THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

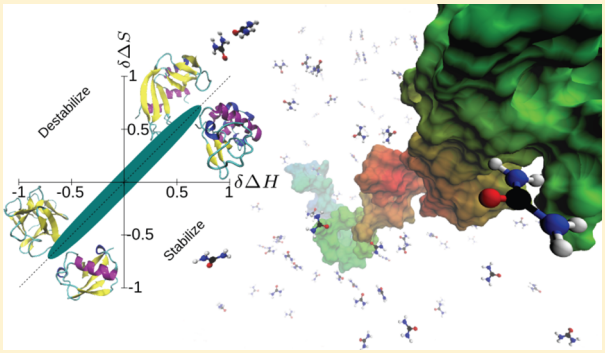
pubs.acs.org/JPCB

Protein Transfer Free Energy Obeys Entropy-Enthalpy Compensation

Eric A. Mills and Steven S. Plotkin*

Department of Physics & Astronomy, University of British Columbia, Vancouver, British Columbia V6T1Z4, Canada

ABSTRACT: We have found significant entropy-enthalpy compensation for the transfer of a diverse set of two-state folding proteins from water into water containing a diverse set of cosolutes, including osmolytes, denaturants, and crowders. In extracting thermodynamic parameters from experimental data, we show the potential importance of accounting for the cosolute concentration-dependence of the heat capacity change upon unfolding, as well as the potential importance of the temperature-dependence of the heat capacity change upon unfolding. We introduce a new Monte Carlo method to estimate the experimental uncertainty in the thermodynamic data and use this to show by bootstrapping methods that entropy-enthalpy compensation is statistically significant, in spite of large, correlated scatter in the data. We show that plotting the data at the transition midpoint provides the most accurate experimental values by avoiding extrapolation errors due to uncertainty in the heat capacity, and that this representation exhibits the strongest evidence of compensation. Entropy-enthalpy compensation is still significant at lab temperature however. We also find that compensation is still significant when considering variations due to heat capacity models, as well as typical measurement discrepancies lab-to-lab when such data is available. Extracting transfer entropy and enthalpy along with their uncertainties can provide a valuable consistency check between experimental data and simulation models, which may involve tests of simulated unfolded ensembles and/or models of the transfer free energy; we include specific applications to cold shock protein and protein L.



1. INTRODUCTION

Enthalpy and entropy play an intimately connected role in the free energy change during numerous biochemical processes. It has long been known, for example, that the transfer of hydrocarbons such as alkanes or alcohols from pure solvent to water is generally exothermic or enthalpically favorable ($\Delta H \approx -$ (3−6) kcal/mol for methane-butane) but entropically unfavorable, with $T\Delta S$ typically 2−3× the magnitude of the enthalpic contribution.[1,2] These opposing thermodynamic forces result in a net free energy change that is often smaller than either of the enthalpic or entropic contributions.

For a variety of physical processes including solute transfer,[3] unfolding of various proteins,[4] and ligand binding, ionization, and hydrolysis,[5] the changes in enthalpy and entropy obey a nearly linear relationship when a variable such as binding ligand is varied; this is referred to as entropy-enthalpy compensation.[5−10] The effect is ubiquitous but not universal.[11] The slope of the enthalpy vs entropy plot, referred to as the compensation temperature T_c, ranges from about 150 K (e.g., for alkane vaporization[12]) to about 300 K (most processes). The difficulty in designing high affinity drugs has been attributed to entropy-enthalpy compensation.[13−15]

As mentioned above, entropy-enthalpy compensation is not an inevitable consequence of statistical mechanics, particularly along chemical reaction coordinates. Long-lived metastable states due to large barriers, and thus the absence of any significant entropy-enthalpy compensation along the reaction coordinate, are fairly common in condensed matter and biophysics. The diamond phase of carbon is metastable to graphite at standard temperature and pressure, with an enormous conversion barrier; allotropes of boron, polymorphs of silica, and martensite in steel are all metastable phases with prohibitive transition barrier; colloidal systems and emulsions have long-lived metastable phases; long-lived structure with slow, glassy dynamics is common in supercooled liquids. The covalent bonds forming the backbones of DNA, RNA, and proteins are metastable to hydrolysis; several proteins have native, functional states that are metastable but simply have enormous kinetic unfolding barriers, including α-lytic protease, subtilisin, *Streptomyces griseus* protease B, and the aspartic peptidase pepsin.[16] In multimeric systems of chain length $\lesssim$100 amino acids, native protein structures have been observed to have higher free energy than the amyloid phase, implying that a significant portion of the proteome is conformationally in metastable equilibrium.[17] In contrast to these observations, as a general rule, entropy-enthalpy compensation does play a critical role in governing the foldability of proteins and resolving the Levinthal paradox.[18] Small barriers in protein folding have been shown to arise due to the locality of interactions and the concomitant loss of entropy in forming stabilizing interactions. If it were not for entropy-enthalpy compensation as protein chain conformations progressed toward native-like folds,

Received: September 21, 2015
Revised: September 30, 2015
Published: September 30, 2015

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

folding barriers would be prohibitively high, proteins could not fold on biological time scales, and life as we know it would not be possible.

