

The physical chemistry of high-sensitivity differential scanning calorimetry of biopolymers

Stephen Leharne¹

Received: 12 October 2016 / Accepted: 1 December 2016 / Published online: 27 December 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract High-sensitivity differential scanning calorimetry (HSDSC) is widely used to examine the thermal behaviour of biomolecules and water-soluble polymers in aqueous solution. The principal purpose of this manuscript is to examine the thermodynamic basis for the signals obtained using HSDSC. It is shown that a combination of the van't Hoff isochore and Kirchhoff's equation are all that is necessary to simulate and curve fit the HSDSC output obtained for the thermally induced unfolding of the protein ubiquitin. The treatment is further developed to show how the temperature dependence of the heat capacity change of unfolding, multiple sequential transitions, and protein dissociation can be incorporated into the thermodynamic description of protein unfolding and how these factors in turn affect the HSDSC signal.

Keywords Calorimetry · Thermodynamics · Biomolecules

Introduction

High-sensitivity differential scanning calorimetry (HSDSC) is widely employed for the study—in aqueous solution—of the thermodynamic parameters associated with processes initiated either by an increase in temperature (up-scan) or by a decrease in temperature (down-scan). Small molecular

mass molecules cannot be examined by HSDSC unless they form aggregate structures showing intermolecular co-operation. On the other hand, biopolymers in aqueous solution, such as proteins, which are cooperatively stabilised by numerous weak forces, can be examined by HSDSC.

Typically HSDSC can be used to examine:

1. Transitions from the physiologically active native form of a protein through intermediate partially unfolded states to the final denatured form of the protein. Very often, such a process is characterised by minimally populated intermediate states and thus approximates to a two-state transition between the initial native form and the final denatured form of the protein [1].
2. Thermally induced co-operative transitions in molecular assemblies of phospholipids, such as multi-lamellar liposomes [2].
3. Melting transitions in DNA and oligonucleotides [3].

In HSDSC, the specific heat of an aqueous system is measured as a function of temperature. For an aqueous solution of a bio-polymer, the apparent specific heat of the solute (S_2) is given by the following expression [1]:

$$S_2 = S_1 + \frac{1}{w_2}(S - S_1) \quad (1)$$

where S is the specific heat of the solution, S_1 is the specific heat of the solvent, and w_2 is the weight fraction of the solute. Because the quantity $(S - S_1)$ is usually very small, a differential mode of measurement [solvent (reference cell) versus solvent plus solute (sample cell)] has to be used. Indeed, given that a major portion of the specific heat change is due to the heating and cooling of the solvent (usually water which has a large heat capacity), it is essential to have a differential arrangement, so that phase transitions in the solute can be observed.

✉ Stephen Leharne
S.A.Leharne@gre.ac.uk

¹ Faculty of Engineering and Science, University of Greenwich, Pembroke, Chatham Maritime, Kent ME4 4TB, UK

HSDSC signals and their interpretation for protein unfolding

DSC instruments measure the power required to maintain the temperature of a sample placed in a designated sample cell at (or close to) the same value as that of a reference cell containing the identical aqueous solvent, but no sample molecule, as the overall temperature of the system, is altered. The cells are located within an adiabatic vacuum chamber. The raw instrumental output conventionally shows power as a function of temperature. To extract data that have more thermodynamic significance, the axes of the trace output are transformed. Power is converted to a molar excess heat capacity using the formula:

$$\frac{dq_p}{dt} \cdot \frac{1}{\sigma M} = C_{p,ex} \quad (2)$$

where q_p is the heat absorbed at constant pressure; t is time, the derivative dq_p/dt represents power; σ is the scan rate (dT/dt , where T is temperature); and M is the number of moles of sample in the sample cell.

A typical DSC experiment normally involves at least two scanning runs. One scan consists of a baseline scan, wherein the sample cell and reference cell both contain the blank aqueous solvent. The second scan is a scan of the solvent (reference cell) against the solvent plus solute (sample cell). The baseline scan is then subtracted from the sample scan.

Figure 1 provides a typical example of an HSDSC trace of the excess heat capacity (the heat capacity difference between the sample and reference cells) as a function of temperature. The signal shown in Fig. 1 was obtained for the protein ubiquitin in buffer solution, at a pH of 2.

Proteins undergo denaturation on heating. The process involves a transition from the physiologically active compact folded form to the normally physiologically inactive unfolded form. Native protein structures in aqueous

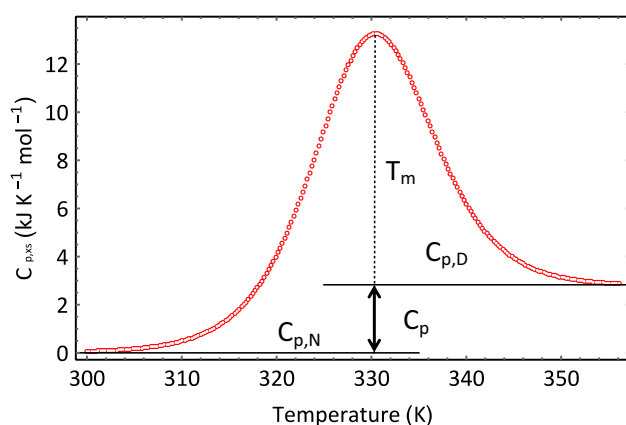


Fig. 1 Data obtained for a 5 mg cm^{-3} solution of ubiquitin at a pH of 2

solution are cooperatively stabilised by numerous intramolecular forces. Disruption of these forces requires an endothermic enthalpy change. The favourable free energy contribution to denaturation is provided by the entropy change that arises from the increased conformational freedom available to the unfolded protein and the increased number of ways of partitioning the increased thermal energy.

A simple pedagogic model of thermally induced unfolding has been described by Dill and Bromberg [4]. Consider a four-bead molecular chain, as shown in Fig. 2. In this model, the ground state is characterised by a compact molecular structure that is held together by an intramolecular bond (the dashed line) between the chain-ends. The first excited microstate is fourfold degenerate—i.e., there are four different unfolded molecular conformational structures, of equal energy that the molecule can adopt. The fractional occupancy of the two different energy states and its functional relationship with temperature can be calculated using the Boltzmann distribution equation:

$$\frac{n_1}{n_0} = \frac{g_1 e^{-\frac{\epsilon_1}{kT}}}{g_0 e^{-\frac{\epsilon_0}{kT}}} \quad (3)$$

The subscripts 0 and 1 denote the ground state and first excited microstate, respectively; n is the number of molecules in a particular state; g is the degeneracy of that state with $g_0 = 1$ and $g_1 = 4$; ϵ is the energy of the state; k is the Boltzmann constant; and T is the absolute temperature.

