Differential scanning microcalorimetry

Alan Cooper

Chemistry Dept., Glasgow University, Glasgow G12 8QQ

Adapted from ref.21: A. Cooper, M. A. Nutley, A. Wadood, Differential scanning microcalorimetry in S. E. Harding and B. Z. Chowdhry (Eds.), Protein-Ligand Interactions: hydrodynamics and calorimetry. Oxford University Press, Oxford New York, (2000) p 287-318.

1. Introduction

Differential scanning calorimetry (DSC) is an experimental technique to measure the heat energy uptake that takes place in a sample during controlled increase (or decrease) in temperature. At the simplest level it may be used to determine thermal transition ("melting") temperatures for samples in solution, solid, or mixed phases (e.g. suspensions). But with more sensitive apparatus and more careful experimentation it may be used to determine absolute thermodynamic data for thermally-induced transitions of various kinds. Formerly this was more the realm of the dedicated specialist, but now with the ready availability of sensitive, stable, user-friendly DSC instruments, microcalorimetry has become part of the standard repertoire of methods available to the biophysical chemist for the study of macromolecular conformation and interactions in solution at reasonable concentrations. And, to the extent that thermal transitions might be affected by ligand binding, DSC can provide useful information about protein-ligand binding. The advantages of calorimetric techniques arise because they are based on direct measurements of intrinsic thermal properties of the samples, and are usually non-invasive and require no chemical modifications or extrinsic probes. Furthermore, with careful analysis and interpretation, calorimetric experiments can directly provide fundamental thermodynamic information about the processes involved.

This document concentrates on the basic theory and practical applications of DSC in the field of protein stability and ligand interactions, with practical examples of its use and details of data analysis and pitfalls. It should be said from the outset, however, that DSC is at best only a rather indirect way of studying protein-ligand interactions, and in most cases other and more direct methods (including isothermal titration calorimetry, ITC) might be better suited to the problem. However, the technique has proved useful in some cases, and can provide preliminary information that might form the basis for more detailed studies by other techniques.

2. DSC Basics

A sketch showing the typical layout of a DSC instrument is shown in Figure 1. In a DSC experiment a

solution of protein (typically 1 mg/ml or less in modern instruments) is heated at constant rate in the calorimeter cell alongside an identical reference cell containing buffer. Differences in heat energy uptake between the sample and reference cells required to maintain equal temperature, correspond to differences in apparent heat capacity, and it is these differences in heat capacity that give direct information about the energetics of thermally-induced processes in the sample. Correct use of such instruments requires careful attention to sample preparation, buffer equilibration, and baseline controls, together with accurate measures of sample concentration if absolute thermodynamic data are required.

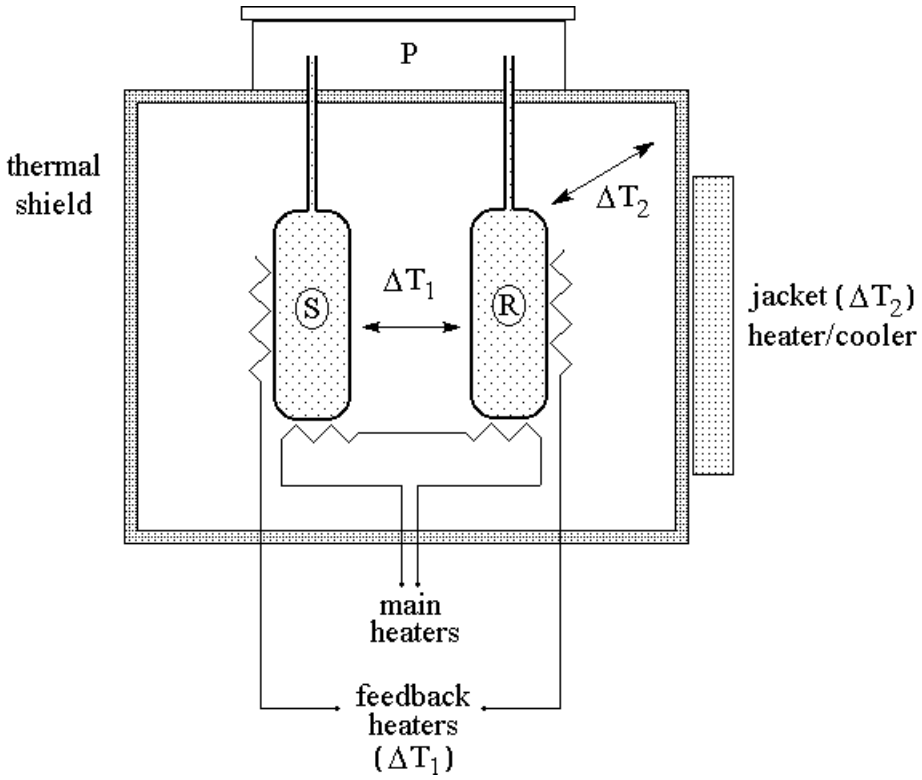

Fig. 1 Sketch diagram of a typical DSC used for thermal studies of dilute solutions of biomolecules (adapted from [2]). Identical, total-fill sample (S) and reference (R) cells (typically 0.5-2 ml volume) containing protein solution and buffer, respectively, are held under elevated atmospheric or inert gas pressure (P) to inhibit bubble formation during heating. During up-scan operation, power is supplied to the main heaters to raise the temperature of the cells at a steady rate, whilst monitoring the temperature differences between sample and reference cells (ΔT_1) and between cells and the surrounding adiabatic jacket (ΔT_2). Feedback through the jacket heater allows the thermal shield temperature to follow that of the cells, and feedback heaters on the cells compensate for any temperature differences between the cells during the scan.

2.1 Instrumentation

To avoid confusion, it must be emphasised that the differential scanning microcalorimeters described here for work on dilute biomolecular solutions are specialised instruments that differ significantly from the possibly more familiar "DSC" or "DTA" instruments commonly used in less demanding thermal analysis measurements. In particular they are designed to accommodate relatively large volumes (0.5 - 2 ml) of dilute solutions (1 mg/ml or less), in true differential mode, rather than the typically 50 µl pans used for DTA studies of solids or pastes. Currently available instruments are based primarily on pioneering work by the Privalov and Brandts groups (1-3), including the following: Microcal MC2, MCS and VP-DSC (Microcal Inc., 22 Industrial Drive E., Northampton, MA 01060-2327, USA. http://www.microcalorimetry.com); CSC NANO II DSC (Calorimetry Sciences Corp., 515 East 1860 South, Provo, Utah 84603-0799, USA. email: CalSCorp@aol.com); DASM-4 (Bureau of Biological Instrumentation, Russian Academy of Sciences, Moscow, Russia). Current versions of these instruments are comparable in sensitivity and stability, though may differ somewhat in cell configuration and control and analysis software options. Slightly less sensitive instruments, but with greater flexibility in sample handling, are available from Setaram (7 rue de l'Oratoire, F-69300 Caluire, France. http://www.setaram.fr). The experiments described here have been done using Microcal equipment, but this does not imply any particular preference.

Protocol 1. DSC of protein unfolding - basic procedure using lysozyme as a model

Equipment and reagents

- DSC instrument (Microcal MCS, VP-DSC, or equivalent)
- Dialysis tubing or cassettes (e.g. Pierce Slide-A-Lyser®)
- Degassing equipment (vacuum desiccator, magnetic stirrer)
- Buffer: 20mM Na acetate, pH 5.2 (Note: most buffers, including organic solvent mixtures, are compatible with DSC, but mercaptoethanol is best avoided because of adverse thermal effects due to oxidation and thermal degradation. Other reducing agents such as DTT or DTE are usually satisfactory, if needed.)
- Protein: Hen egg white lysozyme (e.g. Sigma L-6876), typically 2ml at a concentration of 1mg/ml (for MCS) or 1ml at 0.1 mg/ml (for VP-DSC). For more demanding work, commercial samples of lysozyme may be dialysed against ultra-pure water and lyophilised to remove extraneous salts before use.

A. Sample Preparation

1. Prepare the protein solution and dialyse several changes of appropriate buffer. Each DSC run will

typically require 1-2 ml of protein solution at a concentration of around 1 mg/ml, or less, depending on the DSC instrument.

2. Retain the final dialysis buffer for DSC reference, equilibration and dilutions.
3. Determine the protein concentration by 280nm absorbance or other appropriate method. For lysozyme the molar extinction coefficient, ε_{280} = 37900 ($A_{280}^{1mg/ml}$ = 2.65, mw 14300).
4. Immediately prior to the experiment, degas portions of the sample mixture and buffer for 2-3 minutes under gentle vacuum with gentle stirring. Be careful to avoid excessive degassing or frothing of the mixture at this stage.

B. DSC procedure

1. Load DSC sample and reference cells with degassed buffer and collect baseline scan(s) using appropriate temperature range and scan rate (typically 20 – 100 °C, 60 °C/hr).
2. Allow the DSC cells to cool and refill the sample cell with protein solution.
3. Repeat the DSC scan(s) using the same parameters as in 1. (Depending on circumstances, it may be useful to do repeat scans with the same sample to establish reversibility and reproducibility. It can also be useful to run a preliminary scan, stopping some way before the unfolding transition begins, before cooling and performing the complete scan. This minimises baseline artefacts that can be induced by the thermal shock involved in loading the sample or reference cells.)
4. After final cooling, remove the sample an examine for turbidity, aggregation or other visible changes. (Precaution: traces of aggregated protein or other contaminants in the DSC cell will cause erratic baseline behaviour. Routine vigorous cleaning of the DSC cells, using detergents or strong acids/bases as recommended by the manufacturer, is essential for reliable DSC operation, especially when working with readily aggregating systems.)
5. Process data using instrumental software. This normally involves subtraction of buffer baseline (from 1), concentration normalisation, followed by deconvolution of the resultant thermogram using an appropriate model.

2.2 Quantitative Analysis of DSC Data - Practical Considerations

The typical experimental procedure for following the thermal unfolding of a simple globular protein is described in Protocol 1. Representative data from such an experiment are shown in Figure 2. In this section we shall outline some of the practical aspects related to analysis and interpretation of such data, leaving the theoretical background and justification for some of the points to be described in later sections.

