

Protein unfolding. Thermodynamic perspectives and unfolding models.

Joachim Seelig*¹ and Anna Seelig

¹ Biozentrum, University of Basel, Spitalstrasse 41, CH-4056 Basel, Switzerland

*corresponding author: e-mail: joachim.seelig@unibas.ch; Tel: +41 61 207 2190

Abstract

Protein unfolding is a dynamic cooperative process with many short-lived intermediates. Cooperativity means that similar molecular elements act dependently on each other. The thermodynamics of protein unfolding can be determined with differential scanning calorimetry (DSC). The measurement of the heat capacity provides the temperature profiles of enthalpy, entropy and free energy. The thermodynamics of protein unfolding is completely determined with these thermodynamic properties. We emphasise the model-independent analysis of the heat capacity. The temperature profiles of enthalpy $H(T)$, entropy $S(T)$ and free energy $G(T)$ can be obtained directly by a numerical integration of $C_p(T)$. In evaluating different models for protein unfolding. It is essential to simulate all thermodynamic properties, not only the heat capacity. A chemical equilibrium two-state model is a widely used approximation to protein unfolding. The model assumes a chemical equilibrium between only two protein conformations, the native protein (N) and the unfolded protein (U). The model fits the heat capacity $C_p(T)$ quite well, but fails in simulating the other thermodynamic

properties. In this review we propose a modification of the chemical equilibrium two-state model, which removes these inconsistencies. We also propose a new statistical-mechanical two-state model based on a simple, two-parameter partition function $Z(T)$, from which all thermodynamic parameters can be derived. The thermodynamic predictions of the new models are compared to published DSC-experiments obtained with lysozyme, a globular protein, and β -lactoglobulin, a β -barrel protein. Good fits to all thermodynamic properties are obtained. In particular, the models predict a zero free energy for the native protein, which is confirmed experimentally by DSC. This is in contrast to the often-cited chemical equilibrium two-state model, which predict a positive free energy for the native protein. Two-state models use macroscopic fit parameters, the conformational enthalpy and the heat capacity difference between native and unfolded protein. These simulations provide no molecular insight. The review therefore includes a recently published multistate cooperative model based on physically well-defined molecular parameters only.

Key words: Protein unfolding, differential scanning calorimetry, Zimm-Bragg theory, cold denaturation, free energy

1. Introduction

Many proteins are only marginally stable at room temperature and can be denatured by heating or cooling. The analysis of protein stability is thus an important problem in developing biological therapeutics. Protein unfolding is a cooperative process with many short-lived intermediates. Cooperativity means that similar molecular elements (e.g. amino acid residues) act dependently on each other. Despite the relevance of cooperativity, protein heat- and cold-denaturation was so far analyzed almost exclusively with a chemical two-state equilibrium [1-9]. This model considers only two types of protein conformations in solution, the native protein (N) and the fully unfolded protein (U). No molecular interactions are specified in a two-state model, which therefore must be classified as non-cooperative. The often-cited chemical equilibrium two-state model as presented, for example, in reference [10] fits the heat capacity quite well. However, it is not consistent with the results obtained with differential scanning calorimetry (DSC) regarding the temperature profiles of enthalpy, entropy and free energy. In this review we therefore present new two-state models, which avoid these difficulties. Better agreement with DSC is obtained with an empirical modified chemical equilibrium two-state model, in which all thermodynamic functions are multiplied by the extent of unfolding ($\Theta_U(T)$ -weighted chemical equilibrium two-state model). And even more precise fit of the DSC data is obtained with a completely different statistical-mechanical two-state model, which is based on rigorous statistical thermodynamics.

The fundamental parameter of all thermodynamic properties of protein unfolding is the heat capacity $C_p(T)$, which can be measured with differential scanning calorimetry. In this review we first show how the thermodynamic functions enthalpy $H(T)$, entropy $S(T)$ and free energy, $G(T)$ can be obtained by numerical integration of $C_p(T)$ without the application of any protein unfolding model. We then discuss the standard and the $\Theta_U(T)$ -weighted version of the chemical equilibrium two-state model. Finally, we propose two new statistical-mechanical models based on rigorous thermodynamic partition functions. The models are compared to

DSC measurements of lysozyme and β -lactoglobulin. We emphasise again the model-independent analysis of heat capacity in terms of enthalpy, entropy and free energy. The experimental data are then used to test the different models. The statistical-mechanical models are based on rigorous thermodynamics and provide perfect simulations of all measured thermodynamic properties. The chemical equilibrium two-state model shows discrepancies with respect to entropy and free energy and casts doubt on the physical reality of the postulated positive free energy of the native protein. A modified chemical equilibrium two state model is proposed which corrects most of the insufficiencies. A historical perspective of the chemical equilibrium two-state model can be found in references [7, 8].

2. Method – Differential scanning calorimetry (DSC). Model-independent thermodynamic analysis of protein unfolding experiments

Differential scanning calorimetry (DSC) is the method of choice to study the thermodynamic properties of protein unfolding. Details of the DSC method can be found in references [3, 4, 9, 11]. DSC measures the heat capacity $C_p(T)$ and the relevant literature is focused almost exclusively on the simulation of the heat capacity peak associated with protein unfolding. However, DSC can do more. By numerical integration of the heat capacity the fundamental thermodynamic properties of protein unfolding can be derived. that is [12, 13]

$$\text{enthalpy} \quad \Delta H_{\text{DSC}}(T_i) = \sum_1^i \left[\frac{C_p(T_i) + C_p(T_{i+1})}{2} \right] [T_{i+1} - T_i] \quad (1)$$

$$\text{entropy} \quad \Delta S_{\text{DSC}}(T_i) = \sum_1^i \frac{C_p(T_{i+1}) + C_p(T_i)}{2T_i} [T_{i+1} - T_i] \quad (2)$$

$$\text{free energy} \quad \Delta G_{\text{DSC}}(T_i) = H(T_i) - T_i S(T_i) \quad (3)$$

Note that all thermodynamic properties can be evaluated without resorting to a particular unfolding model.

Native proteins have a substantial heat capacity.[14] After unfolding, the heat capacity is even larger. The increase ΔC_p^0 , is caused essentially by the binding of additional water molecules [15].

A typical DSC-thermogram is shown in figure 1A. $C_p(T)$ starts out almost linearly with the basic heat capacity of the native protein. Unfolding gives rise to a sharp heat capacity peak, which is followed by a region of again rather constant heat capacity of the unfolded protein. The choice of the DSC-baseline is important and is handled quite differently in the literature.[9] The subtraction of a sigmoidal baseline is quite common such that native and denatured protein have zero heat capacities [16]. With this correction the further analysis is limited to the conformational change proper. The increase in the basic heat capacity ΔC_p^0 of the unfolded protein is ignored. However, this has been criticized as “it is clear that in considering the energetic characteristics of protein unfolding one has to take into account all energy which is accumulated upon heating and not only the very substantial heat effect associated with gross conformational transitions, that is, all the excess heat effects must be integrated” [11].

