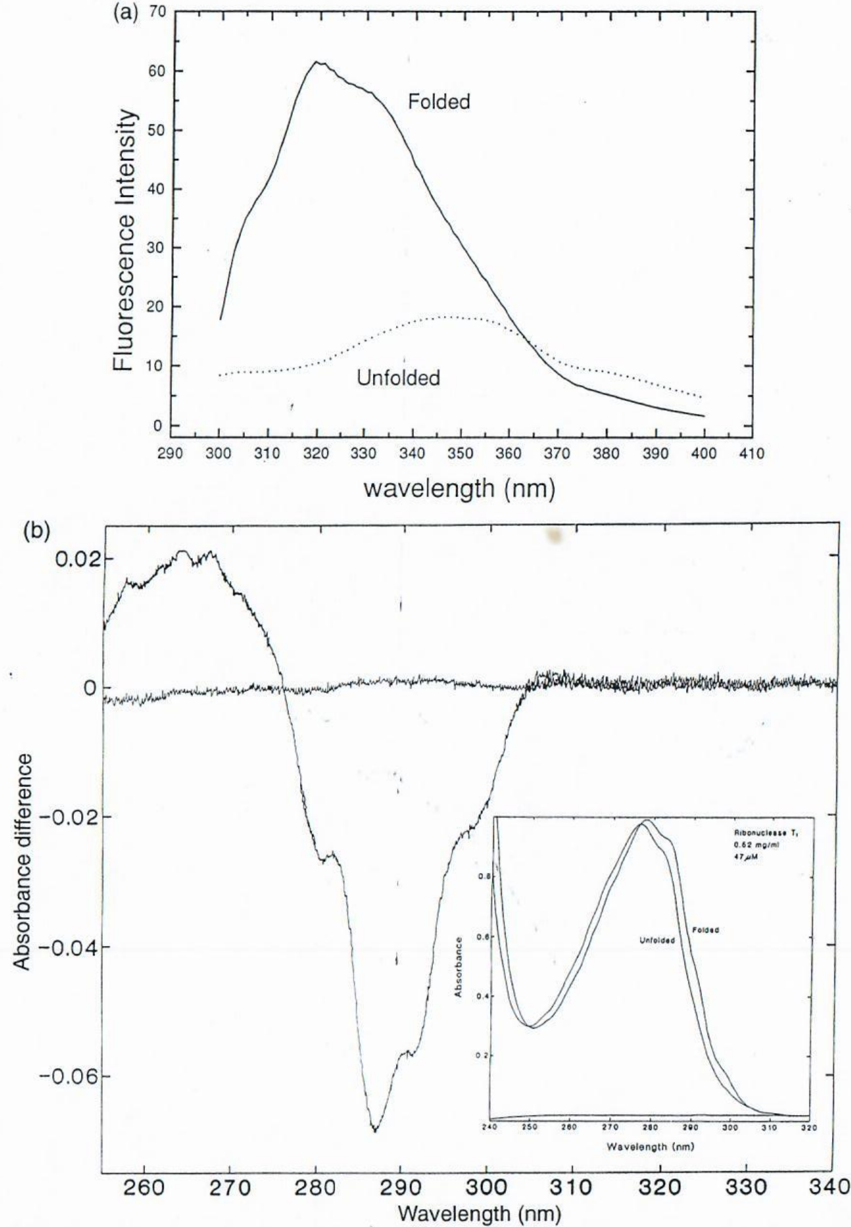


Figure 1. Spectra of folded and unfolded (8 M urea) RNase T1. (a) Intrinsic fluorescence emission spectra with excitation at 278 nm. (b) Difference in absorbance between the two absorbance spectra shown in the inset. The RNase T1 concentration was 0.01 mg/ml for (a) and 0.5 mg/ml for (b).