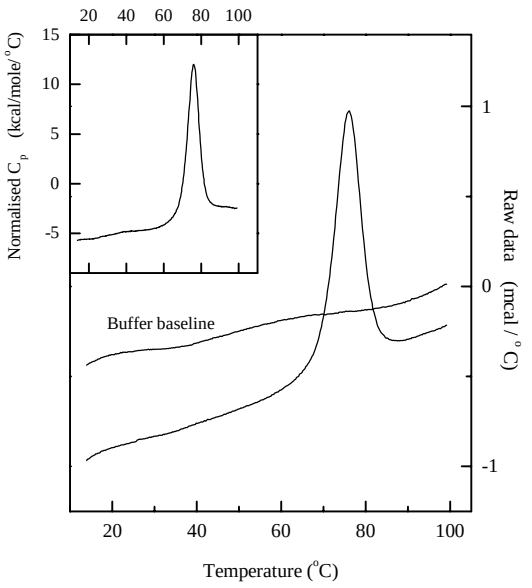

Fig. 2 Raw DSC data for thermal unfolding of hen egg white lysozyme (1.2 mg/ml) in 20mM Na acetate buffer, pH 5.2, at a scan rate of 60 °C hr^{-1}. These data were obtained following Protocol 1, using a Microcal-MCS system. (Similar data are obtained with 10-fold lower concentrations using the more recent VP-DSC instrument.) The inset shows the same data after subtraction of the instrumental (buffer) baseline and concentration normalisation, illustrating also the pre- and post-transition baseline behaviour typical of such processes.

The output from any DSC experiment is a thermogram showing the excess heat capacity (C_p, sample minus reference) as a function of temperature. For a simple globular protein the thermogram comprises three regions: the pre-transition baseline, the endothermic unfolding transition, and the post-transition baseline. At temperatures well below the onset of thermal unfolding, the C_p simply reflects the difference in heat capacity between the protein and the solvent (usually mainly water) it has displaced. Since water has a high heat capacity compared to most organic substances, including proteins, the apparent C_p in this region will normally be negative. For most proteins, this pre-transition baseline also shows a slight positive slope, indicating a gradual increase in heat capacity with temperature - a characteristic also of organic solids. As the protein begins to unfold, the C_p increases as more heat energy is taken up in denaturing the protein, reaching a peak at approximately the mid-point (T_m) temperature of the process (assuming a single cooperative unfolding process), before dropping down to the high temperature baseline. This post-transition baseline, representing the relative heat capacity of the unfolded polypeptide, is usually found at a higher level (positive ΔC_p) and has a lesser slope than the pre-transition baseline. Similar effects are seen for organic liquids also. Consequently, as a first approximation, one might picture the unfolding of a globular protein in water as the “melting” of an organic microcrystal suspended in an aqueous environment.

The Calorimetric Enthalpy (ΔH_{cal}) is the total integrated area under the thermogram peak which, after appropriate baseline correction, represents the total heat energy uptake by the sample undergoing the transition. This heat uptake will depend on the amount of sample present in the active volume of the DSC cell and is, in principle at least, a model-free absolute measure of the absolute enthalpy of the process involved. [It is axiomatic that equilibrium transitions observed in a DSC upscan experiment must be endothermic. Exothermic heat effects can be observed, but when these are encountered it is usually an indication of thermodynamically irreversible, non-equilibrium processes, kinetically activated by elevated temperatures. Aggregation of thermally denatured protein is one such example.]

The van't Hoff Enthalpy (ΔH_{VH}) is an independent estimate of the enthalpy of the transition, based on an assumed model for the process. Here one simply uses the area under the C_p peak at any temperature (see Fig.3), divided by the total area, as a measure of the fraction or extent of unfolding that has occurred at that temperature. In this way one is using the calorimetric signal in just the same way as any other indirect method for following the unfolding transition, such as CD or fluorescence, for example. Assuming a simple two-state model, one can then relate the temperature variation of the fraction unfolded to the apparent enthalpy of the process using the van't Hoff equation (see below). The advantage of this approach is that, since it relies only on ratios of areas under the experimental curve, it does not require any information about concentration or purity of the sample. Ideally, ΔH_{cal} and ΔH_{VH} should be identical in any calorimetric experiment, and comparison of the two can be quite revealing about factors such as the purity and concentration of the sample, and can also give information about the reversibility and apparent mechanism of the process.

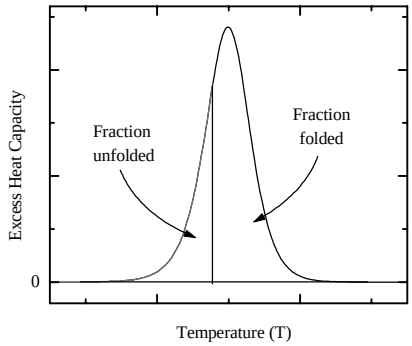

Fig. 3 Sketch showing the use of DSC thermograms to determine the van't Hoff enthalpy (ΔH_{VH}) of a 2-state transition. (This is the peak as it might appear ideally after baseline correction.) At any particular temperature, the extent of unfolding is represented by the area under the thermogram up to that point (shaded area). Consequently, the ratio of shaded:unshaded areas corresponds to K (= [U]/[N]) at that temperature, which can be used in the van't Hoff equation to determine ΔH. Note that, since only ratios of areas are used in this calculation, neither the absolute units of C_p nor the protein concentrations are required. The method is, however, dependent on the validity of the 2-state (or other) model adopted.

2.3 Concentration Measurements

The accuracy of any calorimetric ΔH_{cal} (as opposed to ΔH_{VH}) estimate is critically dependent upon the purity of the sample and on the reliability of the methods used to determine its concentration. For proteins, the most convenient and straightforward method for concentration measurement is usually the UV (280nm) absorbance, provided a reliable molar extinction coefficient (ε_{280}) is available. This is non-destructive and can frequently be done on the actual sample solution prior to insertion in the DSC. ε_{280}may usually be estimated to reasonable precision (typically ± 5%) from the aromatic amino acid (Trp, Tyr) composition of the protein (4). It goes without saying that such measurements should follow good working practices using reliable instrumentation and clean cuvettes, since the entire DSC analysis may depend on this one measurement. In our experience it is unwise to rely on a simple A280 measurement at fixed wavelength, but better to record a complete UV spectrum (240-400 nm), since this can show up immediately any problems due to incorrect baselines, light scattering by aggregated protein, or other impurities. Colorimetric methods of protein estimation (e.g. “Bradford” or other dye-binding assays) are generally less reliable unless previously calibrated for the specific protein under investigation.

One must remember, of course, that most methods of protein estimation will also measure contributions arising from misfolded protein and other protein impurities. For example, if some of the protein sample is already misfolded or unfolded prior to the DSC experiment, and does not contribute to the unfolding transition, then the calorimetric enthalpy for that transition will be reduced accordingly, even though one might be unaware of this problem from simple concentration measurements. Interestingly, the van’t Hoff enthalpy is not affected by such impurities, provided they don’t interfere with the cooperative transition of the correctly folded fraction.

2.4 Units

The SI unit for energy is the joule (J). Consequently, the conventional units are kJ mol^{-1} for molar thermodynamic energies such as enthalpy (H) or free energy (G) and J mol^{-1} K^{-1} for molar entropy (S) or heat capacity (C_p). Despite this, many (particularly in the US) still use the older system of units based on the calorie, and some instruments (e.g. Microcal) are still calibrated in such units. For conversion: 1 calorie = 4.184 J ; the gas constant, R = 1.987 cal mol^{-1} K^{-1} = 8.314 J mol^{-1} K^{-1} .

2.5 Scan rates/reversibility

A scan rate of 60 °C/hr is usually adequate for simple, reversible unfolding transitions, and in theory the DSC thermogram should be unaffected by use of different scan rates. However, there are many instances of kinetically-determined irreversible process (such as aggregation or chemical degradation at

higher temperatures) that can affect the shape of the thermogram, and which are scan rate-dependent. It is always prudent to repeat experiments at different scan rates to determine whether this is a problem in particular instances. Analysis of DSC data in such cases is beyond the scope of the current chapter, but details may be found in (5,6).

2.6 The Baseline Problem

Reliable interpolation of the baseline is crucial to the estimation of both calorimetric and van't Hoff enthalpies of a DSC transition, since both the area under the thermogram and its shape will be affected by this. There are two separate aspects to this that one might refer to as the "instrumental" baseline and the "sample" baseline problems, respectively.

The instrumental baseline is quite straightforward and is just the measured DSC response one would get in the absence of sample. This is typically obtained from scans under identical conditions using sample buffer or appropriate solvent in both cells of the DSC. Since instrumental baselines are susceptible to long-term drift and can vary with ambient conditions, such measurements are best made on a well-equilibrated instrument, both before and after the experimental scans. For samples involving irreversible transitions, some workers prefer to use a second sample scan as baseline. "Annealing" of the sample prior to the transition can also be useful. Since the greatest variation in instrumental baseline usually occurs in the first scan after reloading the DSC cell, due to the relatively large thermal disturbance that this involves, heating the sample in the DSC one or more times to a temperature below the onset of the transition and cooling, prior to execution of the full scan ("annealing"), can give more reliable baseline stability and reproducibility.

Estimation of the sample baseline is a thornier problem. Since the heat capacity baseline of a sample rarely returns to the same level after the transition as it was before, because of ΔC_p effects, one needs to be able to estimate what the sample baseline might have been in the region under the endotherm peak in the absence of the transition. A typical DSC endotherm for a simple globular protein undergoing a cooperative two-state unfolding transition is illustrated in Figure 2. At any point under the transition endotherm, the sample comprises a mixture of folded and unfolded proteins, and the problem is how best to estimate what the heat capacity of this mixture should be. Various strategies have been adopted and are illustrated in Figure 4.