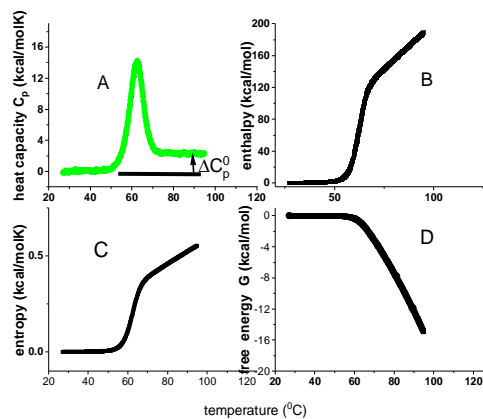


Figure 1. DSC of lysozyme. Model-independent thermodynamic analysis (50 μ M, 20% glycine buffer, pH 2.5). (A) Heat capacity. DSC data (temperature resolution 0.17 $^{\circ}$ C) taken from

reference [17, 18]. (B) Enthalpy $\Delta H_{DSC}(T)$ (eq.1). (C) Entropy $\Delta S_{DSC}(T)$ (eq.2). (D) Gibbs free energy $\Delta G_{DSC}(T)$ (eq.3).

The analysis of DSC experiments as performed in this review is shown in figure 1 for the thermal unfolding of lysozyme [17, 18]. Lysozyme is a 129-residue protein composed of ~25% α -helix, ~40% β -structure and ~35% random coil in solution at room temperature [17]. Upon unfolding, the α -helix is almost completely lost and the random coil content increases to ~60%. Thermal unfolding occurs in the range of $43\text{ }^{\circ}\text{C} < T < 73\text{ }^{\circ}\text{C}$ and is completely reversible. Lysozyme is the classical example to demonstrate two-state unfolding [6, 10, 19].

Figure 1A displays the heat capacity $C_p(T)$ [17, 18]. The midpoint of unfolding is at $T_m = 62\text{ }^{\circ}\text{C}$. Panels 1B -1D show the summation of $C_p(T_i)$ according to equations 1 – 3. Due to baseline correction, the basic heat capacity of the native lysozyme is removed and the figure shows the heat capacity of the unfolding transition proper. The heat capacity is hence zero for the native protein according to basic thermodynamics it follows that all thermodynamic properties must also be zero for $C_p = 0\text{ cal/molK}$.

As shown in figure 1A, the heat capacity $C_p(T_i)$ is a non-linear function of temperature. Consequently, enthalpy $H(T) = \int C_p(T)dT$, entropy $S(T) = \int \frac{C_p(T)}{T} dT$ and Gibbs free energy $G(T) = H(T) - TS(T)$ are also non-linear in temperature T . Indeed, enthalpy and entropy display sigmoidal temperature profiles (figs. 1B,1C). The free energy of lysozyme is zero for the native protein, is slightly negative up to the midpoint temperature T_m , and decreases rapidly beyond T_m (fig. 1D).

Figure 1A is typical for the heat capacity profiles of protein unfolding. Surprisingly, the relevant literature reports only heat capacity measurements, but not the corresponding temperature profiles of enthalpy, entropy or free energy, even though these thermodynamic functions become essential in a model-guided analysis.

Unfolding models generally assume baseline-corrected thermograms with a zero heat capacity for the native protein. Equations, 1-3 are not limited to baseline corrected thermograms. Evaluations which include the substantial heat capacity of the native protein found in a recent publication on cooperative protein unfolding [12].

3. Theory. Models for protein unfolding

3.1. Chemical equilibrium two-state models

Protein unfolding is a cooperative process. Nevertheless, in spite of many short-lived intermediates, protein unfolding is almost exclusively described by a chemical equilibrium between a single native protein conformation (N) and a single denatured molecule (U). The temperature-dependent equilibrium constant is defined as

$$K_{\text{NU}}(T) = \frac{[U]}{[N]} \quad (4)$$

[U] and [N] denote the concentrations of unfolded and native protein, respectively. The temperature dependence of $K_{\text{NU}}(T)$ is handled differently in different models.

3.1.1. Van't Hoff enthalpy model

The early version of the two-state model is based on van't Hoff's law.[1, 3, 4] The temperature dependence of the chemical equilibrium constant is given by

$$\frac{\partial \ln K_{\text{NU}}(T)}{\partial T} = \frac{\Delta H_0}{RT^2} \quad (5)$$

ΔH_0 is the conformational enthalpy. Integration of equation 5 yields

$$\ln K_{\text{NU}}(T) = (-\Delta H_0 / RT) + C$$

$$K_{\text{NU}}(T) = e^{\frac{-\Delta H_0}{R} \left(\frac{1}{T} - \frac{1}{T_m} \right)} \quad (6)$$

The integration constant C was chosen such that the equilibrium constant is unity at the midpoint temperature of unfolding T_m , that is, $K_{NU}(T_m) = 1$. At T_m , native and unfolded protein occur at equal concentrations.

In the van't Hoff model the enthalpy is temperature-independent and the unfolded protein has the same basic heat capacity as the native protein.[3, 4, 20]. To account for the experimentally observed increase in the heat capacity of the unfolded protein, the van't Hoff model was replaced by a more general model with a temperature-dependent enthalpy.

3.1.2. Free energy chemical equilibrium two-state model (“standard model”)

This model is based on the free energy [21, 22]. We follow the common nomenclature [11, 22, 23]. The temperature dependence of the $N \rightleftharpoons U$ equilibrium is calculated with the free energy $\Delta G_{NU}(T)$.

$$K_{NU}(T) = \frac{[U]}{[N]} = e^{\frac{-\Delta G_{NU}(T)}{RT}} \quad (7)$$

The free energy $\Delta G_{NU}(T)$ is composed of the enthalpy $\Delta H_{NU}(T)$ and the entropy $\Delta S_{NU}(T)$.

$$\Delta H_{NU}(T) = \Delta H_0 + \Delta C_p^0(T - T_m) \quad (8)$$

ΔH_0 is the conformational enthalpy proper and ΔC_p^0 is the increase in heat capacity between the native and the unfolded protein. The entropy $S_{NU}(T)$ is defined as

$$\Delta S_{NU}(T) = \Delta S_0 + \Delta C_p^0 \ln \frac{T}{T_m} = \frac{\Delta H_0}{T_m} + \Delta C_p^0 \ln \frac{T}{T_m} \quad (9)$$

The conformational entropy ΔS_0 is evaluated by assuming a first-order phase transition (e.g. melting of ice). In such a phase transition the total heat ΔH_0 is absorbed at a constant temperature T_m and the entropy change is $\Delta S_0 = \Delta H_0/T_m$. With the assumption the free energy follows as

$$\Delta G_{\text{NU}}(T) = -RT \ln K_{\text{NU}}(T) = \Delta H_0 \left(1 - \frac{T}{T_m}\right) + \Delta C_p^0 (T - T_m) - T \Delta C_p^0 \ln\left(\frac{T}{T_m}\right) \quad (10)$$

It should be noted however, that in protein unfolding ΔH_0 is absorbed not at a constant temperature but over a temperature range of 20 - 50 °C.