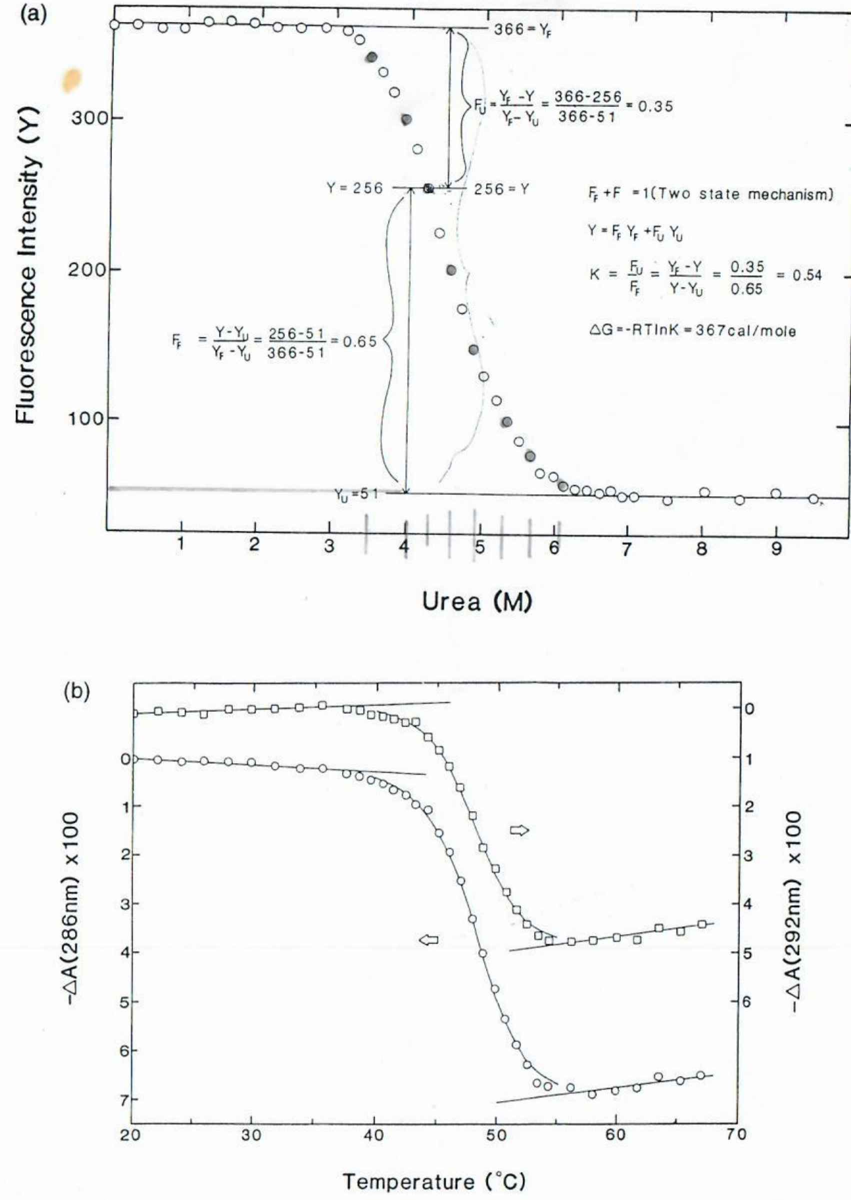


Figure 2. Urea- and heat-induced unfolding transitions of RNase T1. (a) Fluorescence intensity as a function of urea concentration at 25 °C in 30 mM Mops buffer pH 7. The equations used to analyse an unfolding curve for a two-state mechanism are given. (b) Absorbance difference as a function of temperature in 30 mM Mops buffer pH 7. From ref. 48.

ments are less useful for following thermal unfolding, because the pre- and post-transition baselines are steep and temperature-sensitive. UV absorbance measurements are less convenient for following urea and GdmCl unfolding, and the pre- and post-transition baselines are steeper than with most of the other techniques. In general, unfolding curves with steeper pre- and post-transition baselines lead to larger errors in the parameters determined in the analysis (see Section 4).

3. Determining unfolding curves

Figure 2a and *b* show typical urea and thermal unfolding curves. In these cases, fluorescence was used to follow urea unfolding and UV absorbance was used to follow thermal unfolding, but we will refer to the physical parameter used to follow unfolding as y for the discussion that follows. The curves can be conveniently divided into three regions:

(a) The pre-transition region, which shows how y for the folded protein, y_F, depends upon the denaturant.

(b) The transition region, which shows how y varies as unfolding occurs.

(c) 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 less useful 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 is longest at the midpoint of the transition and decreases in both the pre- and post-transition regions. 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 three 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 and the higher the temperature. Similar observations have been made with many other proteins. For this reason, thermal denaturation curves should be determined as quickly as possible.

Proteins that contain free sulfydryl groups present special problems. If the protein contains only free –SH groups and no disulfide bonds, then a reducing agent such as 10 mM dithiothreitol (DTT) can be added to ensure that no disulfide bonds form during the experiments. For proteins containing both free –SH groups and disulfide bonds, disulfide interchange can occur and this may lead to irreversibility. Disulfide 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. 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 (*Table 1* and *Protocol 1*).

Protocol 1. Example of preparation of a urea stock solution[a]

1. Add approx. 60 g of urea to a tared beaker and weigh (59.91 g).[b] Now add 0.69 g of Mops buffer (sodium salt), 1.8 ml of 1 M HCl, 52 ml of distilled water, and weigh the solution again (114.65 g).[b]
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, ΔN = 1.4173–1.3343 = 0.0830.[b]
5. Calculate the urea molarity from ΔN using the equation given in *Table 1*: M = 10.08.[b]
6. Calculate the urea molarity based on the recorded weights. The density is calculated with the equation given in *Table 1*: weight fraction urea (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.[b]
7. The molarities calculated in steps 5 and 6 differed by less than 1%, so this solution was used to determine a urea unfolding curve.

[a] 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.
[b] Experimental result obtained in this example.

The method for determining an unfolding curve is given in *Protocol 2*. 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 practice to measure the pH of the solutions in the transition region. 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 (13). In addition, it is possible to use titration methods to determine urea or GdmCl denaturation curves (14).