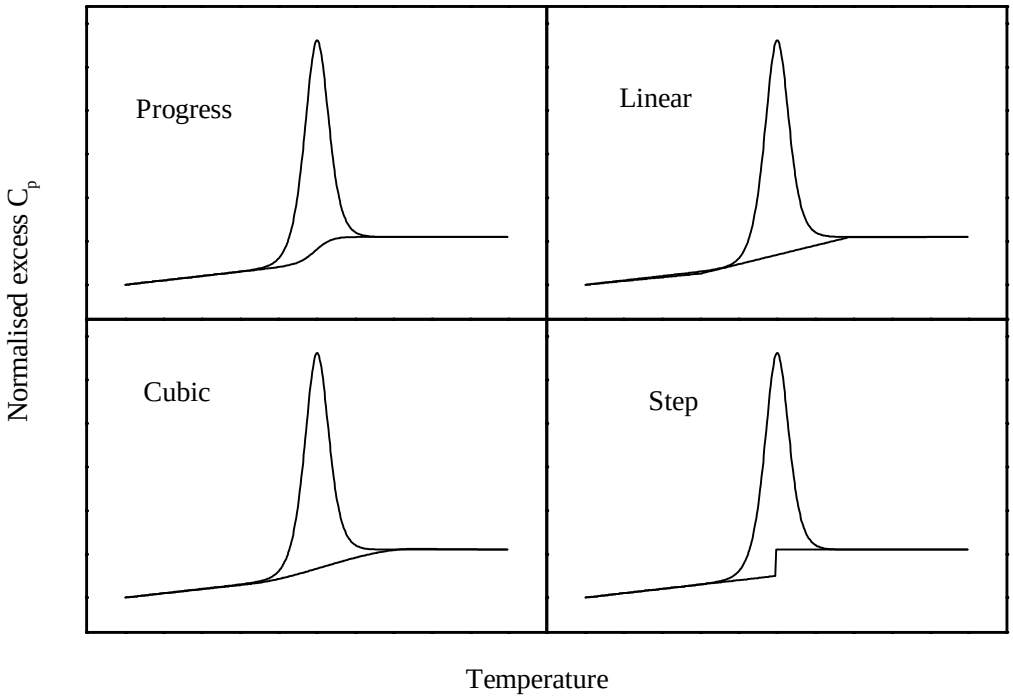

Fig. 4 Examples of the different baseline assumptions that may be used in analysis of DSC transitions. The data here correspond to a typical 2-state cooperative unfolding transition, calculated for a 25 kDa protein with T_m of 50 °C , ΔH_{cal} = ΔH_{VH} = 100 kcal mol^{-1} , ΔC_p (at T_m) = 1.5 kcal mol^{-1} K^{-1}, with a pre-transition slope of 0.025 kcal mol^{-1} K^{-2} and a post-transition slope of zero. Note that this means that the ΔC_p is temperature dependent in this case.

The "progress baseline" is obtained by extrapolation of pre- and post-transition baselines and, at any particular temperature, calculating the baseline heat capacity in proportion to the estimated amounts of folded and unfolded material present at that temperature from the area under the curve. Simpler approaches use a step baseline at the mid-point of the transition - either at the peak of the thermogram or at a halfway point in terms of area under the curve. Alternatively one may choose linear or quadratic interpolations of pre- and post-transition baselines. The next step in any DSC analysis is usually to subtract the chosen baseline prior to area integration or fitting of the unfolding endotherm, and in practice there seems little to choose between the various methods of baseline correction: each of them produces inevitable distortion in the corrected thermogram that can affect both the apparent shape and area under the transition that translate into possible errors in the ΔH_{cal} and ΔH_{VH} estimates (see Table 1 for example). However, for good data these differences are relatively small, and usually smaller than errors arising from concentration estimates or instrumental baseline drifts.

Interestingly, and paradoxically, even in cases where the chosen baseline correction is patently wrong, an empirical relationship has been discovered that combines the apparent computed ΔH_{cal} and ΔH_{VH} to

give and enthalpy close to reality (but only for transitions that are truly two-state and for which experimental data, including protein concentrations, are otherwise correct). This comes about because errors in choice of baseline correction tend to have opposing effects on the calorimetric and van't Hoff enthalpies. Any baseline error that reduces the apparent area under the peak, thus lowering ΔH_{cal}, will also tend to sharpen the peak, raising the estimate of ΔH_{VH} . Conversely, any baseline correction that broadens the transition endotherm, giving a reduced estimate of ΔH_{VH} , will also increase the area under the curve to give a higher apparent ΔH_{cal}. This is illustrated in Figure 5 and Table 1 using ideal calculated data for a representative two-state transition.

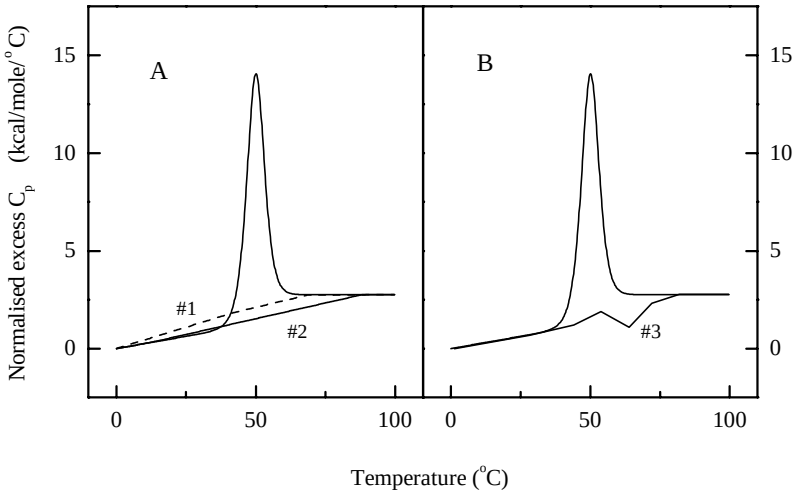

Fig. 5 Some (exaggerated) examples of incorrect baselines which, nevertheless, can give reasonable estimates of ΔH when using the empirical equation for ΔH_{WA} (see Table 1). The data are calculated as in Figure 4. (A) Alternative, incorrect linear baselines. (B) A baseline "glitch", such as might arise experimentally from electronic interference, or from the formation of a small air bubble in the DSC cell, or from convection artefacts due to particulate matter in the sample or reference solution.

Table 1 shows how, even for perfect data, the choice of baseline can lead to distortions that affect subsequent estimates of calorimetric and van't Hoff enthalpies, usually in opposite directions. The differences are small and usually experimentally insignificant between the various conventional progress, step, or interpolation baselines. However, the differences in apparent ΔH_{cal} and ΔH_{VH} are much larger when seriously distorted baselines are involved - the sort of thing that can arise experimentally from baseline fluctuations due to particulate matter in the sample or other instrumental "glitches". Empirically we have found that, even in such pathological cases, by combining the enthalpy estimates using the following formula:

$$\Delta H_{WA} = 0.65 \times \Delta H_{VH} + 0.35 \times \Delta H_{cal}$$

one obtains a weighted average estimate (ΔH_{WA}) remarkably close to the true value. This relationship was first obtained by Haynie (7) as a means of correcting data in special cases, but we have subsequently shown that the formula is more generally applicable to any kind of baseline uncertainty. The reason for the success of this relationship is not fully clear, but probably stems from the mathematical properties of curves of this kind, and the inverse correlation between peak area and peak width during baseline interpolations. Despite this, one must be very cautious in using any such empirical relationship as a substitute for good experimental technique. Differences in ΔH_{cal} and ΔH_{VH} often arise for reasons other than poor baseline correction, and application of this empirical formula would be inappropriate in such cases.

Table 1: Effects of different baseline assumptions on deconvoluted DSC data.

	Baseline	**T_m**	**ΔH_{cal}**	**ΔH_{VH}**	**ΔH_{WA}**
Figure 4:	Progress	49.9	99.3	100	99.8
	Linear	50.1	104	97.5	99.8
	Cubic	50.1	105	96.9	99.7
	Step	49.7	99.5	99.7	99.6
Figure 5:	#1	50.2	109	94.3	99.4
	#2	50.1	96.8	103	100.8
	#3	50.2	110	91.3	97.8
	None	50.3	143	73.8	98.0
	(True)	50	100	100	100.0

Hypothetical data were calculated for a two-state transition with T_m = 50 °C, $\Delta H_{cal} = \Delta H_{VH}$ = 100 kcal mol^{-1} and ΔC_p = 1.5 kcal K^{-1} mol^{-1} at T_m, with pre- and post-transition baseline slopes of 0.025 and 0 kcal K^{-2} mol^{-1}, respectively. The various baselines were then subtracted, and data fitted using Microcal ORIGIN software to a two-state model that independently estimates apparent ΔH_{cal} and ΔH_{VH} values. These were then combined using the empirical equation (see text, Section 2.6) to give the weighted average ΔH_{WA}.

3. Thermodynamic Background

For many applications it is not necessary to have a full understanding of the theoretical thermodynamic background - and it is perfectly possible to use the calorimeter as a convenient qualitative analytical instrument, just as one might use many other devices, without regard to theory. However, in order to fully appreciate the quantitative limitations on experimental observations and their thermodynamic interpretation, it is preferable to have an understanding of at least some of the basics. This is particularly important if one wishes to avoid some of the more common pitfalls in the (over)interpretation of thermodynamic data.

3.1 Heat Capacity, Enthalpy and Entropy

Differential scanning calorimetry of the kind used here measures the excess heat capacity of the sample solution with respect to the reference (usually aqueous) solvent. The heat capacity (or specific heat) of any substance (usually designated C_p at constant pressure) reflects the ability of the substance to absorb heat energy without increase in temperature, and this is central to DSC measurements and to the fundamental underlying thermodynamics. Liquid water has a relatively high C_p because of the extensive ice-like hydrogen-bonded network in the liquid that allows heat energy to be used up in breaking bonds between water molecules rather than increasing their kinetic energy (i.e. temperature). Organic matter, including proteins and nucleic acids, has a lower specific heat than water, except possibly when undergoing some process such as unfolding or melting involving breaking of bonds. Consequently the heat capacity of a dilute biomolecule solution is dominated by the water in the system and great care has to be taken to subtract this in any DSC measurements to give the excess differential heat capacity contribution arising from the process of interest.

Heat capacity is the fundamental property from which all thermodynamic quantities may be derived. In particular, the absolute enthalpies (H) and entropies (S) of any substance are related to the total heat energy uptake involved in the (imaginary) process of heating from absolute zero to temperature T, as represented in the following integrals:

$$H = \int_0^T C_p.dT + H_0$$

where H_0 is the ground state energy (at 0 K) due to chemical bonding and other non-thermal effects, and since classically from the 2nd law of Thermodynamics

$$dS = dH/T = (C_p/T).dT$$

it follows that:

$$S = \int_0^T (C_p/T).dT$$

The molecular interpretation of H (enthalpy, or heat content) is fairly easy to grasp since it is just the total energy (including pressure/volume work terms) taken up in raising the system to temperature T whilst keeping the pressure constant. This will include the energy associated with all the atomic and molecular motions - translation, rotation, vibration, etc. - together with energy taken up in changes in inter- and intra-molecular interactions ("bonds"). By contrast, the absolute entropy (S) is a rather more difficult quantity to comprehend. Usually it is described in terms of "molecular disorder" - the higher the disorder the higher the entropy - but this obscures the connection with heat capacity evident in the above integral definition. Perhaps a better way of viewing entropy is as the multiplicity of ways in which the molecules in a system can take up energy without increasing temperature.

