ELSEVIER

Biophysical Chemistry 119 (2006) 94 – 100

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Heat capacity-independent determination of differential free energy of stability between structurally homologous proteins

David M. LeMaster *

Wadsworth Center, New York State Department of Health and Department of Biomedical Sciences, University at Albany—SUNY, Empire State Plaza, Albany, New York 12201-0509, USA

Received 17 July 2005; accepted 9 August 2005
Available online 25 August 2005

Abstract

Under the assumption of equivalent heat capacity values, the differential free energy of stability for a pair of proteins midway between their thermal unfolding transition temperatures is shown to be independent of ΔC_p up to its cubic term in ΔT_m. For model calculations reflecting the nearly 30 °C difference in T_m for the adenylate kinases from the arctic bacterium *Bacillus globisporus* and the thermophilic bacterium *Geobacillus stearothermophilus*, the resultant error in estimating $\Delta\Delta G$ by the formula 0.5 $[\Delta S_{T_{m1}}(1)+\Delta S_{T_{m2}}(2)]$ ΔT_m is less than 1%. Combined with the analogous thermal unfolding data for the adenylate kinase from *Escherichia coli*, these three homologous proteins exhibit T_m and ΔS_{T_m} values consistent with differential entropy and enthalpy contributions of equal magnitude. When entropy–enthalpy compensation holds for the differential free energy of stability, the incremental changes in T_m values are shown to be proportionate to the changes in free energy.

Keywords: Protein stability; Thermal unfolding; Heat capacity; Entropy–enthalpy compensation

1. Introduction

Accurate determination of the differential free energy of stability is required for the detailed understanding of structural evolution within protein families, in either rational design or the natural biological process. Most experimental techniques which monitor the population ratio of folded vs. unfolded protein are only accurate in the range of approximately 0.1 to 10. Hence comparisons between homologous proteins of substantially differing stabilities require extrapolation between varying conditions of increased destabilization (e.g. temperature, pH or denaturant). For a reversible two-state transition, the variation in stability as a function of temperature has a slope $\delta\Delta G/\delta T = -\Delta S$ and a curvature of $\delta^2\Delta G/\delta T^2 = -\Delta C_p/T$ [1]. The heat capacity determines the temperature dependent variation of both entropy and enthalpy which for the case of a temperature independent ΔC_p take the form:

$$\Delta S = \Delta S_{T_m} + \Delta C_p \ln(T/T_m) \quad (1)$$

and

$$\Delta H = \Delta H_{T_m} + \Delta C_p (T - T_m) \quad (2)$$

For a well behaved protein the thermal transition point can be accurately measured, while the slope of the curve at that point ($\Delta S_{T_m} = \Delta H_{T_m}/T_m$) is generally somewhat less well determined. Accurate determination of the heat capacity is considerably more problematic. Although in principle the difference in heat capacity for the folded and unfolded

* Tel.: +1 518 474 6396; fax: +1 518 473 2900.
E-mail address: lemaster@wadsworth.org.

0301-4622/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2005.08.003

states of the protein can be directly determined calorimetrically from the baselines above and below the unfolding transition region, in practice these data exhibit too large of a variation to provide a reliable measure [2]. In practice, the heat capacity is most commonly estimated calorimetrically by varying the T_m via alteration of the pH and using corresponding variation in ΔH to derive a ΔC_p value [3]. However, this analysis does not incorporate the thermodynamic effects of other coupled pH dependent interactions. As illustration, the change in ΔH as a function of pH with the Sac7d DNA binding protein from the hyperthermophile *Sulfolobus acidocaldarius* yielded an anomolously low ΔC_p value which the authors identified as arising from the neglect of increased anion binding at more acidic pH values [4,5]. A more complete linkage analysis led to a revised ΔC_p that was increased by 40%. More generally, the apparent change in heat capacity has been shown to vary as a function of ionic strength as well as denaturing and stabilizing solute concentration [2].

Differential stability analysis is commonly applied to comparisons between structurally homologous proteins. It is of interest to ascertain under what conditions the differential free energy can be reliably estimated in the absence of a known value of heat capacity. When such an analysis is combined with minimizing the degree of extrapolation required, the accuracy of the derived $\Delta\Delta G$ values may be optimized.