The extent of unfolding $\Theta_U(T)$ is

$$\Theta_U(T) = \frac{K_{\text{NU}}(T)}{1 + K_{\text{NU}}(T)} = \frac{e^{-\frac{\Delta G_{\text{NU}}(T)}{RT}}}{1 + e^{-\frac{\Delta G_{\text{NU}}(T)}{RT}}} \quad (11)$$

The thermodynamic definition of the heat capacity is $C_p(T) = (\partial H(T) / \partial T)_p$. The heat capacity of the chemical equilibrium two-state model is

$$C_p(T) = \frac{\partial(\Theta_U(T) \Delta H_{\text{NU}}(T))}{\partial T} = \Delta H_{\text{NU}}(T) \frac{\partial \Theta_U(T)}{\partial T} + \Theta_U(T) \Delta C_p^0 \quad (12)$$

It should be noted that the unfolding enthalpy $\Delta H_{\text{NU}}(T)$ is multiplied with the extent of unfolding $\Theta_U(T)$ to account for the non-linear temperature profile of the heat capacity $C_p(T)$. Equation 12 is identical to equation 14 in reference [22] and is the hallmark of the standard chemical equilibrium two-state model.

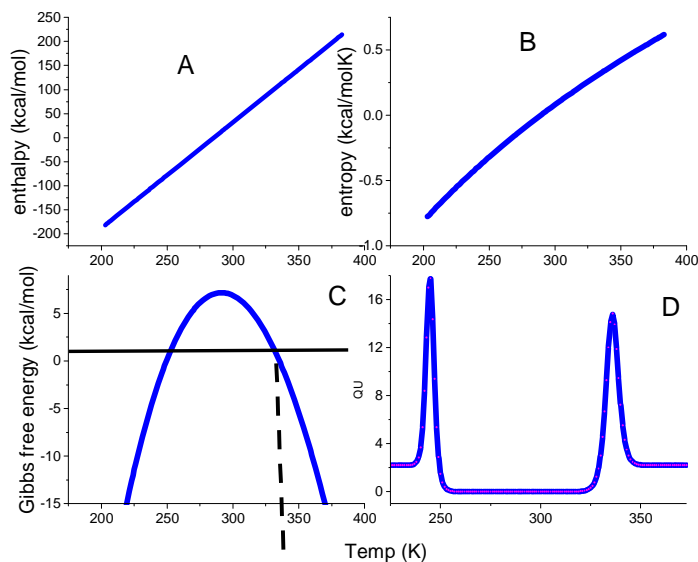


Figure 2. The thermodynamic functions of the standard chemical equilibrium two-state model calculated with $\Delta H_0 = 110$ kcal/mol and $\Delta C_p^0 = 2.2$ kcal/molK. Dashed vertical line: midpoint temperature $T_m = 62^\circ\text{C}$. (A) Enthalpy $\Delta H_{NU}(T)$. Eq. 8. (B) Entropy $\Delta S_{NU}(T)$. Eq. 9. (C) Gibbs free energy $\Delta G_{NU}(T)$. Eq. 10. (D) Heat capacity $C_p(T)$. Eq. 12.but

The thermodynamic predictions of this model are shown in figure 2. The enthalpy is a linear function of temperature, the entropy is almost linear, and the free energy $\Delta G(T)$ has the approximate shape of an inverted parabola. The native protein has a free energy maximum of 7.51 kcal/mol. At 290 K = 17 °C. However, as the heat capacity is zero at the same temperature this reveals a thermodynamic inconsistency. As shown experimentally in figure 1, a zero heat capacity leads to zero values for all thermodynamic functions. The temperatures for heat and cold unfolding are $T_m = 63$ °C and $T_{\text{cold}} = -24$ °C, respectively. At these temperatures, folded and unfolded protein have the equal concentrations. Cold denaturation may not be feasible experimentally but T_{cold} can be estimated as

$$T_{\text{cold}} \approx T_m \left(2e^{\frac{-\Delta H_0}{T_m \Delta C_p^0}} - 1 \right) \quad (13)$$

The temperature difference between heat and cold denaturation is

$$\Delta T \approx 2T_m \left(1 - e^{\frac{-\Delta H_0}{T_m \Delta C_p^0}} \right) \quad (14)$$

ΔT depends essentially on the ratio $\Delta H_0 / \Delta C_p^0$. The two parameters have opposite effects. ΔH_0 increases ΔT , ΔC_p^0 decreases it.

3.1.3. $\Theta_U(T)$ -weighted chemical equilibrium two-state model [24]

The heat capacity of lysozyme in figure 1A shows a non-linear temperature profile. This is taken into account in the chemical equilibrium model by differentiating not $\Delta H_{NU}(T)$ (which would result in a constant heat capacity, eq. 8), but $\Delta H_{NU}(T)\Theta_U(T)$. As $C_p(T)$ is non-

linear other thermodynamic functions must also be non-linear. The solution is to extend the empirical approach of equation 12 to entropy and free energy. We therefore define a new set of $\Theta_U(T)$ -weighted thermodynamic functions.

$$\Delta H_{\Theta}(T) = \Theta_U(T)\Delta H_{NU}(T) \quad (15)$$

$$\Delta S_{\Theta}(T) = \Delta S_{NU}(T)\Theta_U(T) \quad (16)$$

$$\Delta G_{\Theta}(T) = \Delta G_{NU}(T)\Theta_U(T) \quad (17)$$

The heat capacity is equation 12 and is not repeated here. The resulting temperature profiles are shown in figure 3.

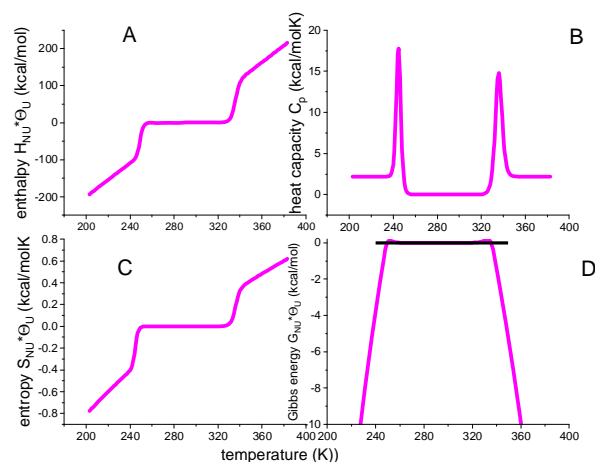


Figure 3. $\Theta_U(T)$ -weighted chemical equilibrium two-state model. (A) Enthalpy (eq. 15).

(B). Heat capacity (eq.12) (C) Entropy (eq. 16). (D) Free energy (eq. 17).. Fit parameters:

$$\Delta H_0 = 107 \text{ kcal/mol}, \Delta C_p^0 = 2.27 \text{ kcal/molK}$$

The weighting factor $\Theta_U(T)$ generates sigmoidal temperature profiles for enthalpy and entropy and a trapezoidal profile for the free energy. $\Delta G_{\Theta}(T)$ of the native protein is zero, which is now consistent with the zero heat capacity. Nevertheless, $\Theta_U(T)$ is an empirical weighting factor, not based on solid thermodynamic reasoning. Indeed, closer inspection of equation 17

reveals residual small positive free energies in the vicinity of the midpoint temperatures T_m and T_{cold} (fig. 3D). These are not confirmed by DSC experiment (cf. fig. 7).