The magnitude of the heat capacity depends on the numbers of ways there are of distributing any added heat energy to the system, and so is related to entropy. Consider the energy required to bring about a 1 degree rise in temperature. If a particular system has only relatively few ways of distributing the added energy, then relatively little energy will be required to raise the temperature, and such a system would have relatively low C_p. If, however, there are lots of different ways in which the added energy can be spread around amongst the molecules in the system (such as different modes of vibration and rotation, or breaking of bonds), then much more energy will be needed to bring about the same temperature increment. Such a system would have a high C_p. In this way, adding heat to anything increases the entropy by giving the molecules more energy to explore many more different ways of arranging themselves (and become "more disordered").

3.2 Equilibrium and Free Energy

Chemical stability and thermodynamic equilibrium represent a balance between two opposing tendencies: firstly the natural trend for systems to move to lower energies (decrease H), and secondly the equally natural tendency at the molecular level for molecules to explore the multiplicity of states available (higher S) under the influence of disruptive thermal motions. This is represented by the Gibbs Free Energy change (ΔG) expression:

$$\Delta G = \Delta H - T.\Delta S$$

which tells us how much work must be done to bring about the desired change. (Changes can occur spontaneously if ΔG is negative, but require the input of energy if positive. Systems are in equilibrium if $\Delta G = 0$.)

Free energies and other thermodynamic parameters are relative quantities that depend on an arbitrary choice of reference or standard state. It turns out that the equilibrium constant (K) for any process is related to the "standard" Gibbs free energy change:

$$\Delta G^\circ = -RT.\ln(K) = \Delta H^\circ - T.\Delta S^\circ$$

(R = gas constant, 8.314 J K^{-1} mol^{-1} or 1.987 cal K^{-1} mol^{-1})

representing the free energy change ΔG°, together with the constituent enthalpy ΔH° and entropy ΔS° changes, that would take place in the (hypothetical) standard state in which reactants and products (initial and final states) were all present at 1 molar concentration (or activity). (This convention adopting 1M concentration for standard states in solution, clearly unrealistic for biomolecular systems, is a consequence of an historical choice of standard units for measuring concentration, but remains an appropriate way of comparing interaction free energies and other parameters on the same scale.) A convenient way to view ΔG° is simply as the equilibrium constant, K, expressed on a logarithmic energy scale. ΔH and ΔH° are practically identical under most conditions, but ΔS and ΔS° will normally differ significantly due to large entropy of mixing effects at different concentrations.

3.3 Temperature Dependence of Thermodynamic Quantities

Changes in enthalpy and entropy (ΔH and ΔS) as a system changes from one state to another ($A \rightarrow B$) at constant temperature follow directly from the integral definitions:

$$\Delta H = H_B - H_A = \int_0^T \Delta C_p \, dT + \Delta H(0)$$

$$\Delta S = S_B - S_A = \int_0^T (\Delta C_p / T) \, dT$$

where $\Delta C_p = C_{p,B} - C_{p,A}$ is the heat capacity <u>difference</u> between states A and B at a given temperature. $\Delta H(0)$ is the ground state enthalpy difference between A and B at absolute zero. Most systems are assumed to have the same (zero) entropy at 0 K (3rd. Law of Thermodynamics).

It is both conventional and convenient to relate these quantities to some standard reference temperature T_{ref} (e.g. T_{ref} = 25 °C or 298 K, rather than absolute 0 K), in which case:

$$\Delta H(T) = \Delta H(T_{ref}) + \int_{Tref}^{T} \Delta C_p \, .dT$$

and $$\Delta S(T) = \Delta S(T_{ref}) + \int_{Tref}^{T} (\Delta C_p / T).dT$$

This illustrates how, if there is a finite ΔC_p between two states (as is the norm, for example, in protein unfolding or other processes involving multiple, weak, non-covalent interactions), then ΔH and ΔS are both temperature dependent.

If ΔC_p is constant, and does not vary with temperature (not altogether true for protein transitions, but usually a reasonable approximation over a limited temperature range), then we can integrate the above to give approximate expressions for the temperature dependence of ΔH and ΔS with respect to some arbitrary reference temperature (T_{ref}):

$$\Delta H(T) \cong \Delta H(T_{ref}) + \Delta C_p \, .(T - T_{ref})$$

$$\Delta S(T) \cong \Delta S(T_{ref}) + \Delta C_p \, .\ln(T/T_{ref})$$

Interestingly (16) these temperature effects will largely cancel in the standard free energy expression to give a ΔG that is relatively much less affected by temperature change. This is an example of "entropy-enthalpy compensation" or "linear free energy" effects that are often, if not universally, found in systems involving multiple weak interactions (8,9,22).

For protein unfolding it is sometimes convenient to take the mid-point (T_m) of the transition as reference temperature. Here, T_m is defined as the temperature at which ΔG for the transition is zero, so that $\Delta H(T_m) = T_m.\Delta S(T_m)$ and consequently the temperature dependence of the unfolding free energy may be written:

$$\Delta G_{unf}(T) = \Delta H_{unf}(T_m).\{1 - T/T_m\} + \Delta C_p.\{T - T_m - T.\ln(T/T_m)\}$$

assuming ΔC_p does not itself change with temperature. Typical data showing this temperature variation for thermodynamic parameters of unfolding of a globular protein are shown in Figure 6. This illustrates how quite large changes in ΔH and ΔS tend to compensate to give relatively much smaller changes in ΔG. This also illustrates how the T_m for the transition, given by the point of intersection of the ΔH and $T.\Delta S$ lines, can be quite sensitive to relatively small changes in either ΔH or $T.\Delta S$, even though they may make relatively little change to ΔG. It is partly for this reason that prediction of the effects brought

about by environmental changes or mutagenesis on protein stability is so difficult.

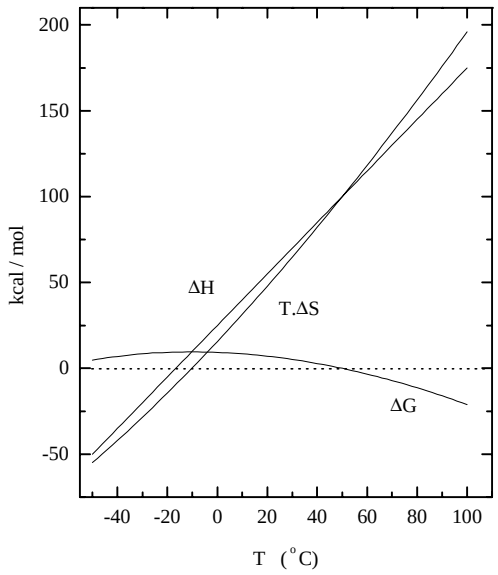

Fig. 6 Temperature variation of the thermodynamic parameters for unfolding of a typical globular protein, calculated for a 25 kDa protein undergoing a 2-state transition with T_m of 50 °C , $\Delta H_{cal} = \Delta H_{VH} = 100$ kcal mol^{-1} , with a temperature-independent $\Delta C_p = 1.5$ kcal mol^{-1} K^{-1} . The mid-point of the transition (T_m) is given by the point of intersection of the ΔH and $T.\Delta S$ lines, where ΔG crosses the zero axis.

3.4 The van't Hoff Enthalpy

Although enthalpy changes (ΔH) can be measured directly by calorimetric techniques, it is frequently convenient to compare these with indirect estimates using the classical van't Hoff equation, arising from the temperature dependence of the equilibrium constant (K) as one might get, for example, from following the thermal unfolding of a protein by fluorescence, circular dichroism, or other indirect technique. Such information can also be obtained from DSC experiments (see below).

Given that: $$-RT.\ln K = \Delta H^\circ - T.\Delta S^\circ$$

it follows that $$\ln K = -\Delta H^\circ/RT + \Delta S^\circ/R$$

and, in the general case where ΔH° and ΔS° may also vary with temperature:

$$d(\ln K)/d(1/T) = -\Delta H^\circ/R - (1/RT)[d(\Delta H^\circ)/(1/T)] + (1/R)[d(\Delta S^\circ)/d(1/T)]$$
$$= -\Delta H^\circ/R$$

since: $$d(\Delta H^\circ)/d(1/T) = -T^2.d(\Delta H^\circ)/dT = -T^2.\Delta C_p$$

and: $$d(\Delta S°)/d(1/T) = -T^2.d(\Delta S°)/dT = -T^2.\Delta C_p/T$$

leads to cancellation of the latter two terms in the above equation.

Thus, a plot of experimental data of lnK vs. 1/T ("van't Hoff plot") gives a line whose slope at any point is the van't Hoff enthalpy ($\Delta H°$ or ΔH_{VH}) divided by R. In simple cases, usually over a limited temperature range, this plot is linear (or is assumed to be so, within experimental error), but in general the temperature dependence of ΔH (due to ΔC_p) will give curvature of the van't Hoff plot that needs more careful analysis (10).

One must be clear about what is meant by the "van't Hoff Enthalpy" and to what it refers. Any van't Hoff analysis is based on a model or hypothesis of the process involved, needed in order to define K. Typically, for protein unfolding transitions, this model will be a "2-state" picture in which the equilibrium constant K is a dimensionless ratio determined, usually indirectly, from spectroscopic, calorimetric, or other measurements. In such a model the molar van't Hoff enthalpy change, ΔH_{VH} , is the enthalpy change per mole of cooperative unit as defined by the model (3,11). Comparison of this with the directly measured calorimetric value can frequently be informative. There are two simple ways in which the van't Hoff enthalpy might differ from the calorimetric enthalpy in protein unfolding. Firstly, the model may simply be wrong. For example, if the unfolding transition is not two-state, but involves one or more intermediate steps, then the transition will appear broader than anticipated, and the ΔH_{VH} will be less than ΔH_{cal}. Alternatively, if the protein unfolds cooperatively as a dimer, or higher oligomer, then the transition will be sharper than anticipated for the two-state transition of a monomer, and the ΔH_{VH} will be correspondingly greater than ΔH_{cal}. One must be wary, however, of placing too much reliance on such comparisons in the absence of supporting evidence from other methods, since other factors can affect the shape of the transition. For example, the irreversible and usually exothermic aggregation of thermally unfolded protein can distort and sharpen the DSC transition, leading to incorrect estimates of both ΔH_{cal} and ΔH_{VH} . Moreover, as indicated in an earlier section, any errors arising out of impurities or incorrect concentration estimates will be reflected in ΔH_{cal} (though not necessarily in ΔH_{VH}) and will give rise to erroneous ΔH_{cal}:ΔH_{VH} ratios.