2. Theoretical analysis

Given a naturally occurring protein of modest thermal stability, a basic thermodynamic question involves whether the increased stability observed for a homologous protein is achieved primarily by an increase in enthalpy, a decrease in entropy, or a combination thereof. For a reversible two-state transition with a temperature independent heat capacity, combining Eqs. (1) and (2) yields the modified Gibbs–Helmholtz formula using the T_m as reference temperature:

$$\Delta G = \Delta S_{T_m}(T_m - T) + \Delta C_p[(T - T_m) + T\ln(T_m/T)] \quad (3)$$

In Fig. 1 is illustrated the stability curve for a protein with a T_m of 45 °C, a ΔS_{T_m} of 0.25 kcal/mol deg and a temperature independent heat capacity of 3.3 kcal/mol deg, consistent with that expected of the adenylate kinase from *Bacillus globisporus* as discussed below. Using purely a change in enthalpy to achieve a 30 °C increase in T_m requires a 12.1 kcal/mol upward shift in the stability curve. The slope of the stability curve as a function of temperature is unaffected so that the entropy of folding for the more thermostable protein at its T_m can be directly calculated from the stability curve of the less thermostable protein according to Eq. (1).

In Fig. 1 a tangent is drawn at the T_m of the stability curve for the less thermostable protein indicating what the

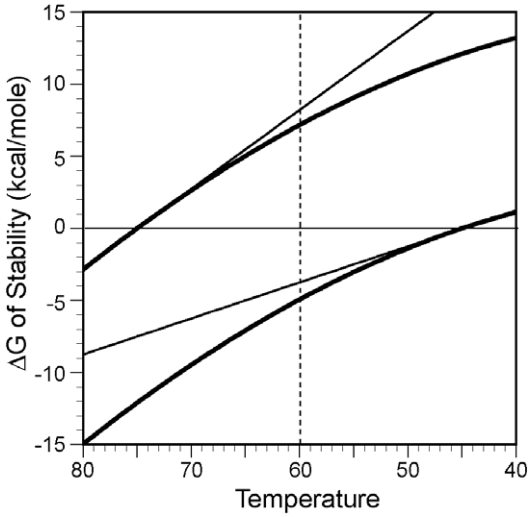

Fig. 1. Differential protein stability achieved by a purely enthalpic change. The free energy of stability curve is illustrated for a protein having a T_m of 45 °C, ΔS_{T_m} of 0.25 kcal/mol · deg and a temperature independent ΔC_p of 3.3 kcal/mol · deg. Also indicated is the stability curve for a thermostable protein with the same heat capacity and a T_m of 75 °C achieved by increasing the ΔG values without alteration of the slope of the curve as a function of temperature. Tangents to the curves at the T_m values indicate the free energy estimate based on a temperature independent entropy.

stability would be if the entropy were temperature independent. In their classic analysis, Becktel and Schellman noted that for a pair of proteins related by a small perturbation in thermodynamic stability this tangent can be used to predict a $\Delta\Delta G$ equal to $\Delta T_m \bullet \Delta S_{T_m}$ [1]. The present example of a 30 °C stabilization is clearly not a small perturbation, and the linear prediction underestimates the differential stability at 75 °C by nearly 40%. Given that the curvature of the stability curve is always negative as a result of the higher heat capacity in the unfolded state, linear extrapolation to higher temperatures will always underestimate the true $\Delta\Delta G$ while extrapolation to lower temperatures will overestimate that value. Calling attention to the temperature midway between the two T_m values, the effects of the two linear extrapolations largely cancel so that the energy difference between these two lines closely approximates the actual energy difference between the two stability curves.

The opposite extreme of a T_m shift arising from purely entropic changes is well approximated by a lateral shift of the stability curve as illustrated in Fig. 2. The modest deviation from a purely entropic change is due to the dependence of curvature on inverse absolute temperature which gives rise to a 9% relative variation in ΔC_p over the illustrated range. The stability curve for the thermostabilized protein differs markedly from that represented in Fig. 1. Nevertheless, at the midpoint between the T_m values, the difference between the two linear extrapolations again closely approximates the $\Delta\Delta G$. As shown below, this similarity between $\Delta\Delta G$ and the combined linear extrapolations is nearly independent of the assumed heat capacity.

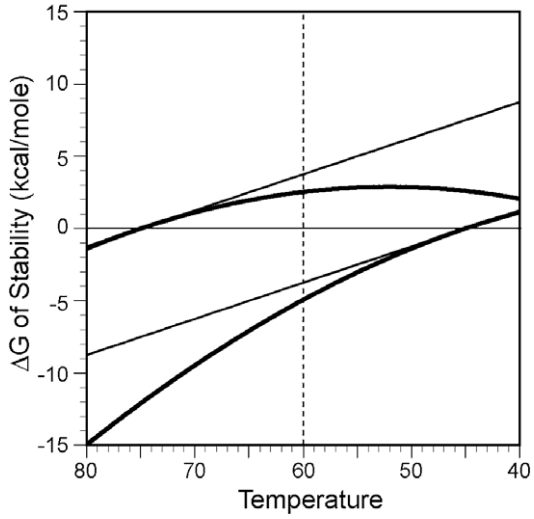

Fig. 2. Differential protein stability achieved by a primarily entropic change. The parameters for the free energy of stability curve of the less thermostable protein are as given in Fig. 1. The stability curve for the more thermostable protein with a T_m of 75 °C is obtained by uniformly shifting the curve to the higher temperature. Tangents to the curves at the T_m values indicate the free energy estimate based on a temperature independent entropy.