3.2. Statistical-mechanical models

3.2.1. Partition function $Z(T)$ and thermodynamic properties

Statistical-mechanics provides a rigorous thermodynamic approach to protein unfolding. The heat capacity $C_p(T)$ is intimately related to the partition function $Z(T)$. Knowledge of the partition function $Z(T)$ then leads to the following thermodynamic relations [25, 26]

$$\text{Helmholtz free energy: } F(T) = -RT \ln Z(T) \quad (18)$$

$$\text{Inner energy: } E(T) = RT^2 \frac{\partial \ln Z(T)}{\partial T} \quad (19)$$

$$\text{Entropy: } S_v(T) = \frac{E(T) - F(T)}{T} \quad (20)$$

$$\text{Heat capacity: } C_v(T) = \left(\frac{\partial E(T)}{\partial T} \right)_v = \frac{\langle E(T)^2 \rangle - \langle E(T) \rangle^2}{RT^2} \quad (21)$$

3.2.2. Statistical-mechanical 2-state model [24]

The problem is to find the partition function of a one-component two-state system. Based on the statistics of the linear Ising model as described in reference[27], the following continuous canonical partition function can be defined

$$Z(T) = \left(1 + e^{\frac{-[\Delta E_0 + C_v(T-T_m)] \left[\frac{1}{T} - \frac{1}{T_m} \right]}{R}} \right) \quad (22)$$

ΔE_0 is the conformational energy of the unfolded protein. It is temperature-dependent with the heat capacity C_v . For convenience the partition function is simplified by introducing

$$Q(T) = e^{\frac{-[\Delta E_0 + C_v(T-T_m)] \left(\frac{1}{T} - \frac{1}{T_m} \right)}{R}} \quad (23)$$

leading to

$$Z(T) = (1 + Q(T)) \quad (24)$$

The fraction of unfolded protein $\Theta_S(T)$ is

$$\Theta_S(T) = \frac{\partial \ln Z(T)}{\partial \ln Q} = \frac{Q(T)}{1 + Q(T)} \quad (25)$$

$\Theta_S(T)$ is included for completeness only. It is not needed to calculate thermodynamic functions. The partition function $Z(T)$ suffices to predict all thermodynamic properties. Figure 4 displays the predictions of equations 18 -22.

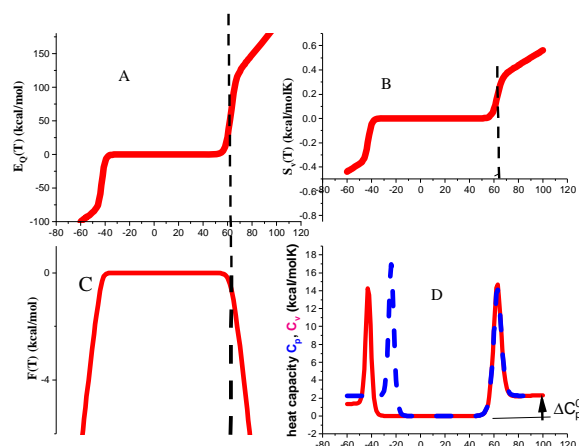


Figure 4 Statistical-mechanical two-state model. Red lines calculated with $\Delta E_0 = 110$ kcal/mol and $C_v = 1.05$ kcal/molK.. (A) Inner energy $\Delta E(T)$. (B) Entropy $\Delta S_v(T)$. (C) Helmholtz energy $\Delta F(T)$. (D) Heat capacity $C_p(T)$. Dashed blue line in panel 4D is the heat capacity calculated with the standard chemical equilibrium two-state model (eq. 12), using, $\Delta H_0 = 107$ kcal/mol, $\Delta C_p^0 = 2.27$ kcal/molK). Dashed vertical lines: midpoint temperature $T_m = 62^\circ\text{C}$.

The native protein is the reference state with all thermodynamic functions being zero. In particular, the free energy $\Delta F(T)$ is zero at 20°C and becomes slightly negative, but never

positive, between T_m and T_{cold} . $\Delta F(T)$ decreases rapidly for temperatures $T > T_m$ and $T < T_{\text{cold}}$ according to

$$\Delta F(T); (E_0 + C_v(T - T_m)) \left(1 - \frac{T}{T_m}\right). \quad (26)$$

The free energy $\Delta F(T)$ of the statistical mechanical two-state model thus also displays a trapezoidal temperature profile. Cold denaturation takes place at

$$T_{\text{cold}} = T_m - \frac{\Delta E_0}{C_v} \quad (27)$$

ΔE_0 and C_v have opposite effects on T_{cold} . Increasing ΔE_0 lowers T_{cold} , increasing C_v leads to an upward shift.

Figure 4D compares the heat capacities predicted by the statistical-mechanical two-state model and the chemical equilibrium two-state model. The high-temperature peaks of the two models overlap precisely but cold denaturation occurs at different temperatures. A discussion of other differences will follow in connection with the protein examples discussed below.

In summary, the statistical mechanical two-state model makes no assumption about the entropy. Also, as mentioned before, no weighting function $\Theta_U(T)$ is needed. to calculate correctly, the thermodynamic properties.

3.2.3. Multistate cooperative unfolding model [12]

The energy of a system with N particles is characterised by its partition function

$$Z(T) = \sum_i g_i e^{\frac{-\epsilon_i}{k_B T}} \quad (28)$$

The partition function is the sum of exponential terms (Boltzmann factor) over all energy levels ϵ_i , multiplied with their degeneracies g_i . The partition function determines the thermodynamic properties of the system (eqs. 18 -21) [25, 26]. Here we use the partition function of the multistate cooperative Zimm-Bragg theory, originally developed for the α -

helix-to-coil transition of polypeptides [28-30]. Its application to protein unfolding has been discussed recently.[12] The Zimm-Bragg theory has been applied successfully to the unfolding of helical and globular proteins of different structure and size [12, 13, 17, 18, 31-35]. Here we use[12]

$$Z(T) = (1 \quad 0) \begin{pmatrix} 1 & \sigma q(T) \\ 1 & q(T) \end{pmatrix}^N \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (29)$$

$$q(T) = e^{\frac{-h(T)}{R} \left(\frac{1}{T} - \frac{1}{T_0} \right)} \quad (30)$$

$$h(T) = h_0 + c_v(T - T_m) \quad (31)$$

h_0 is the energy change of the native \rightarrow unfolded transition of a single amino acid residue. h_0 is temperature-dependent with the heat capacity c_v . N is the number of amino acids participating in the transition. The cooperativity parameter σ determines the sharpness of the transition. The σ parameter is typically in the range of 10^{-3} - 10^{-7} . T_0 is a fit parameter to shift the position of the heat capacity peak. T_0 is usually close to T_m . The temperature difference between heat and cold denaturation is $\Delta T \approx \frac{h_0}{c_v}$.

Figure 5 shows the thermodynamic temperature profiles predicted by equation. 29 in combination with equations 18 - 21. Sigmoidal shapes are predicted for inner energy and entropy, and a trapezoidal shape for the free energy. Figure 5 is very similar to figure 4 of the statistical-mechanical two-state model, but is calculated with molecular parameters only. In particular, the free energy is again zero or negative, but never positive.