4. Effects of Ligand Binding

Application of Le Chatelier's principle, or simple equilibrium considerations, shows that if any ligand (small molecule or other protein or macromolecule) binds preferentially to the folded or native form of the protein, then this will stabilise the folded state, and unfolding of the protein will become progressively less favourable as ligand concentration increases. Conversely, ligands that bind preferentially to the unfolded protein will destabilise the fold and will encourage unfolding. Examples of both are commonly seen (12-15).

4.1 Ligand Binding and Folding Equilibrium

4.1.1 Ligand binds to the folded protein

General theories for binding of multiple ligands and multiple protein may be found in refs.(12,13,16). For the simplest case in which a ligand molecule (L) binds specifically to a single site on the native folded protein (N), the following equilibria apply:

Ligand binding: $N + L \rightleftharpoons NL$; $K_{L,N} = [N][L]/[NL]$

Unfolding: $N \rightleftharpoons U$; $K_0 = [U]/[N]$

where $K_{L,N}$ defines the dissociation constant for ligand binding to the native protein and K_0 is the unfolding equilibrium constant for the protein in the absence of ligand.

In general the effective unfolding equilibrium constant (K_{unf}) is given by the ratio of the total concentrations of unfolded to folded species:

$$K_{unf} = [U]/([N] + [NL]) = K_0/(1 + [L]/K_{L,N}) \approx K_0 K_{L,N}/[L]$$

where the final approximate form applies only at high free ligand concentrations ($[L] > K_{L,N}$). This confirms the expectation that K_{unf} decreases and the folded form becomes more stable with increasing ligand concentration.

This can be expressed alternatively in free energy terms:

$$\Delta G_{unf} = -RT.\ln(K_{unf}) = \Delta G_{unf,0} + RT.\ln(1 + [L]/K_{L,N})$$

$$\approx \Delta G_{unf,0} + \Delta G^\circ_{diss,N} + RT.\ln[L] \qquad ([L] >> K_{L,N})$$

where $\Delta G_{unf,0}$ is the unfolding free energy of the unliganded protein, and $\Delta G^\circ_{diss,N} = -RT.\ln(K_{L,N})$ is the standard Gibbs free energy for dissociation of the ligand from its binding site on the native protein. The approximate form again applies only at high ligand concentrations.

This illustrates how the stabilising effect of bound ligand can be visualised as arising from the additional free energy required to remove the ligand prior to unfolding, together with an additional contribution (RT.ln[L]) from the entropy of mixing of the ligand when released into the bulk solvent.

In the approximate form at high ligand concentrations the free energy can be broken down into the separate enthalpy and entropy contributions:

$$\Delta H_{unf} \approx \Delta H_{unf,0} + \Delta H^{\circ}_{diss,N}$$

and $$\Delta S_{unf} \approx \Delta S_{unf,0} + \Delta S^{\circ}_{diss,N} - R.\ln[L]$$

In many cases the heat of ligand dissociation ($\Delta H^{\circ}_{diss,N}$) might be quite small compared to the heat of unfolding of the protein, especially in the case of small ligands, and can be hard to distinguish in calorimetric unfolding experiments. This can be further complicated when ΔH_{unf} is also varying with temperature due to ΔC_p effects. Entropy effects, particularly those arising from the ligand mixing term (R.ln[L]), will dominate in such cases. [Similar considerations apply at lower concentrations of ligand, though the algebraic expressions are a little more complicated. In such cases the thermodynamic parameters are intermediate between unliganded and fully-liganded values found in the high concentration limit.]

A typical example of the effects of small ligand binding to a globular protein is described in Protocol 2 for the case of 2'-CMP binding to ribonuclease A, with typical DSC data shown in Figure 7.

Protocol 2. Ligand binding to folded protein - DSC of RNAse with 2'-CMP

Equipment and reagents

- DSC and related equipment, as in Protocol 1
- Buffer: 0.1 M Na acetate, pH 4.5
- Protein: Ribonuclease A (RNAse ; Sigma R-5500; mw 13,700 , ε_{280} = 10550)
- Ligand: Cytidine 2'-monophosphate (2'-CMP ; Sigma C-7137)

Experimental procedure

1. Prepare a stock solution of RNAse in buffer (sufficient for several experiments) and determine the concentration by UV absorbance. Take portions of this stock solution and add sufficient 2'-CMP (by weight) to give a series of samples containing 0-1.5 mM ligand. Note: (a) It is not usually necessary to dialyse the protein solution in this case, since commercial samples of RNAse are usually sufficiently pure for these purposes; (b) Since 2'-CMP also absorbs in the 260-280 nm

region, it is necessary to determine the RNAse concentration prior to adding the ligand; (c) For detailed work it will also be necessary to prepare identical concentrations of CMP in buffer alone, to use as reference.

2. For each of the RNAse/2'-CMP mixtures separately, degas and run DSC experiments as described in Protocol 1.
3. Analyse the data using a simple 2-state model to determine T_m , ΔH_{cal} , and ΔH_{VH}

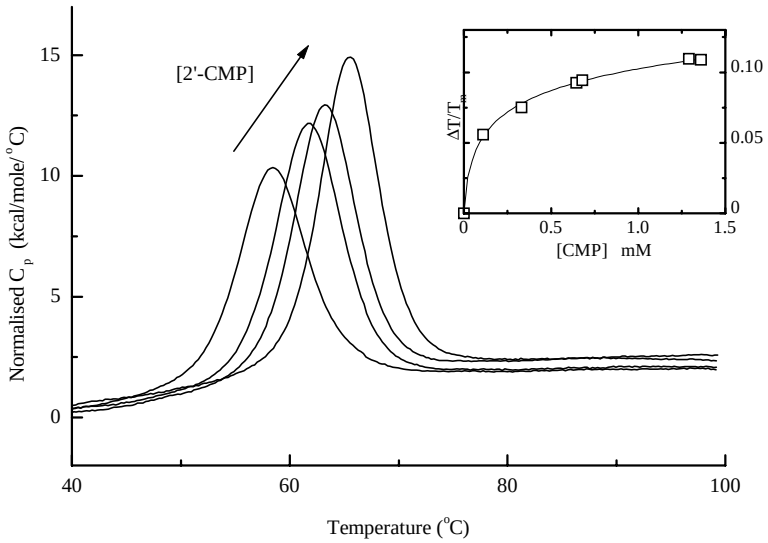

Fig. 7 DSC thermograms of ribonuclease-A unfolding in the presence of various concentrations of 2'-CMP, illustrating the increase in T_m brought about by binding of ligand to the native protein. The inset shows the relative T_m shift as a function of total ligand concentration, with the non-linear regression fit to Equation ?? (Experimental conditions: 0.1M Na acetate buffer, pH 4.5, 80 μM ribonuclease, 0-1.5 mM CMP).

4.1.2 Ligand binds to unfolded protein

The same approach can be applied to situations where ligand binds only to the *unfolded* protein (14,15):

$$U + L \rightleftharpoons UL \qquad ; \qquad K_{L,U} = [U][L]/[UL]$$

for which:

$$K_{unf} = ([U] + [UL])/[N] = K_0.(1 + [L]/K_{L,U}) \approx K_0.[L]/K_{L,U}$$

and: $$\Delta G_{unf} = -RT.\ln(K_{unf}) = \Delta G_{unf,0} - RT.\ln(1 + [L]/K_{L,U})$$

$$\approx \Delta G_{unf,0} - \Delta G^\circ_{diss,U} - RT.\ln[L] \qquad \text{(for high [L])}$$

This illustrates the destabilising effect of a reduction in unfolding free energy as ligand binds to the unfolded polypeptide. Equivalent expressions for the enthalpy and entropy contributions may be written as above, with appropriate sign changes.

An example of this kind of effect is illustrated in Fig. 8 for the unfolding of globular proteins in the presence of cyclodextrins, with practical details given in Protocol 3. The cyclodextrins are a family of toroidal oligosaccharide molecules that form inclusion complexes with small non-polar molecules and therefore bind to exposed aromatic groups on the unfolded protein (14). This gives rise to a decrease in thermal stability of the protein with increasing cyclodextrin concentration that can be analysed in terms of the simple models described here.

Protocol 3. Ligand binding to unfolded protein - DSC of lysozyme with cyclodextrin

Equipment and reagents

- DSC and related equipment, as in Protocol 1
- Buffer: 40mM glycine/HCl, pH 3.0
- Protein: Hen egg white lysozyme (e.g. Sigma L-6876), typically 2ml at a concentration of 1mg/ml (for MCS) or 1ml at 0.1 mg/ml (for VP-DSC).
- Ligand: α-cyclodextrin (Sigma-Aldrich supply a range of cyclodextrins, including some of the more soluble methyl- and hydroxypropyl- derivatives of β-cyclodextrin which may also be used for this experiment.) Cyclodextrins are quite hygroscopic. They should be vacuum dried before use and stored desiccated.

Experimental procedure

1. Prepare 5-15% (w/v) solutions of cyclodextrin in buffer and use these to make up appropriate solutions of lysozyme, retaining sufficient cyclodextrin buffer solution for use as reference.

2. Proceed with DSC experiments, as in Protocols 1 and 2. Cyclodextrins do not normally absorb in the near UV region, so protein concentrations may be determined directly from A280, as before.
3. Analyse the data using a simple 2-state model to determine T_m , ΔH_{cal} , and ΔH_{VH}

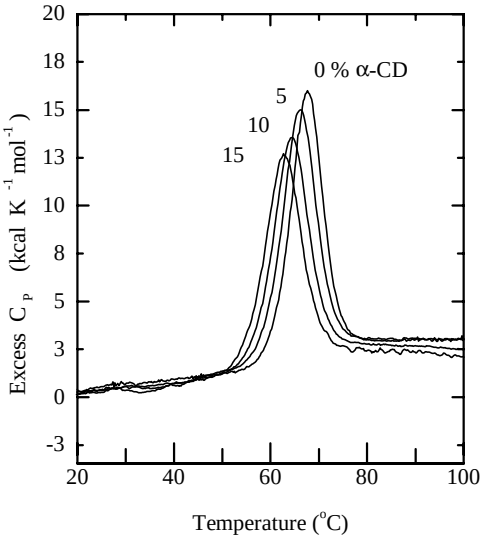

Fig. 8 Normalized DSC thermograms of lysozyme (40 mM glycine/HCl, pH 3.0) in the presence of 0, 5, 10, 15% (w/v) of α-cyclodextrin.