The series expansion of the $T\ln(T_m/T)$ factor in the heat capacity term of the Gibbs–Helmholtz equation is given as follows:

$$\ln(x+1) = \sum(-1)^{n+1}(x)^n/n = x - (x^2/2) + (x^3/3) - \cdots \quad (4)$$

hence

$$T\ln(T_m/T) = (T_m - T) - \left[(T_m - T)^2/2T\right] + \left[(T_m - T)^3/3T^2\right] - \cdots \quad (5)$$

As is well known, when combined with the $\Delta C_p\,(T - T_m)$ term in Eq. (3), the linear term of this expansion cancels so that only the quadratic and higher $(T_m - T)$ terms contribute to the Gibbs–Helmholtz equation.

$$\Delta G = \Delta S_{T_m}(T_m - T) - \Delta C_p[(T_m - T)^2/2T - (T_m - T)^3/3T^2 + \cdots] \quad (6)$$

When this equation is applied at the temperature midway between the T_m values, $\bar{T}_m$, for proteins having similar heat capacity values, their differential free energy is independent of the quadratic term as well.

$$\Delta G_{\bar{T}_m}(1) - \Delta G_{\bar{T}_m}(2) \equiv \Delta\Delta G_{\bar{T}_m} = 0.5[\Delta S_{T_{m1}}(1) + \Delta S_{T_{m2}}(2)]\Delta T_m - \Delta C_p\left[\Delta T_m^3/12\bar{T}_m^2 + \cdots\right] \quad (7)$$

For the examples illustrated in Figs. 1 and 2, the contribution of the heat capacity dependent term is less than 1% of the total differential free energy, well within practical experimental error. Hence even for pairs of homologous proteins which differ substantially in their free energies of stability, the approximate formula

$$\Delta\Delta G_{\bar{T}_m} = 0.5[\Delta S_{T_{m1}}(1) + \Delta S_{T_{m2}}(2)]\Delta T_m \quad (8)$$

can be anticipated to hold.

Although the cancellation of the quadratic term of the heat capacity expansion is based on the assumption of equal ΔC_p values for both proteins, small variations in the heat capacity lead to only modest discrepancies. Several studies [6–9] have indicated a reasonably good correlation between ΔC_p and the surface area of the protein predicted to become accessible to water upon unfolding. In particular, in the study by Myers et al. [8] all of the data points from 45 different proteins lay within ±25% of the line of regression. Under the assumption that the heat capacity of one of the proteins in the above figures is varied by 25%, the resultant error in ignoring the ΔC_p dependence is less than 4%.

For the entropy-dominated comparison of Fig. 2 the slope of the stability curve at the thermal transition is the same for both proteins, while for the enthalpy-dominated comparison of Fig. 1 the slope of the more thermostable protein is increased by $\Delta C_p \ln(T_{m1}/T_{m2})$ as indicated in Eq. (1). An alternate paradigm for thermostabilization has been widely discussed in which proteins from thermophilic organisms have a reduced heat capacity relative to their mesophile homologs. This lowered ΔC_p gives rise to a broadening of the temperature range of positive stability without increasing the maximal free energy of stability. The most direct experimental prediction of this model is that under the same solution conditions the ΔS_{T_m} of the thermophile protein will be less than that of its mesophile homolog [10]. In particular, under the parabolic approximation for the stability curve, the slope of the thermophile protein at its T_m will be reduced by $\sqrt{(\Delta C_p^T/\Delta C_p^M)}$.

To date no such experimental observation of a reduced ΔS_{T_m} for the thermophile homolog has been reported. In cases such as RNase H [11] and CheY [12] for which stability measurements in high concentrations of guanidinium chloride have been used to predict a lowered heat capacity for the thermophile protein, a substantial increase in enthalpy is observed as well. For both of these proteins the resultant stability curves extrapolated to zero denaturant predict the ΔS_{T_m} value for the thermophile protein to be greater than that of its mesophile homolog, indicative of the