Using the mass balance expression $n = n_0 + n_1$ and $\Delta\epsilon = \epsilon_1 - \epsilon_0$, we can rewrite Eq. 3 as

$$\frac{n - n_0}{n_0} = 4e^{-\frac{\Delta\epsilon}{kT}} \quad (4)$$

which gives the following expressions for the fraction of molecules in the ground state:

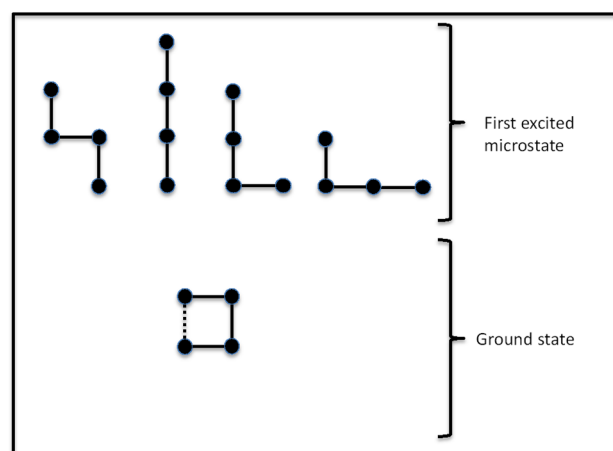


Fig. 2 Four-bead model for molecular unfolding. (Redrawn from Dill and Bromberg [4])

$$\frac{n_0}{n} = \frac{1}{1 + 4e^{-\frac{\Delta\varepsilon}{kT}}} \quad (5)$$

and fraction of molecules in the excited state:

$$\frac{n_1}{n} = \frac{4e^{-\frac{\Delta\varepsilon}{kT}}}{1 + 4e^{-\frac{\Delta\varepsilon}{kT}}}. \quad (6)$$

The temperature dependence of the composition of the system is shown in Fig. 3. This system is an example of a two-state system, i.e., a system within which only two states are significantly populated. At low temperatures, the ground state form predominates. The enthalpy of intramolecular binding is key to this predominance. However, as the temperature rises, the excited state becomes increasingly populated, thereby demonstrating the increasingly important entropic contribution of conformational variety to the system. The statistical thermodynamic description of protein unfolding is far more complex than the four-bead molecular model, but the model does encapsulate one of the reasons as to why proteins unfold upon heating to moderately high temperatures—the large number of excited state conformers. Moreover, just like the model, protein unfolding is, very often, a two-state process. As a consequence, the signal shown in Fig. 1 can be interpreted as showing how the thermal history of the system reflects the changing composition of the aqueous protein system as temperature increases. At low temperatures, the compact native form predominates as the temperature is increased some molecules begin to unfold. The fraction of molecules that have unfolded multiplied by the enthalpy of the unfolding transition provides the basis of the heat signal. Since the enthalpy change is endothermic, the temperature of the sample cell becomes lower than that of the reference cell; and thus, the instrument measures the power needed to raise the temperature to compensate for

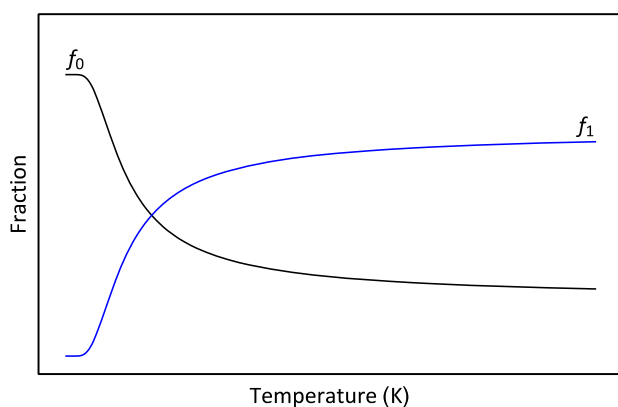


Fig. 3 Changing composition of the four-bead molecular system as a function of temperature. The fraction of molecules in the ground state is designated f_0 , and the fraction of molecules in the first excited state is designated f_1

the temperature difference. This, as we have shown, is easily converted into a molar excess heat capacity.

The DSC signal and initial data analysis

There are several features of the DSC signal, as shown in Fig. 1, which require comment. The transition from the compact physiologically active form of the protein to the more open unfolded physiologically inactive molecular form is shown as an increase in heat capacity of the system, going through a maximum at a temperature designated as T_m and then decreasing to a final higher final heat capacity value. The initial low temperature portion of the scan represents the heat capacity of the native form of the protein in aqueous solution (denoted as $C_{p,N}$). The high-temperature portion of the scan shows the heat capacity in aqueous solution of the unfolded form of the protein (denoted as $C_{p,D}$). In this scan, both heat capacities are assumed to be invariant with the temperature over the temperature range of the experimental run, so the heat capacity change on unfolding given by the expression $\Delta C_p = C_{p,D} - C_{p,N}$ is a constant. Formally, the molar heat capacity is the amount of heat energy required to raise the temperature of 1 mol of substance through 1 K. At the molecular level, the additional heat energy is distributed among the various degrees of freedom and partitioned variously between kinetic energies—including vibrational, rotational, and translational transitions and potential and potential energies—including stretching and bending of molecular bonds [5].

The existence of the heat capacity change indicates that both enthalpy and entropy are functionally dependent upon temperature. If ΔC_p is constant, then we can write:

$$\begin{aligned} \left(\frac{\partial \Delta H}{\partial T}\right)_p &= \Delta C_p \xrightarrow{\text{Integration}} \Delta H(T) \\ &= \Delta H(T_{\text{ref}}) + \Delta C_p(T - T_{\text{ref}}). \end{aligned} \quad (7)$$

Using the second law of thermodynamics, we get a similar expression for the entropy change:

$$\begin{aligned} d\Delta S &= \frac{dq_{\text{rev}}}{T} = \frac{d\Delta H}{T} = \frac{\Delta C_p}{T} dT \xrightarrow{\text{Integration}} \Delta S(T) \\ &= \Delta S(T_{\text{ref}}) + \Delta C_p \ln\left(\frac{T}{T_{\text{ref}}}\right). \end{aligned} \quad (8)$$

There are several reasons why the overall heat capacity change for protein unfolding increases. These include the exposure of hydrophobic amino acid side chains buried in the core of the native form of the protein to water when the protein molecule unfolds. For example, Connelly and Thomson [6] noted that the dissolution of aliphatic and aromatic hydrocarbons in water invariably leads to an increase in heat capacity. However, there are likely to be

other contributions, such as an increase in easily excitable vibrational modes upon unfolding.

The enthalpy of denaturation can be obtained from the experimental data by integration of the peak area. This value is referred to as the calorimetric enthalpy (ΔH_{cal}). To obtain the peak area, we must first draw a baseline to the data. In the example shown in Fig. 4a, a straight line is drawn from what is judged to be the start of the transition to the adjudged end of the transition. Once satisfied that the baseline satisfactorily joins the onset and termination of the thermal transition, it is then subtracted from the HSDSC data (Fig. 4b) to leave a transitional profile with a flat baseline. The resultant signal can then be divided up into evenly spaced segments. The area of the individual segments can then be computed either by the trapezoidal rule or by the Simpson's rule and then summed to give the integrated peak area (Fig. 4c).

The trapezoidal rule equation is

$$\text{Area} = [f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + f(x_n)] \frac{\text{Interval}}{2}$$

whilst the Simpson's rule equation is

$$\text{Area} = [f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \dots + 4f(x_{n-1}) + f(x_n)] \frac{\text{Interval}}{3}$$

Using the trapezoidal rule, ΔH_{cal} was found to be 198 kJ mol^{-1} .

Straight-line baselines are convenient and easily drawn but do not necessarily reflect the true geometry of the underlying baseline. Other baselines can be fitted. In Fig. 4d, the pre- and post-transitional portions of the signal are fitted to a cubic polynomial of the form: $f(x) = ax^3 + bx^2 + cx + d$. Other functions can be used—for example, quartic order polynomials or cubic splines—which may represent the underlying baseline better. Yet, normally, area integration using a straight-line baseline provides values not too dissimilar to values obtained using other baseline functions.