Protocol 2. Determining a urea or GdmCl unfolding curve

1. Prepare three solutions: a denaturant stock solution as described in *Protocol 1*, a protein stock solution, and a buffer solution.
2. Prepare the solutions on which measurements will be made volumetrically (e.g. with Rainin EDP2 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 was measured.
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. Measure the experimental parameter being used to follow unfolding on the solutions in order of ascending denaturant concentration. Leave the cuvette in the spectrophotometer, and do not rinse between samples. Simply remove the old solution carefully with a Pasteur pipette with plastic tubing attached to the tip, and then add the next sample. Leaving the cuvette in position improves the quality of the measurements, and the error introduced by the small amount of old solution is negligible.
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.

Protocol 2. *Continued*

7. Examples of the experimental results obtained are given in *Table 2*.
8. Analyse the results as described in *Protocol 5*.

Both urea and GdmCl can be purchased commercially in highly purified forms (e.g. from United States Biochemical). 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 for recrystallization when it is necessary (15). A procedure for purifying urea has also been described (16). GdmCl solutions are stable for months, but urea solutions slowly decompose to form cyanate and ammonium ions (17) in a process accelerated at high pH. The cyanate ions can react with amino groups on proteins (18). Consequently, a fresh urea stock solution should be prepared for each unfolding curve and used within one 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 *Protocol 1*. 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*.

Table 1. Information for preparing urea and GdmCl stock solutions

Property	Urea	GdmCl
Molecular weight	60.056	95.533
Solubility (25 °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 (19).
[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 ref. 20, and the equation for GdmCl solutions is from ref. 15.

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. Modern instruments may employ thermoelectric devices, such as the Peltier design, that allow for accurate and reliable control over the sample temperature. If the thermoelectric cells are not available, many instruments provide devices that allow water from a constant temperature bath to be circulated to maintain the temperature of the cell. Alternatively, 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 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. Consequently, the solution should be stirred. An alternative 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 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 (*Protocol 3*). Consequently, less protein is required than for urea and GdmCl unfolding experiments. It is a good idea to filter (e.g. with a 0.65 μm 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 imperative 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 remeasure the pH of the solution.

Protocol 3. 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, and acetate and formate buffers are good at lower pH values. In general, ΔH is small for carboxyl groups and large for amino groups.
2. Make certain the instrument to be used to follow unfolding is thoroughly warmed-up and that the system for maintaining and monitoring the temperature of the 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 a 0.65 µm filter.
4. Carefully measure the experimental parameter, y, for the blank solution (all components except the protein) 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 through the transition region. These measurements should be made as quickly as possible, but it is essential that the system reach thermal and chemical equilibrium at each new temperature before the measurement is made. Generally the equilibration time is best determined 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 the T_m. If you are attempting to determine ΔC_P, measurements should be made with great care and 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 remeasure y. If y for the blank is identical to that determined at the lowest temperature, your experiment is probably successful. If your instrument is not very stable, it is necessary to monitor the blank more carefully during the course of the experiment. These blank measurements are very important, especially with single-beam instruments.
9. Cool the sample back down to the starting temperature and measure y again. If y has changed appreciably, this is evidence for a lack of reversibility of the thermal folding transition.
10. Remove the sample and check the pH of the protein solution after the experiment.
11. Analyse the data as described in *Protocol 6*. *Figure 2b* shows a typical thermal unfolding curve and *Protocol 6* shows an analysis of thermal unfolding data.

4. Analysing unfolding curves

The unfolding of many small globular proteins has been found to approach closely a two-state folding mechanism, such as that shown in Equation 1. *Protocol 4* gives a summary of methods that can be used to gain more information about the folding mechanism.

Protocol 4. Investigating the unfolding mechanism

1. If an unfolding curve shows more than one transition, unfolding is more complex than a two-state mechanism and the analysis of the data is more complicated (21). This behaviour is frequently observed with multidomain proteins where the domains unfold independently (13,22,23).
2. If a single-stage unfolding curve is observed, it does not prove that unfolding is a two-state mechanism (24,25).
3. Further insight can be gained by using different techniques or probes to following unfolding. Non-coincidence of plots of f_U as a function of temperature or denaturant concentration determined by different techniques indicates that significant amounts of intermediates are present at equilibrium and hence a simple two-state mechanism can not be used in analysing the data (12,13,24). Unfortunately, coincidence of the unfolding curves is only consistent with, but does not prove, that unfolding follows a two-state mechanism (26).
4. The best evidence that thermal unfolding follows a two-state mechanism is to show that ΔH determined by the van't Hoff relationship (Equation 7) is identical to that determined calorimetrically, where a specific folding mechanism is not required to determine ΔH_{cal}. When $\Delta H_{vH} < \Delta H_{cal}$, it is clear evidence that significant concentrations of intermediates are present at equilibrium (6,26–28).
5. The mechanism of folding of globular proteins is an area of great interest to biochemists and has been reviewed recently (29). For our purposes here, we need not be concerned with kinetic folding intermediates, only those intermediates present at equilibrium.

A two-state folding mechanism will be assumed 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 gives:

$$f_U = (y_F - y)/(y_F - y_U). \quad [2]$$

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

$$K = f_U/f_F = f_U/(1 - f_U) = (y_F - y)/(y - y_U) \quad [3]$$

and

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

where R is the gas constant (1.987 cal mol^{-1} K^{-1}) 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*. Generally, y_F and y_U are found to be linear functions of temperature or denaturant concentration and a least-squares analysis can be used to determine the linear expressions for y_F and y_U.