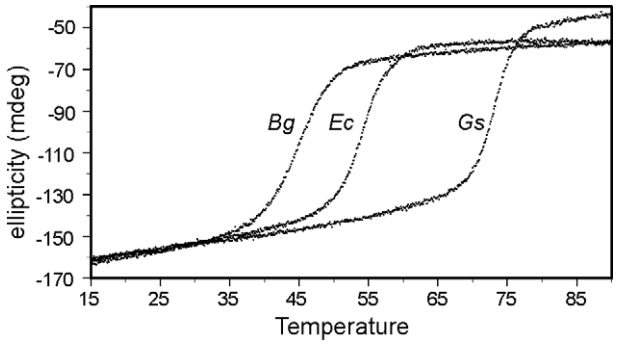

Fig. 3. Thermal CD titration of homologous adenylate kinases. Wild type protein sequences from *Bacillus globisporus* (*Bg*) and *Geobacillus stearothermophilus* (*Gs*) were encoded in de novo synthesized genes (Genscript, Piscataway NJ, USA) utilizing *Escherichia coli* (*Ec*)-optimized codons and expressed in the BL21(DE3)/pET11a system [30] (Novagen, Madison WI, USA). The 10 μM protein samples were equilibrated in 50 mM sodium cacodylate pH 6.80. A temperature ramping rate of 1 °C per minute was used with data collection at 0.1 °C intervals on a JASCO 710 spectropolarimeter. In each case the thermal titrations to 90% unfolded were more than 90% reversible.

fact that the increased enthalpy of the thermophile protein contributes substantially to the increase in T_m.

3. Experimental results

Thermal CD titration experiments were carried out on the structurally homologous adenylate kinases from the psychrophilic arctic bacterium *B. globisporus* (*Bg*), the mesophilic bacterium *Escherichia coli* (*Ec*) and the thermophilic bacterium *Geobacillus stearothermophilus* (*Gs*). A single scan of the raw data from each protein is illustrated in Fig. 3. The structural homology among these proteins has been directly verified by crystallographic analysis [13–15]. In particular, the structures of the two *Bacillus* adenylate kinases are highly homologous with a backbone C^{α} rmsd for the CORE and AMP-binding domains of 0.61 Å and within the reoriented LID domain an rmsd of 0.43 Å [15]. The corresponding sequences exhibit 66% identity without alignment gaps. Both *Bacillus* adenylate kinases contain a zinc-binding cluster in the LID domain which is absent in the Gram negative *E. coli* protein which exhibits a more divergent structure with a C^{α} rmsd of 2.2 Å and 47% sequence identity with respect to the *B. globisporus* enzyme.

The T_m and ΔS_{T_m} values derived from the CD measurements are listed in Table 1. The T_m values agree well with those previously reported for these proteins [15–17]. Although irreversibility has been noted at protein concentrations used in calorimetric measurements [15], reversibility was satisfactory at the 10 μM concentrations used in the present study. Application of the ΔC_p-independent formula for differential stability (Eq. (8)) yields $\Delta\Delta G$ values at the thermal midpoint between the protein T_m values. Not included in the table is the $\Delta\Delta G_{\bar{T}_m}$ value for the comparison between the *E. coli* and *G. stearothermophilus* adenylate kinases which yields 6.34 kcal/mol at 63.5 °C. Combined with the corresponding value of 2.65 kcal/mol at 49.4 °C for the *E. coli* vs. *B. globisporus* comparison, the sum is nearly identical to the $\Delta\Delta G_{\bar{T}_m}$ value (at 58.8 °C) for the two bacilli proteins given in that table.

This additivity of $\Delta\Delta G$ at the various $\bar{T}_m$ values will hold whenever the ΔS_{T_m} of the third protein is the ΔT_m-weighted average of the ΔS_{T_m} values for the other two proteins.

$$\Delta S_{T_{m3}}(3) = [\Delta S_{T_{m1}}(1)\cdot(T_{m3}-T_{m2}) + \Delta S_{T_{m2}}(2)\cdot(T_{m1}-T_{m3})]/(T_{m1}-T_{m2}) \quad (9)$$

Multiplying through by $(T_{m1}-T_{m2})$ and rearranging yields:

$$[\Delta S_{T_{m1}}(1)+\Delta S_{T_{m2}}(2)]\cdot(T_{m1}-T_{m2}) = [\Delta S_{T_{m1}}(1)+\Delta S_{T_{m3}}(3)]\cdot(T_{m1}-T_{m3}) + [\Delta S_{T_{m2}}(2)+\Delta S_{T_{m3}}(3)]\cdot(T_{m3}-T_{m2}) \quad (10)$$

Hence

$$\Delta\Delta G_{\bar{T}_m}(12) = \Delta\Delta G_{\bar{T}_m}(13) + \Delta\Delta G_{\bar{T}_m}(23) \quad (11)$$

Included in Table 1 are the ΔS_{T_m} values predicted for the *G. stearothermophilus* and *E. coli* adenylate kinases assuming an enthalpy-dominated variation from that of the *B. globisporus* protein, using a ΔC_p value of 3.3 kcal/mol deg as estimated from the reported correlation with molecular weight [8]. The predicted increases in ΔS_{T_m} for