Can we use thermodynamics to examine the HSDSC signal?

Thermodynamic information may be obtained from the signal if it can be established that the signal was obtained under conditions of thermodynamic equilibrium. Thermodynamic control of the thermal processes observed in the calorimeter may be established by investigating the reversibility of the system. If the system either reproduces the same trace on rescanning or produces an identical trace

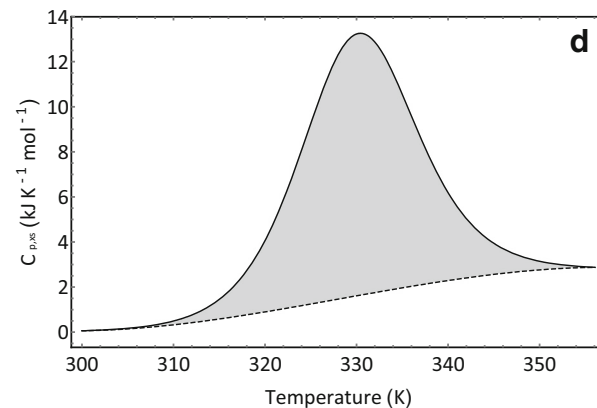
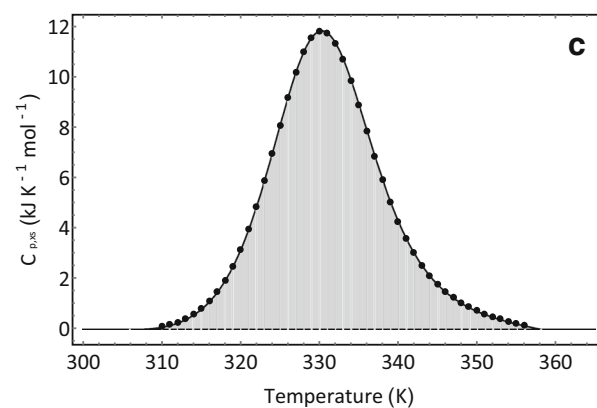
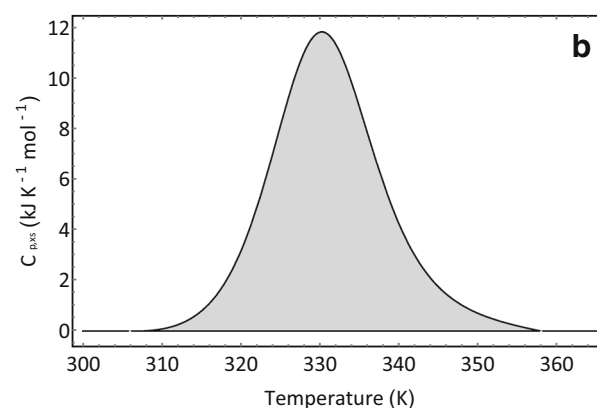
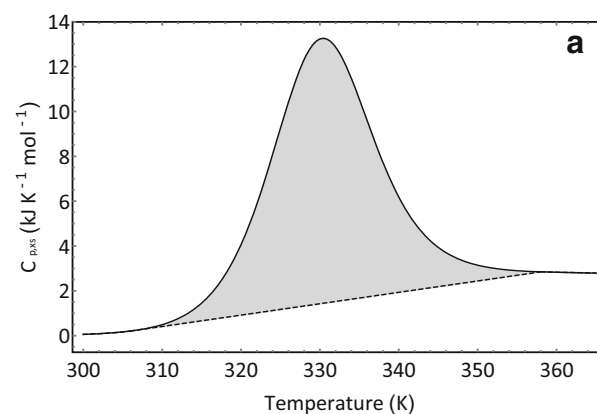


Fig. 4 Integration of the peak area. **a** Baseline is constructed so as to connect the start of the transition and the end of the transition. **b** Baseline is subtracted from the data. **c** Area under the peak is divided up into evenly spaced segments that are then used to calculate the area either using the trapezoidal rule or the Simpson's rule. **d** Baseline fitted to the pre- and post-transitional portions of the signal using a cubic polynomial (see text for details)

on cooling the application of thermodynamic relationships to aid our understanding of the HSDSC trace is justified. A further test of the applicability of thermodynamics is to examine the HSDSC signals for scan rate dependence. Parameters measured for processes under thermodynamic control show no scan rate dependence [7].

For proteins where unfolding is a two-state process the fraction of protein in the native state is given by the following expression that is analogous to the expression for the ground state in the four-bead model:

$$f_N = \frac{1}{1 + e^{-\frac{\Delta G}{RT}}} \quad (9)$$

and similarly for the denatured state

$$f_D = \frac{e^{-\frac{\Delta G}{RT}}}{1 + e^{-\frac{\Delta G}{RT}}} \quad (10)$$

The energy difference between the two states is the free energy of denaturation (ΔG). The ratio of the fraction of the denatured protein to the fraction of native protein is equal to the ratio of the concentrations of the denatured ($[D]$) and native protein ($[N]$):

$$\frac{f_D}{f_N} = \frac{\frac{[D]}{P_t}}{\frac{[N]}{P_t}} = \frac{[D]}{[N]} \quad (11)$$

where P_t is the total concentration of protein.

Using Eqs. 7 and 8, the ratio is also equal to the following expression:

$$\frac{f_D}{f_N} = \frac{\frac{e^{-\frac{\Delta G}{RT}}}{1 + e^{-\frac{\Delta G}{RT}}}}{\frac{1}{1 + e^{-\frac{\Delta G}{RT}}}} = e^{-\frac{\Delta G}{RT}} \quad (12)$$

However, from fundamental thermodynamics, we know that

$$K_p = e^{-\frac{\Delta G}{RT}} \quad (13)$$

where K_p is the equilibrium constant at constant pressure for the unfolding process. Thus, the equilibrium constant for denaturation obtained under constant pressure conditions is equal to the concentration ratio of the native and unfolded forms:

$$K_p = \frac{[D]}{[N]} \quad (14)$$

The thermodynamic basis of the HSDSC signal

The fraction of unfolded protein at temperature, T , multiplied by the enthalpy of unfolding at the same temperature gives the enthalpy needed to unfold f_D of protein at temperature, T . The rate of change in this enthalpy value with temperature gives the excess heat capacity—the heat capacity difference between the sample and reference cells:

$$C_{p,xs} = \frac{d}{dT}(f_D(T)\Delta H_{cal}(T)) \quad (15)$$

To calculate $C_{p,xs}$, we need to be able to calculate the changing composition of the system. This is readily done using the mass balance expression:

$$P_t = [N] + [D] \quad (16)$$

where P_t is the total concentration of protein. Rearranging Eq. 14 to provide an expression for $[N]$ and substituting this in Eq. 16 give

$$P_t = [D] \left(1 + \frac{1}{K_p(T)} \right) \quad (17)$$

which can be rearranged to give the fraction of unfolded protein f_D :

$$\frac{[D]}{P_t} = f_D = \frac{K_p(T)}{1 + K_p(T)} \quad (18)$$

How do we calculate the changing composition of the system as a function of temperature? This is readily achieved using the van't Hoff isochore [4]:

$$\frac{\partial \ln(K_p(T))}{\partial T} = \frac{\Delta H_{vH}(T)}{RT^2} \quad (19)$$

where $\Delta H_{vH}(T)$ is the van't Hoff enthalpy (for a two-state process, this is equal to the enthalpy of unfolding) and R is the universal gas constant. We have already noted that the unfolding process is accompanied by a positive change in heat capacity, which means that the van't Hoff enthalpy is temperature dependent (see Eq. 7).