The calculation of f_U, K, and G from data such as those shown in *Figure 2a* and *b* is illustrated in *Protocols 5* and *6*. 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 within the range ± 1.5 kcal/mol.

4.1 Urea and GdmCl unfolding

Protocol 5. Analysis of a urea unfolding curve

1. Data points from the transition region of an unfolding curve are listed in *Table 2*. 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 below for the pre- and post-transition regions.[a]
2. A least-squares analysis of this data yields the following for Equation 5: ΔG = 5680–(1218) × [urea]. Thus, $\Delta G(H_2O)$ = 5.68 kcal mol^{-1}, m = 1218 cal mol^{-1} M^{-1}, and $[urea]_{1/2}$ = 4.66 M.
3. All of these parameters should be reported to characterize the results from a urea denaturation curve.

[a] $y_F = 112 + 19 \times [urea]$, and $y_U = 765$.

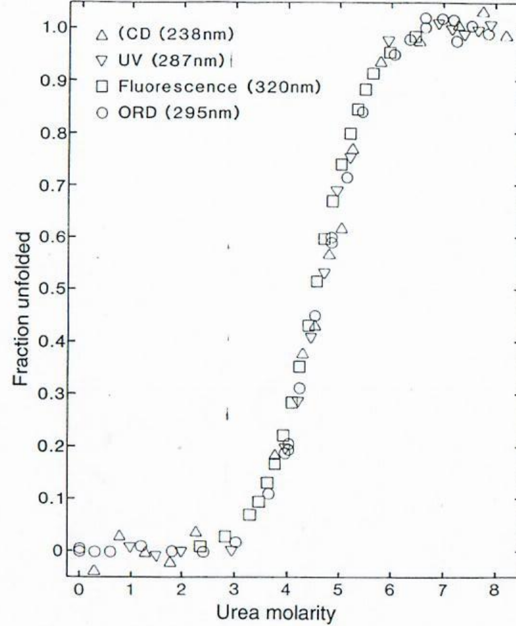


Figure 3. Fraction of RNase T1 unfolded, f_U, as a function of urea concentration. The values of f_U were calculated from data such as those in *Figure 2a* using Equation 2. From ref. 48.

Data such as those shown in *Figure 2a* have been analysed as described in *Protocol 5* and typical results are shown in *Figures 3* and *4*, and *Table 3*. 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 baselines it is often useful to show just f_U as a function of denaturant in the transition region rather than a complete unfolding curve.

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

Table 2. Data of a urea unfolding curve

[Urea] (M)	y	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

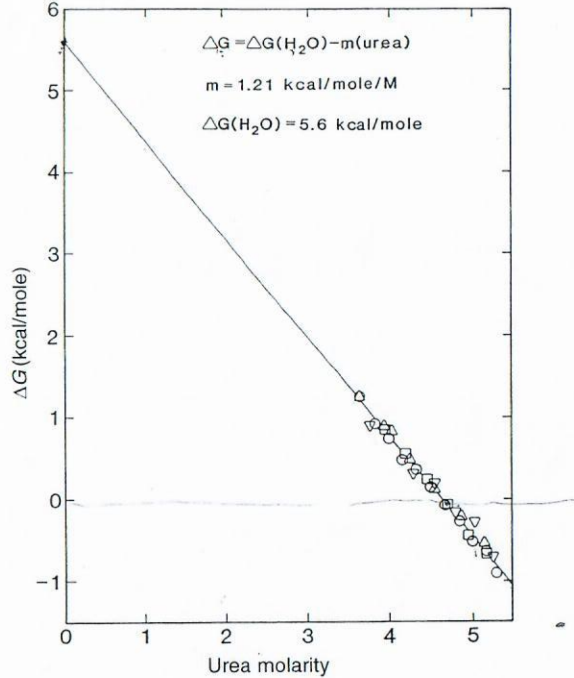


Figure 4. ΔG for RNase T1 unfolding as a function of urea concentration. The experimental values of ΔG were calculated from data such as those in *Figure 2a* using Equation 4. The solid line represents Equation 5 with $\Delta G(H_2O)$ = 5.6 kcal mol^{-1} and m = 1220 cal mol^{-1} M^{-1}. From ref. 48.

have been obtained with many other proteins (30,31). 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:

$$\Delta G = \Delta G(H_2O) - m\,[D] \qquad [5]$$

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

If the linear extrapolation method is assumed for the analysis, it allows a convenient method to be used for analysing a urea or GdmCl denaturation curve (32). The transition region can then be characterized by two parameters, m, which measures the steepness of the transition, and $[D]_{1/2}$, which measures the midpoint of the transition, and the pre- and post-transition

$\Delta G = m[D_{1/2}] - m[D]$

Table 3. Example of experimental results obtained with a urea unfolding experiment

Urea (ml)[a]	Buffer (ml)[b]	Protein (ml)[c]	Urea (M)[d]	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.

regions can also be characterized by two parameters, y_F, and y_U for the intercepts, and m_F and m_U for the slopes. Consequently, the following equation can represent the entire denaturation curve:

$$y = \{(y_F + m_F[D]) + (y_U + m_U\,[D]) \times \exp[m \times ([D]-[D]_{1/2})/RT]\}/(1 + \exp[m \times ([D]-[D]_{1/2})/RT]). \qquad [6]$$

The best fit of the six parameters and their confidence intervals can be determined by using the non-linear least-squares program described by Johnson and Frasier (33). This method can also be used to analyse thermal denaturation curves.