Note: the apparent variation in ΔH_{cal} is predominantly due to the inherent variation of unfolding enthalpy with temperature (ΔC_p effect) rather than the result of ligand binding *per se*.

4.2 Change in T_m

In DSC experiments the most obvious manifestation of ligand binding effects is a change in the apparent T_m of the protein under investigation. This can be calculated from the above expressions in the following way. The Gibbs free energy of unfolding at any temperature, T, is given by:

$$\Delta G_{unf} = -RT.\ln(K_{unf}) = \Delta G_{unf,0} \pm RT.\ln(1 + [L]/K_{L,N \text{ or } U})$$

and this will be zero at the mid-point (T_m) of the transition. Consequently, at T_m:

$$\Delta G_{unf,0} \pm RT_m.\ln(1 + [L]/K_{L,N \text{ or } U}) = 0$$

Ignoring any ΔC_p effects (i.e. assuming, for simplicity, that the enthalpy of unfolding is relatively constant over the temperature range considered here), the free energy of unfolding in the absence of ligand binding may be expressed in terms of the enthalpy of the transition ($\Delta H_{unf,0}$) at the mid-point temperature (T_{m0}):

$$\Delta G_{unf,0} = \Delta H_{unf,0} (1 - T/ T_{m0})$$

So that, after substitution in the previous equation and rearrangement:

$$\Delta T_m /T_m = \pm(RT_{m0} /\Delta H_{unf,0}).\ln(1 + [L]/K_L)$$

in which $\Delta T_m = T_m - T_{m0}$ is the change in unfolding transition temperature, and the ± sign relates to whether ligand stabilises the folded ($K_L = K_{L,N}$) or unfolded form ($K_L = K_{L,U}$), respectively. (Remember: ligand binding to the native fold, N, will increase T_m , whereas binding to the unfolded polypeptide, U, will decrease T_m .)

Important note: In the above derivations, we have defined T_m as the temperature at the mid-point of the transition, where ΔG_{unf} is zero and we have equal populations of folded and unfolded species. In DSC experiments involving simple single transitions, this usually corresponds to the peak of the thermogram (or is very close to this). However, in some situations with strongly binding ligands at less than stoichiometric concentrations (see below) one may see two peaks in the DSC trace, corresponding to separate unfolding of the free protein and protein-ligand complex. In such cases the "T_m" , as defined here for the entire mixture, lies at some point midway between these two peaks.

For the slightly more general case of multiple weakly binding ligands (15) these equations can be extended to give:

$$\Delta T_m /T_m = \pm(nRT_{m0} /\Delta H_{unf,0}).\ln(1 + [L]/K_L)$$

where n is the number of ligand binding sites on the protein (assumed identical). This is the behaviour encountered in the effects of cyclodextrins on folding stability, illustrated in Figure 8, where data fit best to a model assuming multiple binding sites on the unfolded polypeptide. These sites may be identified as the aromatic amino acid sidechains exposed during unfolding and to which the cyclodextrin molecules may attach (14,15).

At low concentrations, with weakly binding ligands ([L]/KL << 1) the expression describing the variation in T_m becomes approximately linear in ligand concentration:

$$\Delta T_m / T_m \approx \pm nRT_{m0}\,[L]/(K_L.\Delta H_{unf,0})$$

It is important to note that the T_m shift continues with increasing ligand concentration even beyond levels where the protein is fully ligand-bound. This is a manifestation of the dominant entropy of mixing contribution described above. A common misconception here is that it is the "bonds" formed between the ligand and the protein that somehow hold the protein in a more stable conformation. But if this were the case, then no further stabilisation would occur once all protein sites were saturated, and this is clearly contrary to observation. The thermodynamic rationalisation is that enhanced stability arises from the additional free energy required to remove the ligand from the protein prior to its unfolding (or vice versa), and this free energy has an important component arising from the entropy of mixing of dissociated ligand. This has nothing to do with protein conformation, but depends on the relative concentration of free ligand in solution.

4.3 Effects when Ligand binds to both N and U

Occasionally cases arise where, contrary to the above argument, the T_m shift does reach a plateau at higher ligand concentrations. This usually signifies binding of L to *both* N and U, albeit with different affinities. For example, a particular ligand might bind strongly to the native protein but less well to the unfolded chain. In such cases the T_m would shift upwards with increasing [L] until the concentration is such that both N and U are fully liganded. An example of this is found with α-lactalbumin, a specific calcium binding protein where increasing $[Ca^{2+}]$ progressively stabilises the native protein up to a limit where weak, non-specific calcium ion binding to the unfolded chain sets in (Figure 9; Robertson, Cooper & Creighton - unpublished observations). Following the same procedures as before, one can show that the T_m shift in such cases is given by:

$$\Delta T_m / T_m = (RT_{m0} / \Delta H_{unf,0}).\ln\{(1 + [L]/K_{L,N})/(1 + [L]/K_{L,U})\}$$

In the high concentration limit ([L] >> $K_{L,N}$ and $K_{L,U}$) this becomes constant, independent of ligand concentration:

$$\Delta T_m / T_m = (RT_{m0} / \Delta H_{unf,0}).\ln\{K_{L,U}/K_{L,N}\}$$

with the plateau value depending on the ratio of ligand binding affinities to the folded and unfolded protein. This also shows that no T_m shift occurs if the ligand binds equally well to both states, and simply the ability to bind to the folded protein is no guarantee that a ligand will stabilise the fold. As a rough rule of thumb, each factor of 10 in the ratio of ligand binding affinities, $K_{L,U}/K_{L,N}$, will give a ΔT_m of about 5°C.

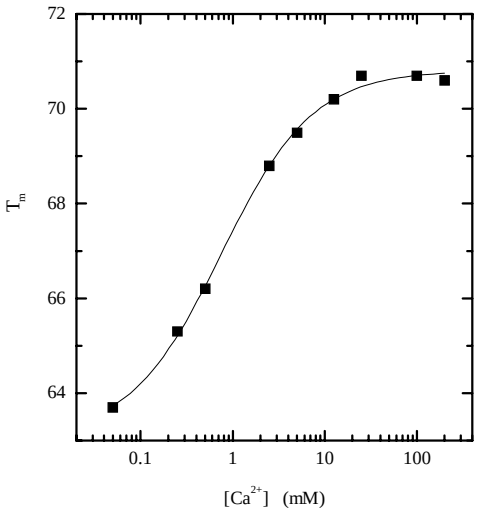

Fig. 9 Effect of calcium ion concentration (log scale) on the T_m of bovine α-lactalbumin, determined from DSC experiments in 0.1M Tris/HCl buffer, pH 8.0, containing 0.1M NaCl. The plateau in T_m at high $[Ca^{2+}]$ indicates binding to the unfolded state, as well as the stabilisation brought about by binding to the native fold.

4.4 One peak or two ?

Analysis of more complex situations involving multiple ligand binding or more tightly binding ligands is generally less straightforward than outlined above, although the same basic principles apply. See refs. (12,16) for details. One interesting problem is related to how many peaks one might expect to see in a DSC endotherm for a mixture of protein plus ligand. In the examples looked at so far, the ligand binding has been relatively weak, and usually the ligand concentration is significantly greater than the protein. In such cases (assuming a simple two-state process) one sees not two peaks corresponding to bound and free protein, but rather a single endotherm, gradually shifting up or down in T_m as the ligand concentration is varied. This is because at any one time the system is in rapid dynamic equilibrium between ligand-bound and ligand-free states, and what we see is a thermodynamic average of the two - assuming, as is usually the case, that on/off exchange of ligand is fast on the DSC time scale.

There are occasions, however, when one might observe separate discrete peaks in a DSC experiment arising from the separate unfolding of apo- and ligand-bound protein, particularly in the case of very tightly binding ligands at sub-stoichiometric concentrations.

One such example of this kind of behaviour has been seen in DSC experiments (Figure 10) of the thermal unfolding of a repressor protein (the methionine repressor, MetJ) in the presence of increasing concentrations of a specific DNA fragment corresponding to its consensus target (17,18). Here, when the protein concentration is in excess over the DNA, we see two peaks in the DSC trace corresponding to separate unfolding of free protein (lower T_m) and protein-DNA complex (higher T_m), respectively. The proportion of the endotherm corresponding to the higher temperature transition increases with increase in DNA fragment concentration until, when the DNA is equimolar with repressor protein, only a single transition corresponding to unfolding of protein-DNA complex is seen. (The melting of the DNA duplex used in these experiments does not occur until even higher temperatures under the experimental conditions used.)

Protocol 4. DSC of protein-DNA complex

Equipment and reagents

- DSC instrument (Microcal MC2, VP-DSC, or equivalent)
- Dialysis tubing or cassettes (e.g. Pierce Slide-A-Lyser®)
- Degassing equipment (vacuum desiccator, magnetic stirrer)
- Buffer.
- Protein + DNA: typically 2ml at a (protein) concentration of 1mg/ml.

A. Sample Preparation

1. Prepare the protein:DNA mixture and dialyse several changes of appropriate buffer. Each DSC run will typically require 1-2 ml of protein solution at a concentration of around 1 mg/ml, together with similar stoichiometric amounts of DNA. Depending on conditions, it may be preferable to dialyse the protein and DNA in separate dialysis bags (though in the same pot) for subsequent mixing.
2. Retain the final dialysis buffer for DSC reference, equilibration and dilutions.
3. Determine the protein and DNA concentrations by 280/260nm absorbance or other appropriate method.
4. Immediately prior to the experiment, degas portions of the sample mixture and buffer for 2-3 minutes under gentle vacuum with gentle stirring. Be careful to avoid excessive degassing or frothing of the mixture at this stage.

B. DSC procedure

1. Load DSC sample and reference cells with degassed buffer and collect baseline scan(s) using appropriate temperature range and scan rate (typically 20 – 100 °C, 60 °C/hr).

2. Allow the DSC cells to cool and refill the sample cell with protein:DNA mixture.
3. Repeat the DSC scan(s) using the same parameters as in 1. (Depending on circumstances, it may be useful to do repeat scans with the same sample to establish reversibility and reproducibility.)
4. After final cooling, remove the sample an examine for turbidity, aggregation or other visible changes.
5. Process data using instrumental software.