Substituting Eq. 7 into Eq. 19 and integration of the resultant expression give

$$\int_{\ln(K(T_{ref}))}^{\ln(K(T))} d \ln(K) = \int_{T_{ref}}^T \frac{\Delta H_{vH,ref} + \Delta C_p(T - T_{ref})}{RT^2} dT \quad (20)$$

Here, $\Delta H_{vH,ref}$ is the value of the van't Hoff enthalpy at T_{ref} and T_{ref} is the reference temperature, which is conveniently defined as the temperature at which the fractions of denatured and native protein are equal. This definition means that $K(T_{ref})$ is equal to unity. Equation 20 is thus written, after integration, as

$$K(T) = K(T_{\text{ref}}) \exp \left[\frac{\Delta H_{\text{vH,ref}}}{R} \left(\frac{1}{T_{\text{ref}}} - \frac{1}{T} \right) + \frac{\Delta C_p}{R} \left(\ln \left(\frac{T}{T_{\text{ref}}} \right) + \frac{T_{\text{ref}}}{T} - 1 \right) \right]. \quad (21)$$

Using Eq. 21 in Eqs. 18 and 16 allows us to calculate how the fractions of denatured and native protein changes vary with temperature.

Simulating and fitting the HSDSC signal

If we complete the differentiation, as shown in Eq. 15, we obtain

$$C_{p,\text{XS}}(T) = \Delta H_{\text{cal}}(T) \frac{df_D(T)}{dT} + f_D(T) \Delta C_{p,\text{cal}}. \quad (22)$$

In this equation, $\Delta C_{p,\text{cal}}$ is the heat capacity change obtained from the signal. We will find it convenient to differentiate between this value and the value of ΔC_p used in the van't Hoff derived equations (Eqs. 20 and 21). The relationship between the two parameters is given by

$$\Delta C_{p,\text{cal}} = \Delta C_p \frac{\Delta H_{\text{cal,ref}}}{\Delta H_{\text{vH,ref}}}. \quad (23)$$

To derive an analytical solution to Eq. 22, we need to find an expression for $\frac{df_D}{dT}$. This is achieved using the following transformation based upon the van't Hoff isochore (Eq. 19):

$$\left(\frac{\partial \ln(K(T))}{\partial f_D(T)} \right)_p \frac{df_D(T)}{dT} = \frac{\Delta H_{\text{vH}}(T)}{RT^2}. \quad (24)$$

Given

$$K(T) = \frac{f_D(T)}{1 - f_D(T)} \quad (25)$$

and expressing Eq. 25 as a logarithmic expression gives

$$\ln(K(T)) = \ln(f_D(T)) - \ln(1 - f_D(T)) \quad (26)$$

which on differentiation gives

$$\left(\frac{\partial \ln(K(T))}{\partial f_D(T)} \right)_p = \frac{1}{f_D(T)} + \frac{1}{1 - f_D(T)} \quad (27)$$

and thus provides

$$\frac{df_D(T)}{dT} = \frac{\Delta H_{\text{vH}}(T)}{RT^2} \frac{1}{f_D(T)} + \frac{1}{1 - f_D(T)}. \quad (28)$$

The excess heat capacity can thus be written as

$$C_{p,\text{XS}} = \frac{\Delta H_{\text{cal}}(T) \Delta H_{\text{vH}}(T)}{RT^2} \frac{1}{\frac{1}{f_D(T)} + \frac{1}{1 - f_D(T)}} + f_D(T) \Delta C_{p,\text{cal}}. \quad (29)$$

We now have an equation for $C_{p,\text{XS}}$ which is functionally related to temperature, T . Equation 29 can be used to fit the data shown in Fig. 1 using a least squares approach. The outcome of fitting Eq. 29 to our ubiquitin data is shown in Figs. 5 and 6. Figure 5 show how the composition of the system changes with temperature and Fig. 6 shows the optimised best fit line through the experimental data. The fitting was conducted in the following way. Initial values were assigned to the following parameters: ΔH_{vH} , ΔH_{cal} , ΔC_p , and T_{ref} . These were then used in the appropriate previously defined equations to calculate an initial set of values of $C_{p,\text{XS}}$ using the temperature data obtained from the data set, as shown in Fig. 1. The differences between the calculated values and the experimental values are calculated, squared, and summed. The sum of the squared differences was then minimised by changing the parameter values using

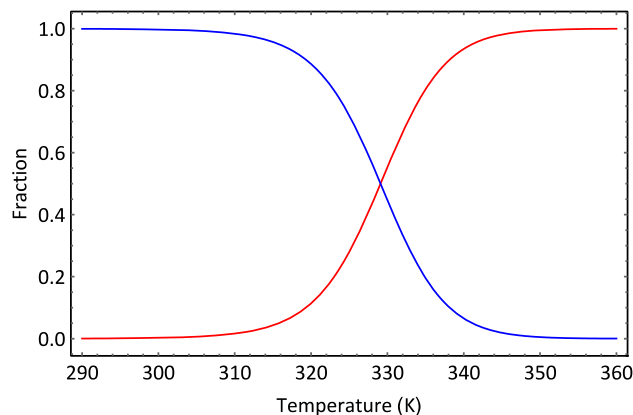


Fig. 5 Changing composition of an aqueous protein system as a function of temperature. The fraction of molecules in the native ground state is shown in blue, and the fraction of molecules in the denatured state is shown in red. The calculation was made using the fitted parameters obtained for ubiquitin unfolding at a pH of 2

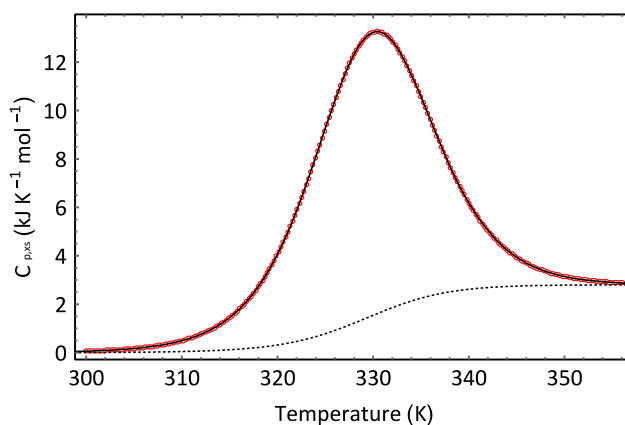


Fig. 6 Best fit line (inner line) to the data (orange open circle) displayed in Fig. 1 using Eq. 29. The dotted line is the baseline given by the expression

Table 1 Optimized fit parameters obtained for a 5 mg cm⁻³ solution of ubiquitin in buffer solution at a pH of 2

Parameter	Estimate
ΔH_{VH}	211 kJ mol ⁻¹
ΔH_{cal}	198 kJ mol ⁻¹
T_{ref}	329.1 K
ΔC_p	3.0 kJ K ⁻¹ mol ⁻¹

Table 2 Optimized fit parameters obtained for a 5 mg cm⁻³ solution of lysozyme in 1.0 M trehalose solution

Parameter	Estimate	Standard error
$\Delta H_{VH,ref}$	418 kJ mol ⁻¹	2.7
$\Delta H_{cal,ref}$	365 kJ mol ⁻¹	3.2
T_{ref}	332.4 K	0.03
a	-188.3 kJ mol ⁻¹ K ⁻¹	3.7
b	0.62 kJ mol ⁻¹ K ⁻¹	0.012
b	46.3 kJ mol ⁻¹ K ⁻¹	7.3
d	-0.08 kJ mol ⁻¹ K ⁻¹	0.02

the optimization routines implemented in the Computer Algebra program Mathematica (<http://www.wolfram.com>). For the example shown in Fig. 6, the NonlinearModelFit routine was used to fit the model to the data and provide a set of optimized parameters. The obtained optimized parameter values are shown in Tables 1 and 2. The standard errors in the parameters were very low—in the order of 0.1%.