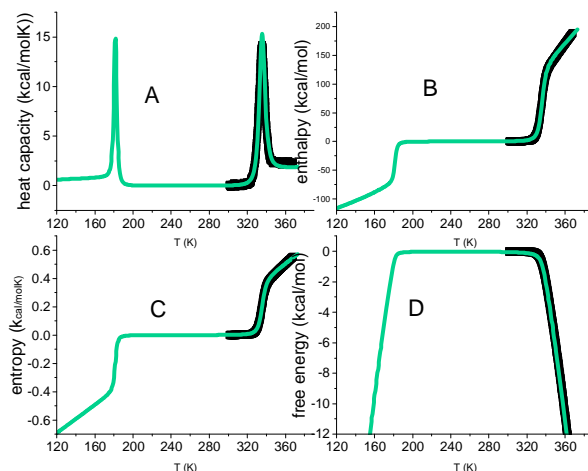


Figure 5. Multistate cooperative model. Green lines calculated with: $N = 129$ amino acid residues. $h_0 = 900$ cal/mol. $c_v = 7$ cal/molK. Cooperativity parameter $\sigma = 5 \times 10^{-7}$. Black data points: thermal unfolding of lysozyme measured with DSC. Same data as in figure 1.

Figure 5 includes the DSC data of lysozyme heat denaturation (fig. 1). An excellent agreement between theory and experiment is obtained.

The multistate cooperative model describes protein unfolding with molecular parameters of well-defined physical meaning. In contrast, two-state models provide macroscopic parameters.

Equation 29 can be applied to proteins of any size, including large antibodies with ~1200 amino acid residues and unfolding enthalpies of ~1000 kcal/mol [12].

4. Results. DSc experiments compared to protein unfolding models

4.1. Lysozyme heat unfolding

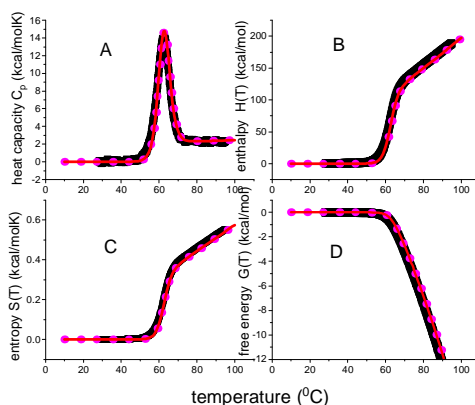


Figure 6. Analysis of lysozyme heat unfolding with 2-state models. Black data points: DSC data of figure 1. Red lines: statistical-mechanical two-state model ($\Delta E_0 = 110$ kcal/mol and $C_v = 1.05$ kcal/molK). Magenta dotted line: weighted chemical two-state model ($\Delta H_0 = 107$ kcal/mol, $\Delta C_p^0 = 2.27$ kcal/molK). (A) Heat capacity. (B) Enthalpy. (C) Entropy. (D) Free energy.

Figure 6 compares the experimental data of lysozyme unfolding with two unfolding models. The heat capacity $C_p(T)$ maximum is at $T_m = 62$ °C and the heat capacity increase is $\Delta C_p^0 = 2.27$ kcal/molK, in agreement with literature data of $\Delta C_p^0 = 1.54 - 2.27$ kcal/mol [6, 10, 15, 17, 19, 36, 37]. The red lines in figure. 6 represent the statistical-mechanical two-state model and were calculated with $\Delta E_0 = 110$ kcal/mol and $C_v = 1.05$ kcal/molK. A perfect fit of the experimental data is achieved.

The magenta dotted lines represent the weighted chemical equilibrium two-state model. The agreement with the experimental data is also very good. However, a small difference to the experimental data is observed near the midpoint of unfolding. Figure 7 displays an an

enlarged

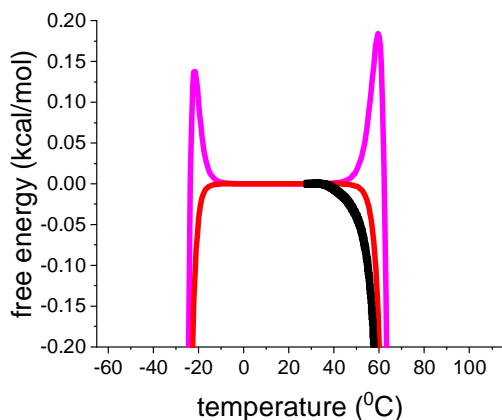


Figure 7. Enlarged view of the free energy. Black data points: DSC results for lysozyme heat unfolding. Red line: statistical mechanical two-state model ($\Delta E_0 = 110$ kcal/mol, $C_v = 1.05$ kcal/mol). Magenta line: $\Theta_U(T)$ -weighted chemical equilibrium two-state model (eq. 17, $H_0 = 1107$ kcal/mol, $\Delta C_p^0 = 2.27$ kcal/molK).

view of this region. DSC reports a zero free energy for the native lysozyme. The free energy becomes immediately negative upon unfolding. This result is correctly reproduced by the statistical-mechanical two-state model. In contrast, the $\Theta_U(T)$ -weighted chemical equilibrium two-state model (eq. 17) predicts spikes of positive free energy at temperatures just before the midpoints of unfolding. While these spikes are small and of no practical importance, they signify a thermodynamic discrepancy. Of course, the difference between experiment and model would be much larger if the parabolic free energy profile (eq. 10, fig 2C) would be included.

The multi-state cooperative model as described in figure 5 also yields a perfect fit of the lysozyme DSC experiment [12].

In summary, three different models provide a good to excellent description of lysozyme DSC unfolding. At the midpoint temperature T_m all three models predict the extent of unfolding exactly as $\Theta_U = 1/2$. Native and unfolded protein occur at equal concentrations. Nevertheless, the free energy is not zero but DSC reports $\Delta G_{DSC}(T_m) = -0.756$ kcal/mol. Indeed, a negative free energy is intuitively plausible as the protein is partially denatured at T_m . This result is supported by two theoretical models. The statistical-mechanical two-state model yields $\Delta F(T_m) = -0.462$ kcal/mol, the multistate cooperative model yields $\Delta F(T_m) = -0.855$ kcal/mol. In contrast, the $\Theta_U(T)$ -weighted chemical equilibrium two-state model predicts $\Delta G_{\Theta}(T_m) = 0$ kcal/mol.

4.2. β -lactoglobulin. Cold and heat denaturation

Bovine β -lactoglobulin (MW 18.4 kDa, 162 aa) folds up into an 8-stranded, antiparallel β -barrel with a 3-turn α -helix on the outer surface. β -Lactoglobulin in buffer without urea displays only heat denaturation (black squares in fig. 8A, data taken from fig.1 of reference [38]). Unfolding takes place between $55^\circ\text{C} < T < 96^\circ\text{C}$ with the $C_p(T)$ maximum at 78°C . The unfolding enthalpy is $\Delta H_{DSC} = 74.5$ kcal/mol (cf. [38], table 1, 0 M urea), the entropy $\Delta S_{DSC} = 0.213$ kcal/molK, and $\Delta H_{DSC}/\Delta S_{DSC} = 349\text{K} = 76^\circ\text{C}$, consistent with the $C_p(T)$ maximum.