Table 1
Thermodynamic stability of adenylate kinases

	T_m (°C)	ΔS_{T_m}	$\Delta\Delta G_{\bar{T}_m}$ [a]	ΔS_{T_m} (ΔC_p) [b]	ΔS_{T_m} (1/2 ΔC_p) [b]
B. globisporus	44.6	0.251	–	–	–
E. coli	54.2	0.302	2.65 (49.9°C)	0.349	0.300
G. stearothermophilus	72.9	0.377	8.89 (58.8°C)	0.532	0.391

[a] Differential free energy at midpoint between T_m values using ΔC_p-independent Eq. (8). Error estimates of 0.006 for ΔS_{T_m} (kcal/mol·deg), 0.3 °C for ΔT_m and 0.14 for $\Delta\Delta G_{\bar{T}_m}$ (kcal/mol) based on three independent scans for each protein.

[b] ΔS_{T_m} predictions assuming enthalpy only or equivalent entropy and enthalpy contributions to the differences in entropy among the proteins using a ΔC_p value of 3.3 kcal/mol · deg.

the thermophile *G. stearothermophilus* and the mesophile *E. coli* adenylate kinases are both twice as large as that which is experimentally observed. Stated differently, for both of these proteins the slope of the stability curve at the thermal transition midpoint is halfway between that predicted by a purely enthalpic vs. an entropic variation of stability with respect to the *B. globisporus* protein.

For the case in which variations in entropy and enthalpy contribute equally to the differential free energy of stabilization, the slope of the thermostable protein at its T_m is given by:

$$\Delta S_{T_{m1}} = \Delta S_{T_{m2}} + 0.5\Delta C_p \ln(T_{m1}/T_{m2}) \quad (12)$$

Utilizing parameters analogous to those of Figs. 1 and 2, Fig. 4 illustrates the condition of entropy–enthalpy compensation in the differential free energy for model calculations closely mimicking the data of the two bacilli adenylate kinases. Applying the ΔC_p-independent formula of Eq. (8) predicts a 0.7% deviation from the exact free energy at $\bar{T}_m$.

Several additional properties are found to hold under the condition in which the differential entropic contributions are compensated for by equivalent variations in enthalpy. At $\bar{T}_m$ the ΔG value of the thermostable protein is equal to $-\Delta G$ for the less stable protein (deviation <0.02% for parameters of Fig. 4). This reflects the similarity in the slope of the stability curve of the thermostable protein between 60 and 75 °C as compared to that of the less stable protein over the range of 45 to 60 °C. The slope $\Delta S_{75°}(Gs)$ is 2% less than $\Delta S_{\bar{T}_m}(Bg)$ while in turn $\Delta S_{\bar{T}_m}(Gs)$ is 2% greater than $\Delta S_{45°}(Bg)$. The equal and opposite free energy values at $\bar{T}_m$ will hold for any two proteins with equal heat capacities and ΔS_{T_m} values which conform to Eq. (12).

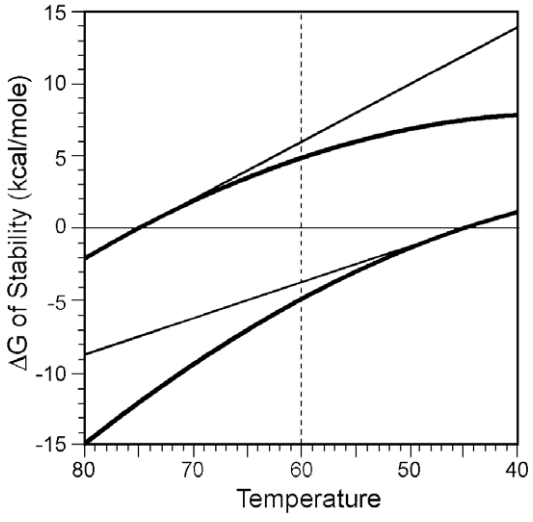

Fig. 4. Differential protein stability achieved by entropy–enthalpy compensation. The parameters for the free energy of stability curve of the less thermostable protein are as given in Fig. 1. The stability curve for the more thermostable protein assumes the same heat capacity. The slope of the curve at the higher T_m (i.e. ΔS_{T_m}) is set to the average of the corresponding slopes from Figs. 1 and 2.