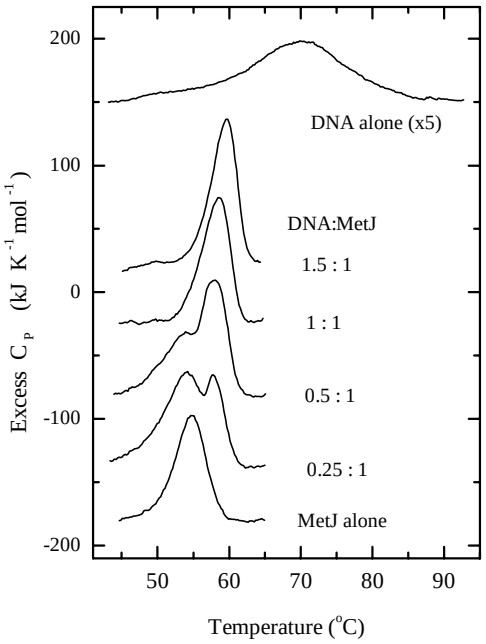

Fig. 10 DSC thermograms (concentration normalized and corrected for buffer baselines) for MetJ protein, repressor DNA 16 bp consensus fragment, and protein-DNA mixtures at different stoichiometric ratios; 25mM phosphate, 0.1M KCl, 1mM DTT, pH 7; DSC scan rate 60 °C hr^{-1}. The data for melting of the DNA alone (upper curve) have been expanded for clarity. (Adapted from ref.6)

In cases where ligand binding is tight, and where separate unfolding peak might be observed at low ligand:protein ratios, the effects can be calculated as follows.

Assume, for simplicity, a standard two-state unfolding equilibrium in which the ligand binds only to the native state. (Equivalent calculations can be done for other scenarios.) As usual we start with the equilibrium expressions describing ligand dissociation and protein unfolding, respectively:

$$NL \rightleftharpoons N + L \quad ; \quad K_{L,N} = [N][L]/[NL]$$

$$N \rightleftharpoons U \qquad ; \qquad K_0 = [U]/[N]$$

We wish to calculate the concentrations of the different species ([N], [U], [NL], [L]) present at equilibrium under any conditions. The total concentrations of protein and ligand, respectively, are written:

$$C_{tot} = [N] + [NL] + [U]$$
$$C_L = [NL] + [L]$$

here we cannot (as we did previously) assume that the free ligand concentration, [L], is the same as the total, CL, and we must rearrange the above equations and solve for each of the species in turn.

Using the equilibrium constant expressions we get:

$$C_{tot} = [NL] + [N](1 + K_0) = [NL]\{1 + K_{L,N}(1 + K_0)/[L]\}$$

which, using the expression for C_L , gives a quadratic in [NL] :-

$$[NL]^2 - [NL]\{C_{tot} + C_L + K_{L,N}(1 + K_0)\} + C_L C_{tot} = 0$$

with the solutions:

$$[NL] = ½[\{C_{tot} + C_L + K_{L,N}(1 + K_0)\} \pm (\{C_{tot} + C_L + K_{L,N}(1 + K_0)\}^2 - 4C_L C_{tot})^{1/2}]$$

By inspection, the negative sign in the ± is the physically realistic choice, so that the concentration of protein-ligand complex under any conditions is given by:

$$[NL] = ½[\{C_{tot} + C_L + K_{L,N}(1 + K_0)\} - (\{C_{tot} + C_L + K_{L,N}(1 + K_0)\}^2 - 4C_L C_{tot})^{1/2}]$$

and the concentrations of other species are obtained by application of the above expressions to give:

$$[N] = (C_{tot} - [NL])/(1 + K_0)$$

$$[U] = K_0[N]$$

$$[L] = C_L - [NL]$$

A procedure for calculating or modelling such behaviour, suitable for simple spreadsheet application, is give in Protocol 5, with typical results shown in Figure 11. Note how in this particular example, simulating a tight-binding ligand present at half the protein concentration, the unfolding occurs in two discrete steps, corresponding to unfolding of N and NL, respectively. Under these conditions, essentially all the ligand remains bound until released by unfolding of the complex.

Protocol 5. Spreadsheet simulation/calculation

This protocol is provided for those who wish step-by-step instructions to construct a spreadsheet or modelling package to simulate the effects of ligand binding on protein stability. Any proprietary spreadsheet or scientific graphics package should be adequate (e.g. Microsoft Excel™, Microcal ORIGIN™). Here we assume that ligand binds only to the native protein, but the procedure is easily modified for other models.

Constants:

R = 1.987 (cal K^{-1} mol^{-1}) or 8.314 (J K^{-1} mol^{-1}) - the Gas constant
T_0 = 273.15 (to correct °C to absolute)

User variables: [typical value]

T_{m0}	=	mid-point T_m without ligand (in °K)	[323 K ≡ 50 °C]
$\Delta H_{unf,0}$	=	unfolding enthalpy at T_{m0}	[100 kcal mol^{-1}]
ΔC_p	=	heat capacity increment	[1.3 kcal mol^{-1} K^{-1}]
$K_L(298)$	=	ligand dissociation constant at 25 °C	[10^{-10} M]
ΔH_L	=	ligand association enthalpy	[-10 kcal mol^{-1}]
C_{tot}	=	total protein concentration	[10^{-5} M]
C_L	=	total ligand concentration	[5 x 10^{-6} M]

Spreadsheet columns:

A: t Range of temperatures (e.g. 10-100 °C, in 1° steps)
B: T = t + T0 (convert to absolute temperature)
C: $\Delta G_{unf} = \Delta H_{unf}(T_{m0}).\{1 - T/T_{m0}\} + \Delta C_p.\{T - T_{m0} - T.\ln(T/T_{m0})\}$
D: $K_0 = \exp\{-\Delta G_{unf}/RT\}$ (unfolding equilibrium at T)
E: $K_L = K_L(298).\exp\{-(\Delta H_L/RT)(1 - T/298)\}$ (ligand equilibrium at T)

F:	X	$= C_{tot} + C_L + K_L(1 + K_0)$	
G:	[NL]	$= ½\{ X - (X^2 - 4C_LC_{tot})^{½}\}$	(concentration of protein-ligand complex at T)
H:	[L]	$= C_L - [NL]$	(concentration of free ligand at T)
I:	[N]	$= K_L[NL]/[L]$	(conc. of native, unbound protein at T)
J:	[U]	$= K[N]$	(concentration of unfolded protein at T)
K:	Cp	$= d[U]/dT$	(unfolding thermogram - arbitrary scale)

Graphical presentation of columns G-K (plotted versus t) will give data such as illustrated in Figure 11.

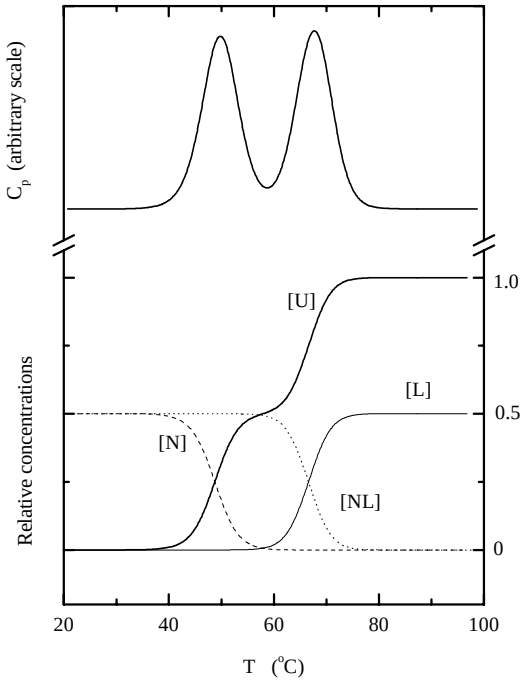

Fig. 11 Simulated data (using Protocol 5) for the unfolding of a protein in the presence of a tightly binding ligand, using the following parameters: C_{tot} = 10 μM, C_L = 5 μM, T_{m0} = 50 °C , $\Delta H_{unf,0}$ = 100 kcal mol^{-1}, ΔC_p = 1.3 kcal K^{-1} mol^{-1} , K_L(at 25 °C) = 10^{-10} M , ΔH_L = - 10 kcal mol^{-1} . The relative concentration of native protein ([N], dashed line) falls with increasing temperature, with a mid-point corresponding to the T_m of the protein in the absence of ligand. The protein-ligand complex ([NL], dotted line) unfolds at a higher temperature, with concomitant release of ligand. The peaks in the C_p thermogram correspond to maxima in the rate of change of [U].

4.5 Hydrogen Ions as Ligands: the effect of pH on protein stability

The effects observed by varying pH in DSC experiments on proteins is just a special case of the ligand-binding consequences described above for situations where ligand binds to both folded and unfolded states. Here the ligands are aqueous hydrogen ions (H^+) that will bind to specific amino acid sidechains or terminal groups (acidic or basic) in both folded and unfolded states. Only if the proton binding affinities are different in the two states will pH have any effect on stability.

For proton binding to a single group on the polypeptide, the acid-base equilibrium for folded and unfolded states may be described separately in terms of the usual equilibrium expression involving the acid dissociation constants $K_{A,N}$ and $K_{U,N}$ for the N and U states, respectively:

$$N + H^+ \rightleftharpoons NH^+ \quad ; \quad K_{A,N} = [N][H^+]/[NH^+]$$

$$U + H^+ \rightleftharpoons UH^+ \quad ; \quad K_{A,U} = [U][H^+]/[UH^+]$$

so that apparent or effective equilibrium constant for protein unfolding is given by:

$$K_{unf} = ([U] + [UH^+])/([N] + [NH^+]) = K_0.(1 + [H^+]/K_{A,U})/ (1 + [H^+]/K_{A,N})$$

where $K_0 = [U]/[N]$ is the equilibrium constant for unfolding of the unprotonated protein.

Consequently it follows that the stability of the folded protein (with respect to unfolded) can only be affected by changes in pH if $K_{A,N}$ is different from $K_{A,U}$, as might arise for example if unfolding gives rise to a change in environment or electrostatic interactions of the ionisable group.

In more realistic situations with multiple ionisable groups the pH-dependence is somewhat more complex, but the same general principles still apply, and changes in pH can only affect folding stability if the ionisable group(s) have different pK_A values in the folded and unfolded states.