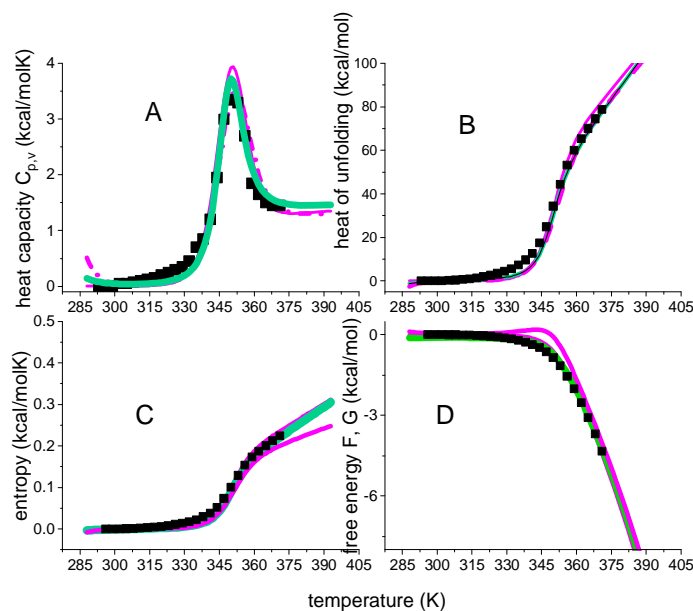


Figure 8. DSC of β -lactoglobulin in 0.1 M KCl/HCl, pH 2.0 buffer. Black data point in in panel A are taken from reference[38](fig. 1). Magenta lines: $\Theta_U(T)$ -weighted chemical equilibrium two-state model ($\Delta H_0 = 50$ kcal/mol, $\Delta C_p^0 = 1.3$ kcal/mol). Red lines: statistical-mechanical two-state model ($\Delta E_0 = 55$ kcal/mol, $C_v = 0.6$ kcal/molK) Green lines: multistate cooperative model ($h_0 = 380$ cal/mol, $c_v = 3$ cal/molK, $\sigma = 7 \times 10^{65}$, $N = 160$). (A) Heat capacity. (B) Unfolding enthalpy. (C) Unfolding entropy). (D) Free energy of unfolding.

The $\Theta_U(T)$ -weighted chemical equilibrium model, the statistical-mechanical two-state model, and the the multistate cooperative model overlap completely. as far as the heat capacity is concerned. On the other hand, some small differences are observed for enthalpy, entropy and free energy. The most conspicuous difference is seen for the free energy. The $\Theta_U(T)$ -weighted chemical equilibrium model predicts a small peak of positive free energy for the native protein, which is not confirmed by the DSC experiment.

The conformational enthalpy is $\Delta H_0 \cong \Delta E_0 = 50 - 55$ kcal/mol, which is small for a protein with 162 amino acid residues. Likewise, the molecular enthalpy parameter $h_0 = 380$

cal/mol is also small compared to the typical 900-1300 cal/mol of globular proteins.[17] The molecular origin of the small unfolding enthalpy of β -lactoglobulin can be traced back to its extensive β -structure content. The enthalpy $h_{0\beta}$ for the reaction β -structure \rightarrow disordered amino acid was measured as $h_{0\beta} = 230$ cal/mol in a membrane environment [39].

A protein can also be unfolded by cooling. Cold denaturation usually occurs at subzero temperatures but can be shifted to above 0 °C by high concentrations of chemical denaturant or extreme pH values. All models discussed above, predict cold denaturation, provided the heat capacities ΔC_p^0 or C_v are non-zero. In fact, the temperature difference between heat and cold denaturation depends strictly on the ratio of conformational enthalpy/heat capacity.

Only a few DSC experiments showing at least partial cold denaturation are available. One of the best examples is DSC-unfolding of β -lactoglobulin in 2.0 M urea solution (figure 9).

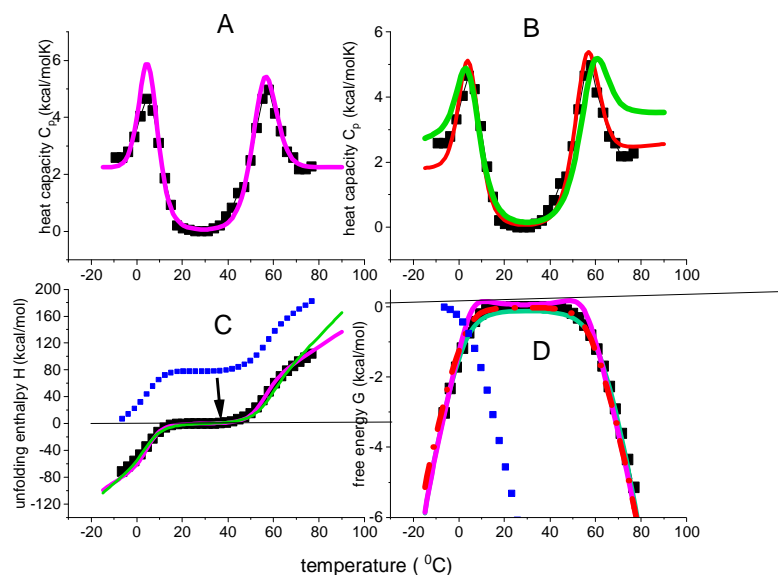


Figure 9. Thermal folding and unfolding of β -lactoglobulin in 2.0 M urea solution. Magenta lines: $\Theta_U(T)$ -weighted chemical equilibrium model. $\Delta H_0 = 56$ kcal/mol; $\Delta C_p^0 = 2.4$ kcal/mol. Red lines: statistical-mechanical two-state model. $\Delta E_0 = 55$ kcal/mol; $C_v = 1.2$ kcal/molK;

Green line: multistate cooperative model. $h_0=0.58$ kcal/mol, $c_v = 13$ cal/molK, $\sigma = 6 \times 10^{-5}$, $N = 80$. (A) DSC heat capacity data taken from reference [38]. Simulation with the $\Theta_U(T)$ -weighted chemical equilibrium model. (B) Same DSC data as in panel A. Simulations with the statistical models (C) Enthalpy $\Delta H(T)_{DSC}$. Integration of the $C_p(T)$ data according to equation 1 generates the blue data points. The data are then shifted by -78.3 kcal/mol, the enthalpy released upon cold denaturation, resulting in the black data points. This scale shift assigns a zero enthalpy to the native protein. (D) Free energy. Evaluation of $C_p(T)$ according to equations 1-3 leads to the blue data points. A related scale shifts for the entropy as for the enthalpy (not shown) and recalculation of the free energy results in the black data points. The free energy of the native protein is now zero. Cold and heat denaturation generate negative free energies (for more details see reference [24]).

The DSC data are taken from figure 2 of reference [38]. The simulation of $C_p(T)$ is shown in figure 9A for the $\Theta_U(T)$ -weighted chemical equilibrium model and in figure 9B for the two statistical models. The DSC experiment begins at -9°C where the protein is in a disordered state. Heating induces a disorder \rightarrow order transition with a heat capacity maximum at 4°C , a folding enthalpy $\Delta H_{DSC} = 78.3$ kcal/mol and an entropy change of $\Delta S_{DSC} = 0.282$ kcal/molK. The ratio $\Delta H_{DSC}/\Delta S_{DSC}$ is $277\text{K} = 4^\circ\text{C}$, consistent with the heat capacity maximum.