In many protein design studies, hybrid sequences are generated from two parental wild type homologous proteins of differing thermal stabilities. Although the T_m values for such hybrids are often cited as a measure of differential stability, concerns regarding a lack of thermodynamic validity are commonly raised. However, for hybrid and parental proteins which share a common ΔC_p value, if their T_m and ΔS_{T_m} differences arise from entropy–enthalpy compensation, then the incremental T_m value of the hybrid protein will be essentially proportionate to its $\Delta\Delta G$ of stability at the parental $\bar{T}_m$. This proportionality between $\Delta\Delta G_{\bar{T}_m}$ and ΔT_m would be rigorous if the entropy–enthalpy compensation condition yielding $\Delta G_{\bar{T}_m}(1)$ being equal to $-\Delta G_{\bar{T}_m}(2)$ were combined with the free energy additivity condition of Eq. (11) (consider a second hybrid with complementary T_m and ΔS_{T_m} properties).

However, the requisite ΔT_m-weighted averaging of the ΔS_{T_m} values is not strictly compatible with the entropy–enthalpy compensation condition of Eq. (12). On the other hand for the model calculations of Fig. 4, the maximum deviation in the ΔS_{T_m} values over the range of 45 to 75 °C calculated by Eq. (9) vs. Eq. (12) is 0.5%, giving rise to an experimentally indistinguishable deviation in the proportionality between $\Delta\Delta G_{\bar{T}_m}$ and ΔT_m.

The proportionality between ΔT_m and $\Delta\Delta G_{\bar{T}_m}$ becomes more approximate as either enthalpy or entropy comes to dominate the differential stability. The deviations from a linear correlation between ΔT_m and $\Delta\Delta G_{\bar{T}_m}$ are maximal for hybrid proteins having T_m values near $\bar{T}_m$. As we [18] have shown for the enthalpy-dominated differential stability, the ΔT_m values of such hybrids are systematically higher than that predicted from the differential free energies. In the present example of the parameters in Fig. 1, a hybrid having a $\Delta G_{\bar{T}_m}$ value midway between that of the parental proteins has a T_m that is 2.6 °C above the $\bar{T}_m$ value of 60 °C. Similarly, for the entropy-dominated differential stability parameters of Fig. 2, the analogous hybrid has a T_m that is 4.0° below $\bar{T}_m$.

4. Discussion

Although protein stability measurements typically require variation in the temperature or solution conditions so as to significantly populate both the folded and unfolded states, subsequent thermodynamic analysis often references the data back to 25 °C standard buffer conditions. In part this reflects the more general convention for biochemical reactions. However in addition, it also closely approximates the temperature of maximum stability for most proteins. On the other hand as noted above, estimation of the maximal free energy of stability generally requires substantial extrapolation from the observed experimental data.

For thermal unfolding data this extrapolation relies upon the accuracy of the measured ΔS_{T_m} and the often more poorly characterized ΔC_p. The analogous limitations apply to extrapolated stability estimates based on increasing denaturant concentrations at a constant temperature. Inaccuracies in measurement of the m-values characterizing the denaturant concentration dependence of free energies contribute to extrapolation errors analogous to those of ΔS_{T_m}. Furthermore, urea vs. guanidinium chloride based extrapolations to the same conditions often yield markedly different stability estimates [19]. This implies a breakdown in the commonly assumed linearity in the m-value extrapolation, presumably largely a reflection of both the ionic strength and solute binding effects of the more powerful guanidinium chloride denaturant [20,21].

An additional appeal to thermodynamic analysis at the point of maximal protein stability is that since $\delta\Delta G/\delta T$ is near zero in this temperature range, entropy–enthalpy compensation necessarily occurs. On the other hand, there is good reason to anticipate entropy–enthalpy compensation to occur under a much wider range of ΔG perturbations. Dunitz [22] has argued that entropy–enthalpy compensation is a general property of weak intermolecular interactions. When the characteristic spacing to the lowest lying excited energy states of an interaction is similar to the average thermal energy, modest perturbations can be expected to elicit entropic and enthalpic contributions of comparable magnitude. Although in the protein stability field, entropy–enthalpy compensation is often discussed in terms of the temperature dependence of hydrophobic interactions, both experimental evidence and theoretical analysis [23,24] indicate the involvement of a wider range of interactions under both isothermal and variable temperature conditions.

In the present study, the free energy perturbation in question is achieved by mutational substitution. The differential stability observed for homologous proteins derived from organisms which live in substantially different temperature environments appears to often arise from a relatively large number of differential interactions. It may be anticipated that, similar to the adenylate kinase proteins, the differential stability between many homologous protein structures will have entropic and enthalpic contributions of comparable magnitude.

The analysis approach described has the conceptual limitation that the reference temperature is determined by the T_m values of the two proteins being compared, although as illustrated in the analysis of *E. coli* adenylate kinase, under certain conditions, comparisons between free energy values determined at differing reference temperatures can still prove useful. On the other hand, since the $\bar{T}_m$ reference lies between the two protein T_m values, it will generally tend to mimic the temperature conditions under which their sequence variations evolved. Furthermore, enzymatic biotechnology applications often operate near the maximum permissible temperature, so that stability at $\bar{T}_m$ will generally reflect these conditions.