Several observations can be made. The fitted calorimetric enthalpy is the same as the calorimetric enthalpy value measured by integration of the peak. The value of T_{ref} is slightly lower than the value that can be interpolated for T_m the temperature at which the excess heat capacity is a maximum.

The values obtained for ΔH_{VH} and ΔH_{cal} are close in value but not the same. This is not necessarily surprising. The units of both parameters are kJ mol⁻¹. However, as we have shown, the raw calorimetric data are converted into an excess heat capacity in Eq. 2 using a mass value as measured by the experimenter. In this experiment, the required mass of protein was weighed and dissolved in buffer. The aqueous protein solution was then injected into a cell of known fixed volume. In this way, the number of moles (M) is readily calculated and inserted into Eq. 2. On the other hand, the molar unit in the van't Hoff enthalpy is supplied by the universal gas constant. We can show this using dimensional analysis:

$$\left(\frac{\partial \ln[K_p(T)]}{\partial T}\right)_p = \frac{\Delta H_{VH}(T)}{RT^2} \rightarrow \text{Dimensional analysis } \frac{1}{K}$$

$$= \frac{\frac{\text{kJ}}{\text{mol}}}{\frac{\text{kJ}}{\text{mol K}} \times K^2} \quad (30)$$

It is important to note that logarithmic terms are dimensionless. Thus, in order for the equality to be true, the

molar unit used in the van't Hoff enthalpy must be identical to the molar unit supplied by the gas constant.

Therefore, why should ΔH_{cal} be less than ΔH_{VH} ? There are a number of explanations and it is indeed a matter of interest in protein thermochemistry. We make the assumption that the only thermal process on which the calorimetric enthalpy reports is the thermal unfolding of the protein. However, other thermal events may occur because of the presence of the protein in solution, which is not observed when running the baseline scans. However, it is also possible and indeed extremely likely that in our particular case not all the mass of protein placed in the cell was protein or that some of the protein placed in the cell did not undergo thermal unfolding. Both events would lead to overestimation of active protein and thus an underestimation of the calorimetric enthalpy.

We shall see later this article that there are cases, where $\Delta H_{cal} > \Delta H_{VH}$. This arises when protein unfolding involves the appearance of substantial populations of intermediates.

There are examples, where $\Delta H_{VH} \gg \Delta H_{cal}$. One example is the thermally driven change from the P_β gel phase to L_α in phospholipid multi-lamellar vesicles, where the ratio of the van't Hoff enthalpy to the calorimetric enthalpy can be as high as 200. Such numbers suggest that something like 200 phospholipid molecules are acting together as a co-operative unit.

Protein unfolding signals when the heat capacity change is dependent upon temperature

So far, we have assumed that thermally induced unfolding is characterised by a temperature invariant heat capacity. However, it is very often the case that the pre-transitional heat capacity shows a marked functional relationship with temperature, whilst the post-transitional is somewhat flatter—less influenced by temperature. Figure 7a provides an excellent example of this kind of thermal behaviour. The signal shown was obtained in laboratory class practical for the protein lysozyme in a 1.0 M aqueous solution of trehalose. The objective of the practical was to examine the behaviour of proteins in aqueous sugar solutions. The temperature dependence of the pre- and post-transitional heat capacities of the signal is readily incorporated into our analysis.

We assume that the temperature dependence of the pre- and post-transitional heat capacities can be described by a linear relationship. The heat capacity of the native protein is given by: $C_{p,N} = a + bT$, and for the unfolded protein, $C_{p,D} = c + dT$. This provided a heat capacity change that is temperature dependent:

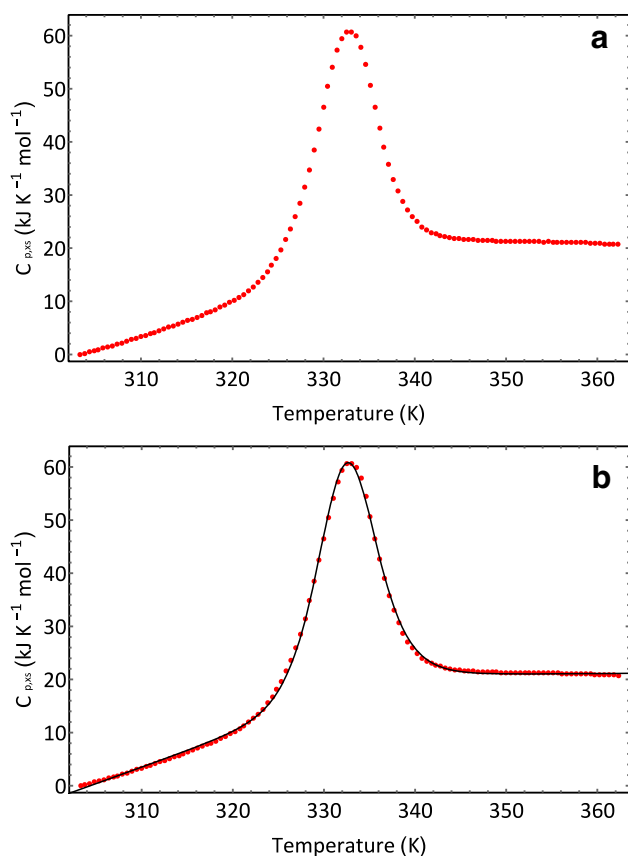


Fig. 7 **a** HSDSC data obtained for a 5 mg cm^{-3} solution of lysozyme in 1.0 molar solution of trehalose; and **b** line of best fit to the data

$$\Delta C_p(T) = (b - d)T + (c - a). \quad (31)$$

This is then used to provide a modified form of the Kirchoff equation:

$$\left(\frac{\partial \Delta H}{\partial T}\right)_p = \Delta C_p(T) = (b - d)T + (c - a) \longrightarrow$$

$$\text{Integration } \Delta H(T) = \Delta H_{\text{ref}} + (c - a)(T - T_{\text{ref}}) + (b - d)(T - T_{\text{ref}}^2). \quad (32)$$

Which then provides a modified form of Eq. 21 wherein, as before, $K_p(T_{\text{ref}}) = 1$:

$$\frac{K_p(T)}{K_p(T_{\text{ref}})} = \exp \left[\frac{\Delta H_{VH,\text{ref}}}{R} \left(\frac{1}{T_{\text{ref}}} - \frac{1}{T} \right) + \frac{(c - a)}{R} \left(\ln \left(\frac{T}{T_{\text{ref}}} \right) + \frac{T}{T_{\text{ref}}} - 1 \right) \right] \dots + \frac{(d - b)}{2R} \left(T - T_{\text{ref}} - T_{\text{ref}}^2 \left(\frac{1}{T_{\text{ref}}} - \frac{1}{T} \right) \right) \quad (33)$$

The temperature dependence of the heat capacity change results in the addition of another term to Eqs. 7 and 21.