At $\sim 18-30^\circ\text{C}$ the protein is in the native-like conformation. Cooling reverses the before mentioned process and returns a disordered conformation with a simultaneous release -78 kcal/mol (cf. fig. 9C). The heat capacity peak at 4°C is hence the mirror image of cold denaturation (see fig. 2 in [38]). Heating β -lactoglobulin above 30°C destroys the native structure. The order \rightarrow disorder transition has a heat capacity maximum at 57°C , $\Delta H_{DSC} = 104$ kcal/mol and $\Delta S_{DSC} = 0.312$ kcal/molK. The ratio $\Delta H_{DSC}/\Delta S_{DSC}$ of heat denaturation is $333\text{K} = 60^\circ\text{C}$.

β -Lactoglobulin is less stable in urea solution as the midpoint temperature T_m is shifted from 78 °C to 57 °C. Such a decrease in temperature is common in chemical denaturants. However, it is usually associated with a decrease in enthalpy, not an increase [34, 35, 40]. In the present case the enthalpy increases by ~40%, the entropy by ~50%, resulting, in turn, in a 20°C downshift of T_m .

The DSC data were analysed with the $\Theta_U(T)$ -weighted chemical equilibrium model and the two statistical models. The conformational enthalpy ΔH_0 is equal to the inner energy ΔE_0 with $\Delta H_0 \approx \Delta E_0 = 56$ kcal/mol. These parameters are also identical to those obtained for β -lactoglobulin in buffer. In contrast, the heat capacities ΔC_p^0 , C_v and c_v are 2 – 3 times large, most likely due to the binding of urea molecules.

The three models discussed above, provide good simulation of all experimental data. In particular, they reproduce the trapezoidal temperature profile of the free energy (black squares in fig. 9D). The native protein has a zero heat capacity and, in turn, a zero free energy. Unfolding leads to negative free energies, both for heat and cold denaturation. The two statistical model exactly the produce this profile. The $\Theta_U(T)$ -weighted chemical equilibrium two-state model displays small positive peaks in the vicinity of T_m and T_{cold} , which are not supported by DSC

5. Conclusions

Under equilibrium conditions protein stability is determined by the midpoint of heat denaturation, by the temperature difference between heat and cold denaturation, and by the width and cooperativity of the unfolding transitions. These parameters are intimately connected to the thermodynamic properties of the system. The thermodynamics of protein unfolding is completely characterised by the temperature profiles of enthalpy, entropy and free energy. The building stone of these thermodynamic properties is the heat capacity $C_p(T)$,

which can be measured precisely with differential scanning calorimetry. In this review we have emphasised the almost completely ignored concept of the direct and model-independent evaluation of the heat capacity in terms of the thermodynamic functions $H(T)$, $S(T)$ and $G(T)$. The evaluation is straightforward and simple. It is hence quite surprising why this approach is not considered in the relevant literature.

Thermodynamic unfolding models should predict not only the heat capacity $C_p(T)$, but also the complete set of thermodynamic functions. The DSC experiment reveals a sigmoidal temperature profiles for enthalpy and entropy and a trapezoidal profile for the free energy. Focusing on the unfolding transition proper, the heat capacity of the native protein is zero and all thermodynamic functions are equally zero. No positive free energy is measured for the native protein, which is in contrast to the prediction of the often-cited chemical equilibrium two-state model. Experimental temperature profiles were shown for the heat-induced unfolding of the globular protein lysozyme and for the heat and cold denaturation of the β -barrel protein β -lactoglobulin. The experimental results are compared to the predictions of four different models, that is, two chemical equilibrium two-state models and two statistical mechanical models. All four models described the heat capacity equally well. The popular chemical equilibrium two state model fails however in the simulation of the thermodynamic temperature profiles. This model was therefore modified into the $\Theta_U(T)$ -weighted chemical equilibrium model by multiplying the thermodynamic functions with the extent of unfolding. This is an empirical approach which fits all thermodynamic data quite well, but displays a small discrepancy to the experimental results in the vicinity of the unfolding transition. The two statistical-mechanical models have a rigorous thermodynamic foundation and avoid this difficulty. They provide the best simulations of the experimental data.

Two-state models are non-cooperative approximations to a cooperative multistate protein folding \leftrightarrow unfolding equilibrium. They describe protein unfolding in terms of two

macroscopic parameters, the conformational enthalpy ΔH_0 ; ΔE_0 and the heat capacity ΔC_p^0 : $2C_v$. . In contrast, the multistate cooperative unfolding model uses molecular parameters, that is, the enthalpy h_0 per amino acid residue, the heat capacity c_v , the cooperativity parameter σ , and N , the number of participating amino acid residues. The multistate cooperative model can be applied to proteins of any length, e.g. antibodies with 1200 amino acid residues and unfolding enthalpies of 1000 kcal/mol.[12, 35]. In contrast, two-state unfolding models are best suited for unfolding enthalpy is of 50-200 kcal/mol, typically found for small proteins only.

Protein unfolding is characterised by large enthalpies and entropies but a small free energy (enthalpy-entropy compensation). As discussed in detail by comparing lysozyme and lactoglobulin. The free energy is not a good criterion to judge protein stability. Better parameters are defined above. The-beaded chemical equilibrium two state model, the statistical-mechanical two state model and the multistate cooperative model provide quantitative thermodynamic interpretations of these parameters.

Funding: Stiftung zur Förderung der biologischen Forschung, Basel, Switzerland

References

1. Brandts, J. F., Thermodynamics of Protein Denaturation .I. Denaturation of Chymotrypsinogen. *Journal of the American Chemical Society* **1964**, 86, (20), 4291-&.
2. Brandts, J. F., Thermodynamics of Protein Denaturation .2. Model of Reversible Denaturation + Interpretations Regarding Stability of Chymotrypsinogen. *Journal of the American Chemical Society* **1964**, 86, (20), 4302-&.
3. Jackson, W. M.; Brandts, J. F., Thermodynamics of Protein Denaturation - a Calorimetric Study of Reversible Denaturation of Chymotrypsinogen and Conclusions Regarding Accuracy of 2-State Approximation. *Biochemistry-U.S* **1970**, 9, (11), 2294-&.