The utility of this ΔC_p-independent analysis will be diminished when there is a substantial difference in heat capacity between the two proteins being compared. Although several compilations have indicated that the correlation between molecular weight and heat capacity exhibits a fairly modest variance among naturally occurring proteins, mutational variants have been described for which the change in ΔC_p is substantially larger than predicted by these correlations. A notable example is that of the studies by Makhatadze and coworkers [25,26] in which three mutations of ubiquitin were generated by substitution of the similar sized asparagine for buried nonpolar valine or leucine residues. For all three mutations the resultant ΔC_p changed by more than 0.2 kcal/mol deg from the wild type ubiquitin, while in turn each mutated protein was destabilized by more than 3 kcal/mol as compared to the parental protein. Comparison to the parental ubiquitin structure indicated no suitably positioned hydrogen bonding partners for the amide containing sidechain of these mutants. In naturally occurring proteins, buried protein hydrogen bond donor and acceptor groups are found to be unpaired only ~ 1% of the time [27]. Hence, although such mutations can substantially alter the heat capacity from that predicted by the linear surface area correlations, their adverse effects on the protein stability or structure render it unlikely that they will be propagated through the evolutionary process.

Given the approximate estimate of heat capacity that can be derived from the surface area correlations, comparison of the slope of the free energy curves at the thermal transition (i.e. ΔS_{T_m}) can provide a basis for assessing whether the differential stability between a homologous pair of proteins arises from predominantly entropic or enthalpic contributions or from a similar magnitude contribution of both. Extending the previously illustrated free energy curves to lower temperatures under the assumption of a common ΔC_p value predicts differing temperatures of maximum stability T^*. For the entropy-dominated differential represented in Fig. 2 the difference in T^* is equal to the difference in T_m values, while the maximum stability temperature is unchanged for the enthalpy-dominated process (Fig. 1). In the presence of entropy–enthalpy compensation the differential for the maximum stability temperatures is one half that between the T_m values (Fig. 4).

In comparing estimated T^* values for a set of hyperthermophile vs. mesophile proteins, Rees and Robertson [28] observed average T^* values of 37 and 12 °C, respectively. Given the small number of hyperthermophile protein measurements available at the time of that study, the corresponding average T^* value is not well determined. Nevertheless, it is of interest to compare this differential to that which would be predicted from the entropy–enthalpy compensation model illustrated in Fig. 4. A recent survey [29] of homologous mesophile and thermophile protein pairs indicated that for host organisms with optimal growth temperatures differing by more than 30 °C, the ratio of $\Delta T_m / \Delta T_{\text{growth opt}}$ lay in the range of 0.5 to 0.8. For a maximum

$\Delta T_{\text{growth opt}}$ of ~70 °C between mesophilic and hyperthermophilic organisms, this predicts ΔT_m values in the range of 35 to 56 °C. Hence, the 25 °C differential for average T^* values reported for these classes of organisms is consistent with the model of entropy–enthalpy compensation occurring in the differential stabilization of evolutionarily derived homologous proteins.

Acknowledgment

I wish to acknowledge the use of the Wadsworth Center Biochemistry core facility and the technical assistance of Leslie Eisele.