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 ref. 21 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.

4.2 Thermal unfolding

We will now consider how to analyse a thermal unfolding curve, such as the one shown in *Figure 2b*, in order to determine the conformation stability, $\Delta G(H_2O)$ (see ref. 11 for an excellent discussion of this subject). This

requires extrapolating ΔG measurements from the narrow temperature range where unfolding occurs to a reference temperature, T, such as 25 °C. *Protocol 6* illustrates the calculation of K and ΔG from results such as those in *Figure 2b*.

Protocol 6. Analysis of a thermal unfolding curve

1. The data are from a thermal unfolding curve for ribonuclease T1 determined under the same conditions as the urea unfolding curve analysed in *Protocol 5*. Values of f_U, K, and ΔG were calculated using Equations 2, 3, and 4 respectively (see *Table 4*). A least-squares analysis of the pre- and post-transition regions gave $y_F = 87.2 + 0.66 \times T$ and $y_U = 646$.
2. The slope of the plot of ΔG versus T is –300.3 cal mol^{-1} K^{-1}, and $T = T_m = 48.3$ °C at $\Delta G = 0$. Since $\Delta G = 0$ at T_m, $H_m = T_m \times \Delta S_m$. Therefore, $\Delta H_m = (-1) \times (48.3 + 273.2) \times (-300.3) = 96.5$ kcal mol^{-1}. The slope of the van't Hoff plot (Equation 7) is –48 631. Therefore, $\Delta H_m = -(1.987) \times (-48\,631) = 96.6$ kcal mol^{-1}. The van't Hoff plot shows no curvature over the temperature range of the transition, so clearly these data can not be used to determine ΔC_P with Equation 8.
3. To estimate ΔG at 25 °C use Equation 9. If it is assumed that $\Delta C_P = 0$, the second term in the equation is 0, and ΔG(25 °C) = 7.0 kcal mol^{-1}. If the average of eight experimental values is used, ΔC_P = 1460 cal mol^{-1} K^{-1} and ΔG(25 °C) = 5.7 kcal mol^{-1}. If ΔC_P is calculated using Equation 10, ΔC_P = 1680 cal mol^{-1} K^{-1} and ΔG(25 °C) = 5.5 kcal mol^{-1}. Note the excellent agreement with the value of ΔG(25 °C) = 5.7 kcal mol^{-1} from the analysis of a urea unfolding curve in *Protocol 5*.

The familiar method used to obtain the enthalpy change, ΔH, from these measurements is with the van't Hoff equation:

$$d(\ln K)/d(1/T) = -\Delta H/R \quad [7]$$

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

$$d(\Delta H)/d(T) = C_P(U) - C_P(F) = \Delta C_P \quad [8]$$

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. Therefore, both ΔC_P and ΔH are required to calculate ΔG as a function of temperature. Since ΔH is needed at only a single temperature, 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

Table 4. Thermal unfolding data for ribonuclease T1

T (°C)	y	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			

parameters, the equation used to calculate ΔG at any temperature T, $\Delta G(T)$, is:

$$\Delta G(T) = \Delta H_m(1 - T/T_m) - \Delta C_P[(T_m - T) + T \ln (T/T_m)]. \quad [9]$$

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 to use $T_m = T$ at $\Delta G = 0$, and $\Delta H_m = -1 \times (T_m \text{ in K}) \times$ (slope at T_m). Calculations based on this approach and on the van't Hoff equation are illustrated in *Protocol 6*.

The determination of ΔC_P is more difficult. One approach is to use a calorimeter to determine ΔC_P. In our best cases, ΔC_P can be measured directly to about ± 10% with a differential scanning microcalorimeter. A non-calorimetric approach that has been used successfully with T4 lysozyme (11), the chymotrypsinogen family (34) and other proteins (35), is to measure T_m and ΔH_m as a function of pH. T_m usually varies with pH with a concomitant variation in ΔH_m. The slope of the plot of ΔH_m versus T_m will provide ΔC_P, as shown by Equation 8. This assumes that ΔH and ΔC_P do not vary significantly with pH, which appears to be the case normally (26,33,34). This approach can also be used to determine ΔC_P with a calorimeter (27). In favourable situations, ΔC_P can be determined directly using data such as those shown in *Figure 2b*. Equations 7 and 8 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 (36), but we have not been able to determine a reasonable and reproducible value of ΔC_P for RNase T1 using this method. Brandts

used this approach to determine ΔC_P for chymotrypsinogen in his pioneering studies in this area (37). All the methods described here give comparable values of ΔC_P for chymotrypsinogen (38).