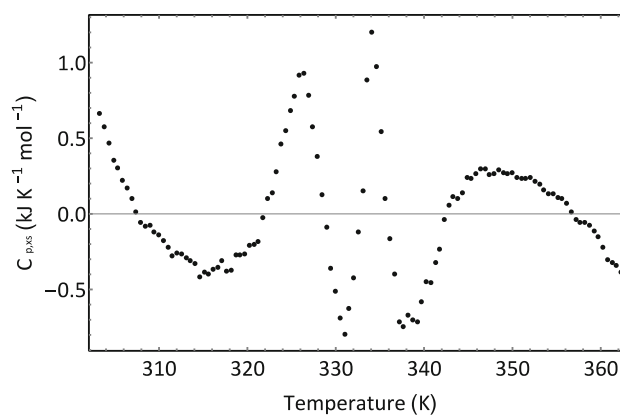


Fig. 8 Residual plot for the fit shown in Fig. 7b

Equations 31, 32, and 33 can then be used in a modified form of Eq. 29 to fit the lysozyme data. The modification is made to the last term in Eq. 29 and takes into account the changing contribution of the temperature-dependent heat capacities of the folded and unfolded forms make to the underlying base line:¹

$$\begin{aligned} f_D(T)C_{p,U}(T) + f_N(T)C_{p,N}(T) \\ \rightarrow f_D(T)C_{p,U}(T) + (1 - f_D(T))C_{p,N}(T) \\ \rightarrow f_D(T)\Delta C_p(T) + C_{p,N}(T). \end{aligned}$$

Equation 29 is thus rewritten as

$$\begin{aligned} C_{p,xs} = \frac{\Delta H_{\text{cal}}(T)\Delta H_{vH}(T)}{RT^2} \frac{1}{\frac{1}{f_D(T)} + \frac{1}{1-f_D(T)}} \\ + \frac{\Delta H_{\text{cal}}}{\Delta H_{vH}} (f_D(T)\Delta C_p(T) + C_{p,N}(T)). \end{aligned}$$

It will be recalled that the calorimetric enthalpy van't Hoff enthalpy ratio appears to appropriately scale the contribution of the underlying baseline to calorimetric signal. This modified form of Eq. 29 was fitted to data obtained for an aqueous solution of lysozyme in 1 M trehalose solution. The solution concentration was 5 g dm^{-3} . The results of this fit are shown in Fig. 7b. The optimised fit parameters are displayed in Table 2. The adjusted R^2 value for the fit is 0.999; the following optimized fit parameters were obtained. Both the adjusted R^2 value and Fig. 7b seem to suggest that the fit is extremely good. However, it is always good practice to look at a plot of the residuals. The residuals are calculated as the difference between the measured value for $C_{p,xs}$ and the calculated value $C_{p,xs}$ using the best fit parameters. A residual plot is shown in Fig. 8. If the residuals arise purely from the uncertainties in measurement—for example, instrumental noise, then it would be expected that the residuals would be located at random about the $C_{p,xs}$ axis.

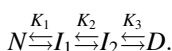
¹ In the treatment of ubiquitin the heat capacity of the native form was arbitrarily set to zero after the subtraction of the instrumental baseline and so does not appear in Eq. 29.

The fact that they are not reveals that there are systematic errors either in the data or in an error(s) in the model used to fit the data. Most probably, the analysis has neglected some other minor thermal events.

Of further note is the rather large discrepancy between the calorimetric and van't Hoff enthalpies. The calorimetric enthalpy is 87.3% of the value of the van't Hoff enthalpy. The lysozyme sample used was obtained from Sigma Aldrich who claims that its purity is $\geq 90\%$. For the experiment, the lysozyme was used as received which may thus explain the discrepancy.

Multiple transitions

The residual plot in Fig. 8 suggests that the model used to fit the data may have been too simple in that other thermal events may also occur which have not been incorporated into the model. If these thermal events are independent of the main transition, then the overall thermal transition is a simple arithmetic addition of the two underlying events. For example, small amounts of an impurity, which is also calorimetrically observable, may be present in the sample. It is also possible that in the case of a multi-sub-unit protein, the sub-units unfold independently of each other. Again, the thermal signal would be a composite of the underlying transitions. However, protein unfolding may involve the native protein undergoing a transition to one or several intermediate states before ultimately adopting the final unfolded form. Protein unfolding, under equilibrium conditions, by such a mechanism can be represented by the following mass action expression in the case, where there are two intermediates formed in significant quantities:



To be able to calculate the fraction of each species at any particular temperature, T , we formulate the following mass balance expression:

$$P_t = [N] + [I_1] + [I_2] + [D] \quad (34)$$

where P_t is the total protein concentration and $[]$ terms represent the equilibrium concentrations of the respective species. If we divide Eq. 34 by the $[N]$ and invert, we obtain the following expression for the fraction of native protein:

$$\frac{[N]}{P_t} = \alpha_N = \frac{1}{1 + \frac{[I_1]}{[N]} + \frac{[I_2]}{[N]} + \frac{[D]}{[N]}}. \quad (35)$$

Because unfolding occurs under equilibrium conditions, we can write the following equilibrium equations:

$$K_1(T) = \frac{[I_1]}{[N]} \quad K_2(T) = \frac{[I_2]}{[I_1]} \quad K_3(T) = \frac{[D]}{[I_2]}. \quad (36)$$

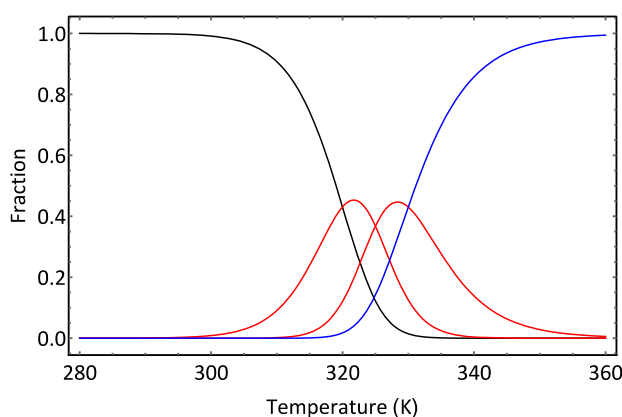


Fig. 9 Graph showing how the fraction of protein species varies with temperature using the model data in Table 3

Using these expressions in Eq. 35, we obtain

$$\alpha_N(T) = \frac{1}{1 + K_1(T) + K_1(T)K_2(T) + K_1(T)K_2(T)K_3(T)}. \quad (37)$$

Similar expression is readily derived for the fractions of the other intermediate and denatured species:

$$\begin{aligned} \alpha_1(T) &= K_1(T)\alpha_N(T) \\ \alpha_2(T) &= K_1(T)K_2(T)\alpha_N(T) \\ \alpha_D(T) &= K_1(T)K_2(T)K_3(T)\alpha_N(T). \end{aligned} \quad (38)$$

The equilibrium constants are calculated using Eq. 21. Given the following model thermodynamic data, the fractional composition of an aqueous protein solution is depicted in Fig. 9.