References

[1] W.J. Becktel, J.A. Schellman, Protein stability curves, Biopolymers 26 (1987) 1859–1877.
[2] Y. Liu, J.M. Sturtevant, The observed change in heat capacity accompanying the thermal unfolding of proteins depends on the composition of the solution and on the method employed to change the temperature of unfolding, Biochemistry 35 (1996) 3059–3062.
[3] P.L. Privalov, Stability of proteins. Small globular proteins, Adv. Protein Chem. 33 (1979) 167–241.
[4] B.S. McCrary, S.P. Edmondson, J.W. Shriver, Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d, J. Mol. Biol. 264 (1996) 784–805.
[5] B.S. McCrary, J. Bedell, S.P. Edmondson, J.W. Shriver, Linkage of protonation and anion binding to the folding of Sac7d, J. Mol. Biol. 276 (1998) 203–224.
[6] K.P. Murphy, E. Freire, Thermodynamics of structural stability and cooperative folding behavior in proteins, Adv. Protein Chem. 43 (1992) 313–361.
[7] R.S. Spolar, J.R. Livingstone, J.M.T. Record, Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water, Biochemistry 31 (1992) 3947–3955.
[8] J.K. Myers, C.N. Pace, J.M. Scholtz, Denaturant m values and heat-capacity changes: relation to changes in accessible surface areas of protein folding, Protein Sci. 4 (1995) 2138–2148.
[9] G.I. Makhatadze, P.L. Privalov, Energetics of protein structure, Adv. Protein Chem. 47 (1995) 307–425.
[10] B.M. Beadle, W.A. Baase, D.B. Wilson, N.R. Gilkes, B.K. Shoichet, Comparing the thermodynamic stabilities of a related thermophilic and mesophilic enzyme, Biochemistry 38 (1999) 2570–2576.
[11] J. Hollien, S. Marqusee, A thermodynamic comparison of mesophilic and thermophilic ribonucleases H, Biochemistry 38 (1999) 3831–3836.
[12] W.A. Deutschman, F.W. Dahlquist, Thermodynamic basis for the increased thermostability of CheY from the hyperthermophile *Thermotoga maritima*, Biochemistry 40 (2001) 13107–13113.
[13] C.W. Mueller, G.E. Schulz, Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor AP5A refined at 1.9 Å resolution — a model for a catalytic transition-state, J. Mol. Biol. 224 (1992) 159–177.
[14] M.B. Berry, G.N. Phillips, Crystal structures of *Bacillus stearothermophilus* adenylate kinase with bound Ap(5)A, Mg2+ Ap(5)A and Mn2+ Ap(5)A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg2+ and Mn2+, Protein Struct. Funct. Genet. 32 (1998) 276–288.
[15] E. Bae, J.G.N. Phillips, Structures and analysis of highly homologous psychrophilic, mesophilic and thermophilic adenylate kinases, J. Biol. Chem. 279 (2004) 28202–28208.
[16] M. Monnot, A.M. Gilles, I.S. Giron, S. Michelson, O. Barzu, S. Fermandjian, Circular dichroism investigation of *Escherichia coli* adenylate kinase, J. Biol. Chem. 262 (1987) 2502–2506.
[17] P. Glaser, E. Presecan, M. Delepierre, W.K. Surewicz, H.H. Mantsch, O. Barzu, A.M. Gilles, Zinc, a novel structural element found in the family of bacterial adenylate kinases, Biochemistry 31 (1992) 3038–3043.
[18] D.M. LeMaster, J. Tang, D.I. Paredes, G. Hernández, Contribution of the multi-turn segment in the reversible thermal stability of hyperthermophile rubredoxin: NMR thermal chemical exchange analysis of sequence hybrids, Biophys. Chem. 116 (2005) 57–65.
[19] C.N. Pace, Determination and analysis of urea and guanidine hydrochloride denaturation curves, Methods Enzymol. 131 (1986) 266–280.
[20] G.I. Makhatadze, Thermodynamics of protein interactions with urea and guanidinium hydrochloride, J. Phys. Chem. B 103 (1999) 4781–4785.
[21] P.E. Mason, G.W. Neilson, J.E. Enderby, M.L. Saboungi, C.E. Dempsey, J.A.D. MacKerell, J.W. Brady, The structure of aqueous guanidinium chloride solutions, J. Am. Chem. Soc. 126 (2004) 11462–11470.
[22] J.D. Dunitz, Win some, lose some: enthalpy–entropy compensation in weak intermolecular interactions, Curr. Biol. 2 (1995) 709–712.
[23] A. Cooper, C.M. Johnson, J.H. Lakey, M. Nollmann, Heat does not come in different colors: entropy–enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions, Biophys. Chem. 93 (2001) 215–230.
[24] A. Cooper, Heat capacity effects in protein folding and ligand binding: a re-evaluation of the role of water in biomolecular thermodynamics, Biophys. Chem. 115 (2005) 89–97.
[25] V.V. Loladze, D.N. Ermolenko, G.I. Makhatadze, Heat capacity changes upon burial of polar and nonpolar groups in proteins, Protein Sci. 10 (2001) 1343–1352.
[26] V.V. Loladze, D.N. Ermolenko, G.I. Makhatadze, Thermodynamic consequences of burial of polar and non-polar amino acid residues in the protein interior, J. Mol. Biol. 320 (2002) 343–357.
[27] I.K. McDonald, J.M. Thornton, Satisfying hydrogen bonding potential in proteins, J. Mol. Biol. 238 (1994) 777–793.
[28] D.C. Rees, A.D. Robertson, Some thermodynamic implications for the thermostability of proteins, Protein Sci. 10 (2001) 1187–1194.
[29] S. Kumar, C.-J. Tsai, R. Nussinov, Thermodynamic differences among homologous thermophilic and mesophilic proteins, Biochemistry 40 (2001) 14152–14165.
[30] F.W. Studier, B.A. Moffatt, Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, J. Mol. Biol. 189 (1986) 113–130.