Another useful technique for determining ΔC_P without a calorimeter has been described by Pace and Laurents (39). In this method, the $\Delta G(H_2O)$ values from urea denaturation experiments at low temperatures are combined with ΔG values from the transition region of a thermal unfolding curve. A least-squares fit of the $\Delta G(T)$ data to Equation 9 provides a measure of ΔH_m, T_m, and ΔC_P. This method is illustrated in *Figure 5* for a small protein, HPr (40). In favourable cases, such as this, when the temperature of maximum stability is above 0 °C, the ΔC_P value can be determined with an accuracy of 5–10%. If the temperature of maximum stability is below 0 °C, there is more uncertainty in the ΔC_P value determined from this method; however, since this method does not employ a calorimeter, almost any laboratory can use this method to determine ΔC_P using the simple methods discussed in this chapter.

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 (41). A recent report (42) provides a simple way to estimate the ΔC_P for the protein unfolding reaction using only information from the primary sequence of the protein:

$$\Delta C_P \approx 172 + 17.6 \times N - 164 \times SS \text{ (in cal mol}^{-1}\text{ K}^{-1}) \quad [10]$$

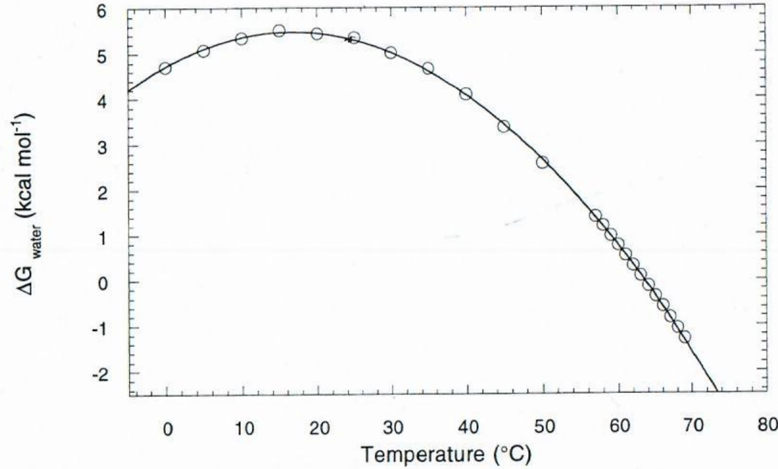


Figure 5. A stability curve for the small bacterial protein HPr using the method of Pace and Laurents (39). The data at $T \leq 50$ °C are $\Delta G(H_2O)$ values from urea denaturation curves, while the data at higher temperatures are from a thermal unfolding curve in the absence of denaturant. The line represents the best fit to Equation 8 with $T_m = 63.5 \pm 0.2$ °C, $\Delta H_m = 70.4 \pm 1.4$ kcal mol^{-1}, and $C_P = 1.45 \pm 0.08$ kcal mol^{-T} K^{-1}. The data are from Nicholson and Scholtz (40).

where N is the number of amino acid residues and SS is the number of disulfide cross-links in the protein. If the structure of your protein is known, there are several other equations that can be used to provide even better estimates for ΔC_P (see ref. 42 for a comparison of the various empirical methods). We have used this approach in *Protocol 6* to obtain a value of ΔC_P for RNase T1 folding to illustrate the use of Equation 9.

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 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. *Table 5* presents results from urea and thermal unfolding studies of wild-type RNase T1 and a mutant that differs 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 (21). 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 change these values. Consequently, differences in stability determined by comparing the $\Delta G(H_2O)$ values can have large errors. 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 5*. *Table 6* illustrates 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 at 25 °C, but a much higher concentration of GdmCl is needed to denature lysozyme because the m value is much smaller. Likewise, lysozyme and cytochrome c unfold at about the same temperature, even though lysozyme has a much larger value of $\Delta G(H_2O)$ at 25 °C.

A second approach is to compare the proteins in the presence of urea by taking the difference between the $[D]_{1/2}$ 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 5*. The rationale here is 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. However, substantial differences in m values between proteins differing by only one amino acid in sequence have been observed with some proteins (43).

Becktel and Schellman have provided a simple method for the analysis of

Table 5. Analysis of urea unfolding curves (at 25°C) and thermal unfolding curves for wild-type RNase T1 and a mutant which differs by one amino acid

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	6.41		1210	5.30		
Tyr11Phe	4.52	1.9	1270	3.56	1.74	2.2
Protein	ΔS_m[g]	ΔH_m[h]	T_m[i]	$\Delta(T_m)$[j]	$\Delta(\Delta G)$[k]	
Wild-type	339	110	50.9			
Tyr11Phe	317	101	44.9	6.0	2.0	