Simulating the HSDSC signal using the parameters in Tables 3 and 4 is slightly more complicated than our previous examples. The greater complexity comes from correctly identifying the enthalpy changes that accompany the formation of each species. Essentially, all enthalpy changes are calculated with the native form as the low energy form of the protein. Thus, the enthalpy change accompanying the formation of I_1 is $\Delta H_{VH,1}$; the enthalpy change accompanying the formation of I_2 is $\Delta H_{VH,1} + \Delta H_{VH,2}$; and the enthalpy of denaturation is given by $\Delta H_{VH,1} + \Delta H_{VH,2} + \Delta H_{VH,3}$. Thus, the following expression can be then be used to calculate the excess heat capacity assuming $\Delta H_{cal} = \Delta H_{vH}$:

$$\begin{aligned} C_{p,XS} &= \Delta H_{VH,1}(T) \frac{d\alpha_1(T)}{dT} + \alpha_1(T)\Delta C_{P,1} + \\ &\dots(\Delta H_{VH,1}(T) + \Delta H_{VH,2}(T)) \frac{d\alpha_2(T)}{dT} + \alpha_2(T)(\Delta C_{P,1} + \Delta C_{P,2}) + \\ &\dots(\Delta H_{VH,1}(T) + \Delta H_{VH,2}(T) + \Delta H_{VH,3}(T)) \frac{d\alpha_3(T)}{dT} + \alpha_3(T) \\ &\quad (\Delta C_{P,1} + \Delta C_{P,2} + \Delta C_{P,3}). \end{aligned} \quad (39)$$

Table 3 Model data used to show the effect of temperature upon the composition of an aqueous protein solution using Eqs. 21, 37, and 38

Transition	ΔH_{vH}	T_{ref}	ΔC_p
1	190	320	0.8
2	220	325	1.2
3	170	330	0.6

Table 4 Parameter values used to simulate the HSDSC signal shown in Fig. 12

Parameter	Assumed value
$\Delta H_{vH,\text{ref}}$	300 kJ mol ⁻¹
$\Delta H_{\text{cal},\text{ref}}$	295 kJ mol ⁻¹
T_{ref}	330 K
a	-30 kJ mol ⁻¹ K ⁻¹
b	0.11 kJ mol ⁻¹ K ⁻¹
b	9.0 kJ mol ⁻¹ K ⁻¹
d	0.01 kJ mol ⁻¹ K ⁻¹

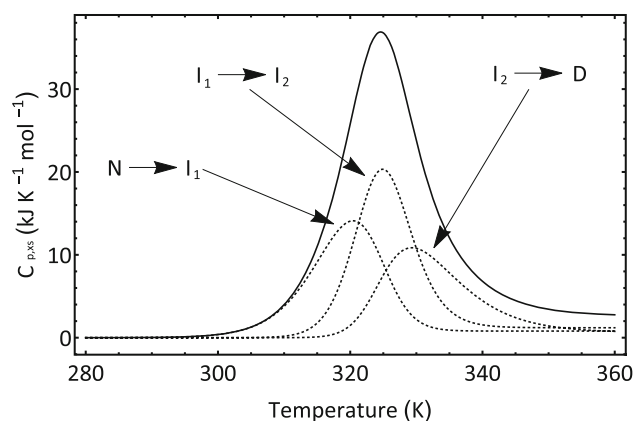
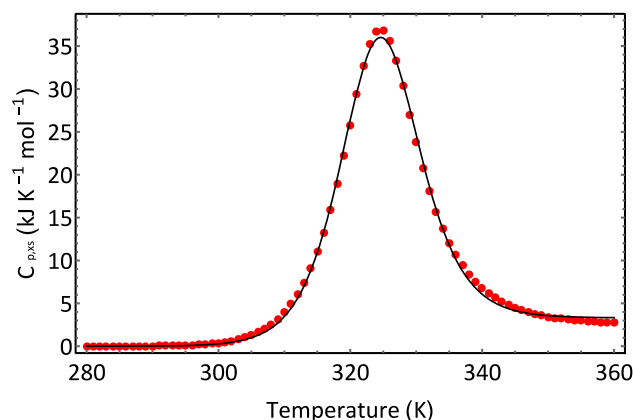
If we collect together the appropriate terms, then Eq. 39 can be written in terms of the underlying component transitions as follows:

$$\begin{aligned}
 C_{p,rs} = & \Delta H_{vH,1}(T) \left(\frac{d\alpha_1(T)}{dT} + \frac{d\alpha_2(T)}{dT} + \frac{d\alpha_3(T)}{dT} \right) \\
 & + (\alpha_1(T) + \alpha_2(T) + \alpha_3(T)) \Delta C_{p,1} + \\
 & \dots \Delta H_{vH,2}(T) \left(\frac{d\alpha_2(T)}{dT} + \frac{d\alpha_3(T)}{dT} \right) + (\alpha_2(T) \\
 & + \alpha_3(T)) \Delta C_{p,2} + \\
 & \dots \Delta H_{vH,3}(T) \frac{d\alpha_3(T)}{dT} + \alpha_3(T) \Delta C_{p,3}.
 \end{aligned} \quad (40)$$

The derivatives in Eqs. 39 and 40 can be estimated using a centred finite difference approximation:

$$\frac{d\alpha(T)}{dT} \approx \frac{\alpha(T + \delta T) - \alpha(T - \delta T)}{2\delta T}. \quad (41)$$

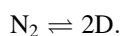
The simulated DSC signal using the data in Table 3 and Eqs. 21, 37, 38, 39, 40, and 41 is shown in Fig. 10. It is worth noting that the shapes of the component transitions are not symmetrical. The overall thermal transition can be fitted to a two-state model as was the data obtained for ubiquitin. This is shown in Fig. 11. The fit is not especially poor and could lead inexperienced experimenters to conclude that the transition is two states. However, the optimized fit parameters show immediately that the use of the two-state model is incorrect. The van't Hoff enthalpy value is 218 kJ mol⁻¹, whilst the calorimetric enthalpy is 548 kJ mol⁻¹. As we pointed previously in the manuscript, if $\Delta H_{\text{cal}} > \Delta H_{vH}$, then the presence of significant populations of intermediate states in the transition is inferred.

**Fig. 10** Simulated DSC signal for a model protein system wherein unfolding involves the formation of two intermediates. The component transitions are identified and shown. A value of was used in the simulation**Fig. 11** Fit of the overall simulated signal shown in Fig. 10 to a two-state model

Transitions involving dissociation

Many proteins have a quaternary structure which involves the association of several folded molecular sub-units to form a multiple sub-unit complex. The simplest complex is a dimer. One such example, examined by Sturtevant and co-workers, is the tryptophan repressor obtained from *Escherichia Coli*, which shows unusual thermal stability at pH 7.5 [8]. Their DSC traces show that the heat capacity change is temperature dependent and that the peak itself shows a significant amount of asymmetry. Furthermore, the signal shows some concentration dependence. To develop a thermodynamic model that can encapsulate these observations, we need to use Eq. 33 that incorporates the temperature effects upon heat capacity. However, we need to be extremely careful about the equilibrium constant equations that we use. We shall find it expedient to formulate these equations in terms of the fraction of protein that has undergone dissociation/denaturation.

We assume that the following equilibrium equation adequately describes the dissociation/unfolding process:



In other words, dissociation and unfolding occur at more or less the same time. Or to use the language, we used in the previous section on multiple transitions the population of dissociated sub-units in the native form is extremely if not vanishingly small.