The converse of this argument shows that, where the stability of the folded protein is sensitive to pH, the folding/unfolding transition must be accompanied by an uptake or release of hydrogen ions. This is one example of the general theory of linked thermodynamic functions (19,20), which can be viewed as the rigorous expression of Le Chatelier's principle. Here can be used to derive the change in number of H^+ ions bound (δn_{H+}) when the protein unfolds in terms of the pH dependence of the unfolding equilibrium:

$$\delta n_{H+} = - \partial \log K_{unf}/\partial pH$$

Shifts in pK (designated δpK) also correspond to changes in standard free energy of proton ionisation of the group ($\delta\Delta G°_{ion}$) which are numerically related by:

$$\delta\Delta G°_{ion} = -2.303RT.\delta pK$$

where R is the gas constant (8.314 J K^{-1} mol^{-1}) and T the absolute temperature. This corresponds to a free energy change of almost 6 kJ mol^{-1} for each unit shift in pK at room temperature.

Figure 12 shows typical DSC data for the effect of pH changes on the thermal unfolding of a simple globular protein (lysozyme). The major change in T_m in the low-pH region occurs over the pH 2-3 range, consistent with anticipated protonation of carboxylate side chains, and the variation corresponds to a cumulative uptake (δn_{H+}) of about 3 hydrogen ions per protein molecule in the middle of this pH range where the effect is most pronounced. The slight decrease in T_m above pH 4.5 may reflect titration of a histidine group, but may also be an artefact arising from the tendency of unfolded lysozyme to aggregate in the neutral pH region.

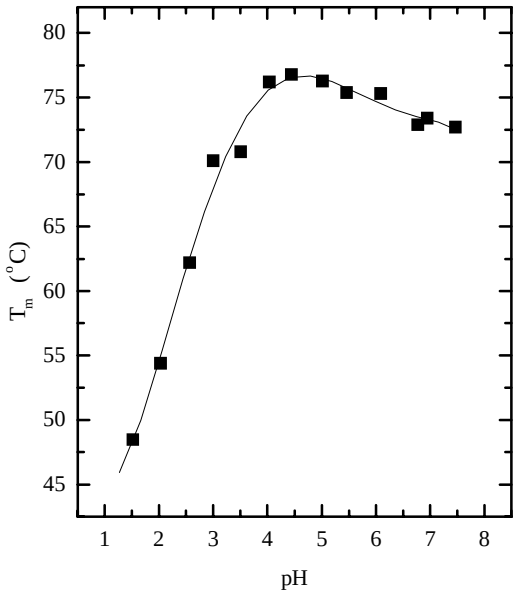

Fig. 12 Variation in T_m for unfolding of hen egg white lysozyme with pH, reflecting preferential binding of protons to residues in the unfolded chain. Determined by DSC (Protocol 1) in the pH 1.5 - 7.5 range, using the following buffers: KCl/HCl (pH 1.5-2.0), glycine/HCl (pH 2.5-3.0), sodium citrate/citric acid (pH 3.5), sodium acetate/acetic acid (pH 4.0-5.5), MES (pH 6.1 - 6.8) and MOPS (pH 7.0 - 7.5).

4.6 Non-Specific Metal Ion Effects

In addition to the specific binding effects mentioned above, there are a variety of ways in which non-specific interactions of metal ions in solution may affect the stability of a protein and, therefore, show up in DSC experiments. Some of these are illustrated in Figure 13, taking lysozyme as an example, and we mention them here in order to illustrate the caution sometimes needed in interpreting results.

Firstly, there are simple ionic strength effects. In the case of lysozyme, most metal ions (salts) in solution tend to reduce the T_m in ways that might, at first sight, suggest direct ionic interaction with the unfolded polypeptide. Divalent metal ions seem to be more effective than simple 1:1 electrolytes in this respect. However, when plotted using the more appropriate ionic strength (I) scale, rather than concentration, we find that the T_m reduction effects for most simple salts fall on a common line, suggesting that this is just a non-specific electrostatic screening effect on protein stability. (Note: ionic strength, $I = \frac{1}{2}\Sigma z_i^2 c_i$, where c_i is the concentration of each ionic species and z_i is its valency (integral charge), is the appropriate quantity replacing concentration in ionic solutions to take account, at least at low concentrations, of ionic screening effects.) This might reflect the weakening of charge-charge interactions between groups in the protein, that might both weaken the native fold and allow greater flexibility of the unfolded chain.

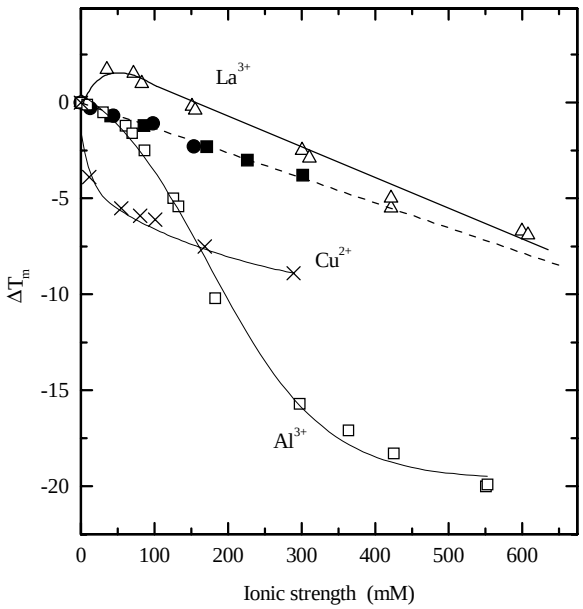

Fig. 13 Effects of various metal ions/ionic strength on the thermal stability of lysozyme. ΔT_m is the change in T_m determined from the peak of the DSC thermogram, following experimental procedures described in Protocol 1 with the addition of appropriate salts (chloride or sulphate) to the buffer. The solid symbols (dashed line) indicate the simple ionic strength effect for simple mono- or di-valent salts such as KCl and other alkali metal salts, $MgCl_2$, and $ZnCl_2$. Other data are for $LaCl_3$ (open triangles), $AlCl_3$ (open squares), and $CuSO_4$ (crosses). Data are plotted on a common ionic strength (I) scale.

Deviations from this general ionic strength effect are, however seen in some instances. For example (Figure 13), addition of lanthanum ions to the solution produces an initial rise in T_m for lysozyme, followed by a more gradual fall in T_m that parallels the ionic strength effect. This is consistent with the known binding of lanthanide ions to native lysozyme, confirmed by separate ITC and fluorescence methods, which will initially stabilise the folded state, superimposed on the more general ionic destabilising effect common to all salts. Other metal cations - Cu^{2+} and Al^{3+}, for example - show the opposite effect, giving destabilisation of the protein in excess of that expected from simple ionic strength effects. This suggests additional binding or other interactions of these particular metal ions with the unfolded polypeptide. In the case of copper ions the effect is the same even in the presence of 100mM KCl, suggesting that the interaction of these ions with exposed protein or peptide groups is not simply electrostatic. This is consistent with the known complexation properties of Cu^{2+} with peptide groups, known classically as the "biuret reaction". For Al^{3+}, the effect is accompanied by irreversible precipitation of the unfolded protein, which probably exaggerates the T_m reduction seen here. Non-equilibrium kinetic effects such as these are difficult to interpret using DSC alone.

Acknowledgement:

Much of the work reported here has been done under the auspices of the UK Biological Microcalorimetry Facility in Glasgow, funded jointly by BBSRC and EPSRC.

References

1. Privalov, P.L. and Potekhin, S.A. (1986). In *Methods in Enzymology*, Vol 131, pp. 4-51. Academic Press, London.
2. Plotnikov, V.V., Brandts, J.M., Lin, L-N. and Brandts, J.F. (1997). Analytical Biochemistry, **250**, 237-244.
3. Jackson, W.M. and Brandts, J.F. (1970). Biochemistry, **9**, 2294-2301.
4. Gill, S.C. and von Hippel, P.H. (1989). Anal.Biochem. **182**, 319-326.
5. Sanchez-Ruiz, J.M., Lopez-Lacomba, J.L., Cortijo, M. and Mateo, P.L. (1988). Biochemistry, **27**, 1648-1652.
6. Galisteo, M.L., Mateo, P.L. and Sanchez-Ruiz, J.M. (1991). Biochemistry, **30**, 2061-2066.
7. Haynie, D.T. and Freire, E. (1994) Anal. Biochem. **216**, 33-41.
8. McPhail, D. and Cooper, A. (1997). *J.Chem.Soc. Faraday Trans.*, **93**, 2283.
9. Dunitz, J.D (1995). *Chemistry & Biology*, **2**, 709.
10. Naghibi, H., Tamura, A. and Sturtevant, J.M. (1995). *Proc.Natl.Acad.Sci.USA*, **92**, 5597.
11. Sturtevant, J.M. (1974). Ann.Rev.Biophys.Bioeng. **3**, 35-51.
12. Sturtevant, J.M. (1987). Ann.Rev.Phys.Chem. **38**, 463-488.
13. Fukada, H., Sturtevant, J.M. & Quiocho, F.A. (1983). J.Biol.Chem. **258**, 13193-13198.
14. Cooper, A. (1992). J.Amer.Chem.Soc. **114**, 9208-9209.
15. Cooper, A., and McAuley-Hecht, K.E. (1993). Phil.Trans.R.Soc.Lond. A, **345**, 23-35.

16. Brandts, J.F. & Lin, L.-N. (1990). Biochemistry **29**, 6927-6940.
17. Cooper, A., McAlpine, A. and Stockley, P.G. (1994). *FEBS Lett*., **348**, 41-45.
18. Cooper, A. (1998). Microcalorimetry of protein-DNA interactions, in "Protein-DNA Interactions: A Practical Approach", ed. A.Travers & M.Buckle (Oxford Univ.Press).
19. Wyman, J. (1964). *Adv.Protein Chem*., **19**, 223-286.
20. Wyman, J. & Gill, S.J. (1990). *Binding and Linkage: functional chemistry of biological macromolecules*. University Science Books, Mid Valley, CA.
21. A. Cooper, M. A. Nutley, A. Wadood, *Differential scanning microcalorimetry* in S. E. Harding and B. Z. Chowdhry (Eds.), Protein-Ligand Interactions: hydrodynamics and calorimetry. Oxford University Press, Oxford New York, (2000) p 287-318.
22. A. Cooper, Heat capacity of hydrogen-bonded networks: an alternative view of protein folding thermodynamics, Biophys. Chem. (2000) 85, 25-39.