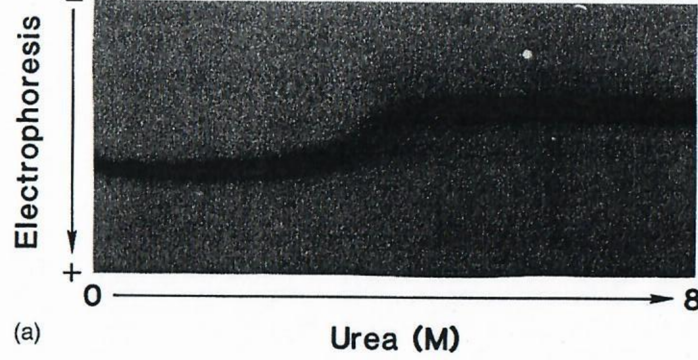
[a] From Equation 5, in kcal mol^{-1}.
[b] Difference between the $\Delta G(H_2O)$ values in kcal mol^{-1}.
[c] From Equation 5, in cal mol^{-1} M^{-1}.
[d] Midpoint of the urea unfolding curve in M.
[e] Difference between the $[urea]_{1/2}$ values in M.
[f] From $[urea]_{1/2} \times 1240$ (the average of the two m values) in kcal mol^{-1}.
[g] Negative of the slope of ΔG versus T at T_m in cal mol^{-1} K^{-1}.
[h] $\Delta H_m = T_m(K) \times (\Delta S_m)$ in kcal mol^{-1}.
[i] Midpoint of the thermal unfolding curve in °C.
[j] Difference between the T_m values.
[k] $\Delta(\Delta G) = (\Delta T_m) \times \Delta S_m(wt) = \Delta T_m \times [\Delta H_m(wt)/T_m(wt)]$, where ΔT_m is the difference in the midpoints of the thermal transitions and $\Delta S_m(wt)$ and $\Delta H_m(wt)$ are the values for the better characterized, usually wild-type, protein (see ref. 11).

thermal denaturation curves in order to determine differences in stability (11). We have used their suggestions in calculating the $\Delta(\Delta G)$ values from thermal unfolding in *Table 5*. For these proteins, the estimates of $\Delta(\Delta G)$ are in good agreement, which is reassuring. This good agreement, however, is not always observed. 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 (44) is described in Chapter 8. *Figure 6a* shows a urea gradient gel for RNase T1. $[D]_{1/2}$ values can be estimated from these gels reasonably well using a ruler. In *Figure 6b*, we show that $[D]_{1/2}$ values determined by urea gradient gel electrophoresis are in surprisingly good agreement with values determined by conventional urea unfolding

Table 6. Unfolding curve midpoints of different proteins

Protein	$[GdmCl]_{1/2}$ (M)	$\Delta G(H_2O)$ (kcal mol^{-1}) at 25°C
Myoglobin	1.7	12
Lysozyme	3.1	12
Protein	T_m (°C)	$\Delta G(H_2O)$ (kcal mol^{-1}) at 25°C
Cytochrome c	80	8
Lysozyme	80	12



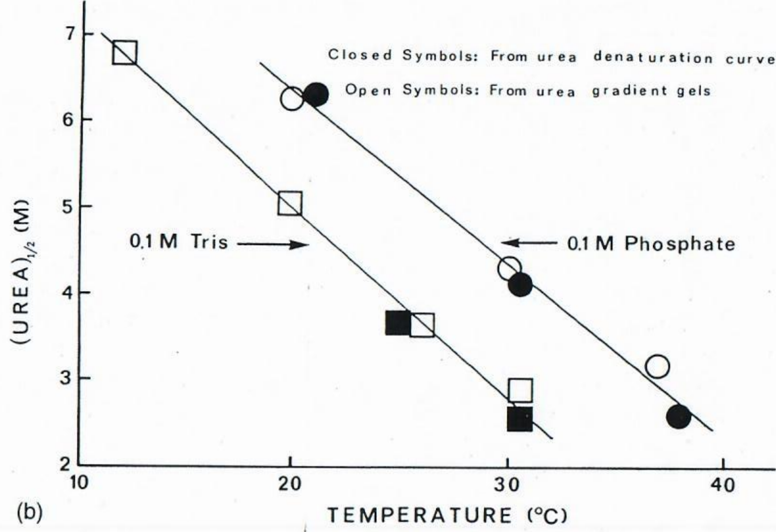


Figure 6. Unfolding of RNase T1 measured by urea gradient electrophoresis. (a) An example of a urea gradient gel (see Chapter 8). (b) Comparison of the unfolding measured by urea gradient gels (open symbols) and by conventional urea unfolding curves (closed symbols), using two different buffers. The midpoints of the urea denaturation curves are given at various temperatures.

curves. These results show that RNase T1 is much more stable in a phosphate buffer than 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 $\Delta G(H_2O)$ (45).

6. Concluding remarks

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 (48). 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 (46). However, GdmCl is a more potent denaturant, and this is sometimes very useful.

Studies of mutant proteins over the past few years have improved our understanding of the forces that contribute to the conformational stability of globular proteins (10). In addition, our understanding of how to make proteins more stable and more useful is improving, as is our ability to design new proteins (47). However, much remains to be learned and the techniques described in this chapter are essential tools for doing so.

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 refs 2–11 for more information on the topics discussed here.

Acknowledgements

We thank Janis Chmiel for her excellent assistance in preparing this chapter. We also thank the NIH (GM 37039 and GM 52483), the Robert A. Welch Foundation (A-1060 and A-1281), the American Cancer Society Junior Faculty Research Award program (JFRA-577), and the Tom and Jean McMullin Professorship for financial support of the work in our laboratories.