To make the simulations simple, we shall assume that both sub-units comprising the dimer are the same, so that we can use the following mass action description of the equilibrium process:

$$K_p(T) = \frac{[D]^2}{[N_2]}.$$

If P_t is the total concentration of the dimeric protein, then we can write the following:

$$[D] = 2f_D P_t. \quad (42)$$

Here, f_D is the fraction of dimer that has undergone dissociation/denaturation. Similarly

$$[N_2] = (1 - f_D) P_t. \quad (43)$$

The equilibrium constant is thus written as

$$K_p(T) = \frac{[D]^2}{[N_2]} = \frac{4f_D^2 P_t}{(1 - f_D)}. \quad (44)$$

We now need to define $K(T_{\text{ref}})$ T_{ref} is the temperature at which half the protein has undergone dissociation/denaturation, i.e., $f_D = 0.5$. We shall, however, also define T_{ref} in terms of a reference concentration P_{ref} :

$$K_p(T_{\text{ref}}) = \frac{4 \times 0.5^2 \times P_{\text{ref}}}{(1 - 0.5)} = 2P_{\text{ref}}. \quad (45)$$

Substituting Eqs. 44 and 45 into Eq. 33 gives

$$\frac{2f_D^2}{(1 - f_D)} \frac{P_t}{P_{\text{ref}}} = \exp \left[\frac{\Delta H_{VH,\text{ref}}}{R} \left(\frac{1}{T_{\text{ref}}} - \frac{1}{T} \right) + \left(\frac{c - a}{R} \right) \left(\ln \frac{T}{T_{\text{ref}}} \right) + \frac{T_{\text{ref}}}{T} - 1 \right] \dots + \frac{(d - b)}{2R} \left(T - T_{\text{ref}} - T_{\text{ref}}^2 \left(\frac{1}{T_{\text{ref}}} - \frac{1}{T} \right) \right) \quad (46)$$

Equation 46 is a quadratic expression in terms of f_D . As before, if we define all the parameters on the right-hand side (the values are shown below) as well as the concentrations, we can calculate f_D using the normal solution for quadratic equations. These values are then used in Eq. 29 to simulate the HSDSC signal.

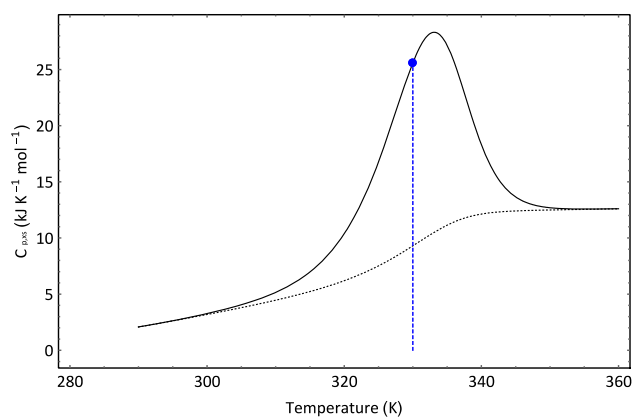


Fig. 12 Simulation of dissociative unfolding of a dimer protein complex. The blue line shows the location of T_{ref} the temperature at which half the protein has undergone the thermal transition

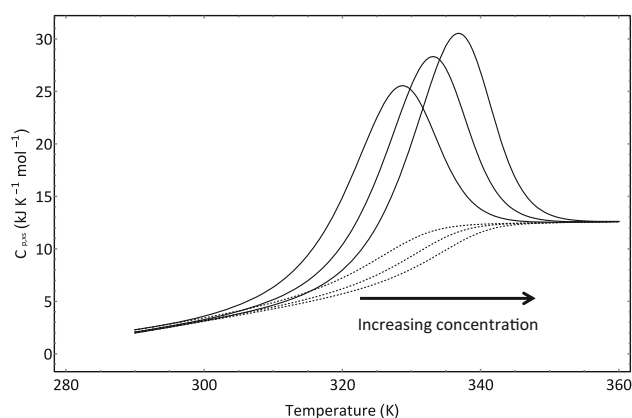


Fig. 13 Simulation showing the effect of concentration upon thermal transition shown in Fig. 12

For the simulation of the HSDSC signal shown in Fig. 12, the data displayed in Table 4 was used and it was assumed that $P_t = P_{\text{ref}}$. Comparison between the data provided by Sturtevant et al. [8] and Fig. 13 shows that the simulation captures the major features of the experimental data. The heat capacity is temperature dependent, and the signal shows distinct asymmetry. Moreover, in Figs. 12 and 13, it is shown that T_{ref} does not correspond to the temperature of maximum excess heat capacity. If we change the protein concentration, then we can show that the signal shifts to higher temperature ranges when the concentration is increased and to lower temperature ranges when the concentration is lowered in line with experimental observations [8]. The observant reader will no doubt detect that the signals appear larger at higher concentrations. This is to be expected, since the transitions occur over higher temperature ranges at higher protein concentrations and the positive heat capacity change thus results in increases in both the calorimetric and van't Hoff enthalpies.

Concluding remarks

The aim of this article has been to examine the thermodynamics of temperature induced changes in aqueous protein systems as detected by scanning calorimetry. I have tried to show how through the use of simple models of protein unfolding and through the application of familiar thermodynamic relationships, the scanning calorimetric signals can be simulated and fitted to these models. The text, however, does come with a caveat. Calorimetric signals can be, and very often are, over interpreted. The model selected must fit the known attributes of the thermally induced transition. It is not uncommon to see novices try to fit a dissociation transition (that always shows a distinct asymmetric peak) to a model that involves several independent transitions using the software supplied by the instrument manufacturer. The better the novice understands, the underpinning science of signal creation the more likely they shall be able to correctly interpret that signal.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Sturtevant JM (1987) Biochemical applications of differential scanning calorimetry. *Annu Rev Phys Chem* 38:463–488
2. Mabrey S, Sturtevant JM (1976) Investigation of phase transitions of lipids and lipid mixtures by sensitivity differential scanning calorimetry. *Proc Natl Acad Sci* 73(11):3862–3866
3. Spink CH (2008) Differential scanning calorimetry. In: Correia JJ, Detrich HW (eds) *Methods in cell biology*, vol 84. Academic Press, Cambridge, MA, pp 115–141
4. Dill KA, Bromberg S (2011) *Molecular driving forces: statistical thermodynamics in biology, chemistry, physics and nanoscience*, 2nd edn. Garland Science, New York
5. Cooper A (2010) Protein heat capacity: an anomaly that maybe never was. *J Phys Chem Lett* 1(22):3298–3304. doi:[10.1021/jz1012142](https://doi.org/10.1021/jz1012142)
6. Connelly PR, Thomson JA (1992) Heat capacity changes and hydrophobic interactions in the binding of FK506 and rapamycin to the FK506 binding protein. *Proc Natl Acad Sci USA* 89(11):4781–4785
7. Sanchez-Ruiz JM, Lopez-Lacombe JL, Cortijo M, Mateo PL (1988) Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin. *Biochemistry* 27(5):1648–1652
8. Bae SJ, Chou WY, Matthews K, Sturtevant JM (1988) Tryptophan repressor of escherichia-coli shows unusual thermal-stability. *Proc Natl Acad Sci USA* 85(18):6731–6732