4. Privalov, P. L.; Khechinashvili, N. N., Thermodynamic Approach to Problem of Stabilization of Globular Protein Structure - Calorimetric Study. *Journal of Molecular Biology* **1974**, 86, (3), 665-684.
5. Pfeil, W.; Privalov, P. L., Thermodynamic investigations of proteins. II. Calorimetric study of lysozyme denaturation by guanidine hydrochloride. *Biophys Chem* **1976**, 4, (1), 33-40.
6. Privalov, P. L.; Gill, S. J., Stability of Protein-Structure and Hydrophobic Interaction. *Advances in Protein Chemistry* **1988**, 39, 191-234.
7. Zhou, Y.; Hall, C. K.; Karplus, M., The calorimetric criterion for a two-state process revisited. *Protein science : a publication of the Protein Society* **1999**, 8, (5), 1064-74.
8. Ibarra-Molero, B.; Sanchez-Ruiz, J. M., Statistical Differential Scanning Calorimetry: Probing Protein Folding-Unfolding Ensembles. *Protein Folding, Misfolding and Aggregation: Classical Themes and Novel Approaches* **2008**, 85-105.
9. Ibarra-Molero, B.; Naganathan, A. N.; Sanchez-Ruiz, J. M.; Munoz, V., Modern analysis of protein folding by differential scanning calorimetry. In *Methods in enzymology*, Elsevier: 2016; pp 281-316.
10. Privalov, P. L., Thermodynamics of protein folding. *Journal of Chemical Thermodynamics* **1997**, 29, (4), 447-474.
11. Privalov, P. L.; Dragan, A. I., Microcalorimetry of biological macromolecules. *Biophysical chemistry* **2007**, 126, (1-3), 16-24.
12. Seelig, J.; Seelig, A., Molecular Understanding of Calorimetric Protein Unfolding Experiments. *Biophysical Reports* **2022**, 2, (March 9, 2022), 1-11.
13. Seelig, J., Free energy in thermal and chemical protein unfolding. In *Gibbs energy and Helmholtz energy: Liquids, Solutions and Vapours*, Letcher, E. W. a. T. M., Ed. Royal Society of Chemistry: London, 2022; pp 363-378.
14. Gomez, J.; Hilser, V. J.; Xie, D.; Freire, E., The heat capacity of proteins. *Proteins* **1995**, 22, (4), 404-12.
15. Myers, J. K.; Pace, C. N.; Scholtz, J. M., Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein science : a publication of the Protein Society* **1995**, 4, (10), 2138-48.
16. Durowoju, I. B.; Bhandal, K. S.; Hu, J.; Carpick, B.; Kirkitadze, M., Differential Scanning Calorimetry - A Method for Assessing the Thermal Stability and Conformation of Protein Antigen. *Jove-Journal of Visualized Experiments* **2017**, (121).
17. Seelig, J.; Schönfeld, H.-J., Thermal protein unfolding by differential scanning calorimetry and circular dichroism spectroscopy Two-state model versus sequential unfolding. *Quarterly Reviews of Biophysics* **2016**, 49, e9 (24 pages).
18. Li-Blatter, X.; Seelig, J., Thermal and Chemical Unfolding of Lysozyme. Multistate Zimm-Bragg Theory Versus Two-State Model. *J Phys Chem B* **2019**, 123, (48), 10181-10191.
19. Privalov, G.; Kavina, V.; Freire, E.; Privalov, P. L., Precise Scanning Calorimeter for Studying Thermal-Properties of Biological Macromolecules in Dilute-Solution. *Analytical Biochemistry* **1995**, 232, (1), 79-85.
20. Privalov, P. L., Stability of proteins: small globular proteins. *Adv Protein Chem* **1979**, 33, 167-241.
21. Privalov, P. L.; Griko, Y. V.; Venyaminov, S. Y.; Kutysenko, V. P., Cold Denaturation of Myoglobin. *Journal of Molecular Biology* **1986**, 190, (3), 487-498.
22. Privalov, P. L., Cold Denaturation of Proteins. *Critical Reviews in Biochemistry and Molecular Biology* **1990**, 25, (4), 281-305.
23. Becktel, W. J.; Schellman, J. A., Protein stability curves. *Biopolymers* **1987**, 26, (11), 1859-77.

24. Seelig, J.; Seelig, A., Protein unfolding. DSC and 2-state models. *J Phys Chem B* **2023**, submitted.
25. Baumann, R. P., Evaluation of Thermodynamic Properties. In *Modern Thermodynamics with Statistical Mechanics*, McConnin, R. A., Ed. Macmillan Publishing Company: New York, USA, 1992; p 341.
26. Eisenberg, D.; Crothers, D., Calculation of the energy using the system partition function. In *Physical Chemistry with Applications to the Life Sciences*, Rhamer, B.; Moore, M., Eds. The Benjamin/Cummings Publishing Company, Inc.: Menlo Park, CA, 1979; p 675.
27. Davidson, N., In *Statistical Mechanics* Mac Graw-Hill: New York, 1962; p 385.
28. Zimm, B. H.; Bragg, J. K., Theory of the Phase Transition between Helix and Random Coil in Polypeptide Chains. *Journal of Chemical Physics* **1959**, 31, (2), 526-535.
29. Zimm, B. H.; Doty, P.; Iso, K., Determination of the parameters for helix formation in poly-gamma-benzyl-L-glutamate. *Proceedings of the National Academy of Sciences of the United States of America* **1959**, 45, (11), 1601-7.
30. Zimm, B. H.; Rice, S. A., The Helix-Coil Transition in Charged Macromolecules. *Molecular Physics* **1960**, 3, (4), 391-407.
31. Zehender, F.; Ziegler, A.; Schonfeld, H. J.; Seelig, J., Thermodynamics of Protein Self-Association and Unfolding. The Case of Apolipoprotein A-I. *Biochemistry* **2012**, 51, (6), 1269-1280.
32. Schulthess, T.; Schonfeld, H. J.; Seelig, J., Thermal Unfolding of Apolipoprotein A-1. Evaluation of Methods and Models. *Biochemistry* **2015**, 54, 3063-3075.
33. Seelig, J., Cooperative protein unfolding. A statistical-mechanical model for the action of denaturants. *Biophys Chem* **2018**, 233, 19-25.
34. Eckhardt, D.; Li-Blatter, X.; Schonfeld, H. J.; Heerklotz, H.; Seelig, J., Cooperative unfolding of apolipoprotein A-1 induced by chemical denaturation. *Biophys Chem* **2018**, 240, 42-49.
35. Garidel, P.; Eiperle, A.; Blech, M.; Seelig, J., Thermal and Chemical Unfolding of a Monoclonal IgG1 Antibody: Application of the Multistate Zimm-Bragg Theory. *Biophys J* **2020**, 118, (5), 1067-1075.
36. Privalov, P. L.; Makhatadze, G. I., Heat capacity of proteins. II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects. *Journal of molecular biology* **1990**, 213, (2), 385-91.
37. Rosgen, J.; Hinz, H. J., Response functions of proteins. *Biophysical Chemistry* **2000**, 83, (1), 61-71.
38. Griko, Y. V.; Privalov, P. L., Calorimetric study of the heat and cold denaturation of beta-lactoglobulin. *Biochemistry* **1992**, 31, (37), 8810-5.
39. Meier, M.; Seelig, J., Thermodynamics of the coil \rightleftharpoons beta-sheet transition in a membrane environment. *Journal of molecular biology* **2007**, 369, (1), 277-89.
40. Hedoux, A.; Krenzlin, S.; Paccou, L.; Guinet, Y.; Flament, M. P.; Siepmann, J., Influence of urea and guanidine hydrochloride on lysozyme stability and thermal denaturation; a correlation between activity, protein dynamics and conformational changes. *Physical chemistry chemical physics : PCCP* **2010**, 12, (40), 13189-96.