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. (1975). *Crit. Rev. Biochem.*, **3**, 1.
6. Privalov, P. L. (1979). *Adv. Protein Chem.*, **33**, 167.
7. Pace, C. N. (1990). *Trends Biochem. Sci.*, **15**, 14.
8. Creighton, T. E. (ed.) (1992). *Protein folding*. Freeman, New York.
9. Shortle, D. (1992). *Q. Rev. Biophys.*, **25**, 2.
10. Matthews, B. W. (1993). *Adv. Protein Chem.*, **46**, 249.

11. Becktel, W. J. and Schellman, J. A. (1987). *Biopolymers*, **26**, 1859.
12. Wong, K. P. and Tanford, C. (1973). *J. Biol. Chem.*, **248**, 8518.
13. Saito, Y. and Wada, A. (1983). *Biopolymers*, **22**, 2105.
14. Scholtz, J. M. (1995). *Protein Sci.*, **4**, 35.
15. Nozaki, Y. (1972). In *Methods in enzymology* (ed. C. H. W. Hirs and S. N. Timasheff), Vol. 26, p. 43. Academic Press, New York.
16. Prakash, V., Loucheux, C., Scheufele, S., Gorbunoff, M. J., and Timasheff, S. N. (1981). *Arch. Biochem. Biophys.*, **210**, 455.
17. Hagel, P., Gerdling, J. J. T., Fieggen, W., and Bloemendal, H. (1971). *Biochim. Biophys. Acta*, **243**, 366.
18. Stark, G. R. (1965). *Biochemistry*, **4**, 1030.
19. Kawahara, K. and Tanford, C. (1966). *J. Biol. Chem.*, **241**, 3228.
20. Warren, J. R. and Gordon, J. A. (1966). *J. Phys. Chem.*, **67**, 1524.
21. Pace, C. N. (1986). In *Methods in enzymology* (ed. C. H. W. Hirs and S. N. Timasheff), Vol. 131, p. 266. Academic Press, New York.
22. Waheed, A., Qasim, M. A., and Salahuddin, A. (1977). *Eur. J. Biochem.*, **76**, 383.
23. Azuma, T., Hamaguchi, K., and Migita, S. (1972). *J. Biochem.*, **72**, 1457.
24. Kuwajima, K. (1989). *Proteins: Struct., Funct., Genet.*, **6**, 87.
25. Price, N. C. (1994). In *Mechanisms of protein folding* (ed. R. H. Pain), p. 160. IRL Press, Oxford.
26. Lumry, R., Biltonen, R., and Brandts, J. F. (1966). *Biopolymers*, **4**, 917.
27. Privalov, P. L. and Khechinashvili, N. N. (1974). *J. Mol. Biol.*, **86**, 665.
28. Freire, E. and Biltonen, R. L. (1978). *Biopolymers*, **17**, 463.
29. Kim, P. S. and Baldwin, R. L. (1990). *Annu. Rev. Biochem.*, **59**, 631.
30. Greene, R. F. and Pace, C. N. (1974). *J. Biol. Chem.*, **249**, 5388.
31. Ahmand, F. and Bigelow, C. C. (1982). *J. Biol. Chem.*, **257**, 12935.
32. Santoro, M. M. and Bolen, D. W. (1989). *Biochemistry*, **27**, 8063.
33. Johnson, M. L. and Frasier, S. G. (1985). In *Methods in enzymology* (ed. C. H. W. Hirs), Vol. 117, p. 301. Academic Press, New York.
34. Shiao, D. F., Lumry, R., and Fahey, J. (1971). *J. Am. Chem. Soc.*, **93**, 2024.
35. Swint, L. and Robertson, A. D. (1993). *Protein Sci.*, **2**, 2037.
36. Pace, C. N. and Tanford, C. (1986). *Biochemistry*, **7**, 198.
37. Brandts, J. F. (1964). *J. Am. Chem. Soc.*, **86**, 4291.
38. Jackson, W. M. and Brandts, J. F. (1970). *Biochemistry*, **9**, 2294.
39. Pace, C. N. and Laurents, D. V. (1989). *Biochemistry*, **28**, 2520.
40. Nicholson, E. M. and Scholtz, J. M. (1996). *Biochemistry*, **35**, 11369.
41. Edelhoch, H. and Osborne, J. C. Jr. (1976). *Adv. Protein Chem.*, **30**, 183.
42. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995). *Protein Sci.*, **4**, 2138.
43. Shortle, D. (1995). *Adv. Protein Chem.*, **46**, 217.
44. Creighton, T. E. (1986). In *Methods in enzymology* (ed. C. H. W. Hirs and S. N. Timasheff), Vol. 131, p. 156. Academic Press, New York
45. Hollecker, M. and Creighton, T. E. (1982). *Biochim. Biophys. Acta*, **701**, 395.
46. Pace, C. N. and Grimsley, G. G. (1988). *Biochemistry*, **27**, 3242.
47. Betz, S. F., Bryson, J. W., and DeGrado, W. F. (1995). *Curr. Opin. Struct. Biol.*, **5**, 457.
48. Thomson, J. A., Shirley, B. A., Grimsley, G. A., and Pace, C. N. (1989). *J. Biol. Chem.*, **264**, 11614.