The entropy is often obtained from measurements of the enthalpy and free energy by subtraction; one complication however is that errors in enthalpy are often much larger than errors in free energy. In these cases the error in enthalpy can induce a spurious linear relation to the entropy, for example in early measurements for oximation reactions of alkyl thymyl ketones,[19] correlated entropy-enthalpy errors were sufficiently large that they could account for the whole measurable effect of entropy-enthalpy compensation, which could then not be definitively proven, regardless of whether or not it existed. Correlated errors generally have an effective compensation temperature equal to the temperature at which the measurements were taken. Compensation exists however when large entropy and enthalpy changes either cancel or compensate each other to yield a relatively small net free energy gain for a given process, regardless of the compensation temperature, and includes cases where the compensation temperature equals the lab temperature.[6,9,20] As well, even if the compensation temperature is quite different from the temperature of the experiments, if the correlated scatter is sufficiently large, it can rule out the significance of the effect. Thus, there is a need to introduce more rigorous error analysis to judge the significance of any observed entropy-enthalpy compensation.

In this paper we investigate how broadly this effect applies to macromolecular systems, by analyzing the experimentally derived enthalpy and entropy of transferring two-state proteins from water, perhaps with buffers and at some pH which need not be 7, into the same solution but in the presence of various *cosolutes*. These cosolutes can be osmolytes, denaturants, crowders, salts, or other proteins; i.e. we place no restriction on the size or on how relatively favorable or unfavorable the interactions are with the protein.[21]

In what follows, we begin by introducing various thermodynamic equations that define the two-state model in Section 2. In Section 3.1 we introduce a Monte Carlo procedure for estimating the experimental uncertainty of thermodynamic quantities obtained from calorimetry assays. In Section 3.2.1 we show that entropy-enthalpy compensation occurs for the transfer of a diverse set of two-state proteins to various solvents. This may be the most general class of systems that have been observed to obey compensation. While it may be intuitive that a given protein and solvent series may compensate, it is not obvious that there would exist compensation across *both* solvents and proteins. For example, the excluded volume component of transfer is generally noncompensated and different across protein−solvent systems. We use the Monte Carlo and bootstrapping methods derived in Section 3.2.2 to confirm that entropy-enthalpy compensation is indeed a significant effect. We then discuss a phenomenon in Section 3.2.3 that is present across proteins, wherein the enthalpic contribution to the free energy tends to be the dominant "leading" term, and the entropy "follows" and only partially compensates. We show in Section 3.2.4 how enthalpy and entropy extracted from simulation data allows for a test of the Tanford transfer model. In Sections 3.2.5 and 3.2.6, we emphasize the importance of accounting for the (often neglected) concentration-dependence of the heat capacity change upon unfolding. In Section 3.2.7, we plot the experimental data at lab temperature, which also exhibits significant entropy-enthalpy compensation across a diverse set of proteins and cosolutes. We discuss the consequences of several sources of error in Sections 3.2.8−3.2.10, including lab-to-lab variance, and variance due to different models of the temperature-dependence of ΔC_p. Finally we conclude in Section 4.

2. METHODS AND THEORY

2.1. Thermodynamic Equations for Protein Unfolding.

Two-state models in protein folding have have a long and rich history and have empirical validity for many proteins;[22−24] various aspects of two-state folding, including applications to protein denaturation, protein stability, and the prediction of so-called *m*-values, are described elsewhere.[18,23−34]

Here, we adopt the two-state model for a set of proteins that either have been shown previously to satisfy the van't Hoff two-state criterion[34] or that have comparably small residuals when fit to a two-state model.

The changes in enthalpy ΔH, entropy ΔS, and free energy ΔG upon unfolding can be obtained if the change of heat capacity upon unfolding $\Delta C_p = C_{pu} - C_{pn}$ is measured. Here C_{pu} and C_{pn} are the unfolded and native state heat capacities respectively, which generally may depend on temperature. A temperature-independent unfolding heat capacity is often used as a good approximation,[35,36] while others have considered a linear temperature-dependence of the unfolding heat capacity.[37,38] Here, we adopt the most general temperature-dependence of the unfolding heat capacity, following the method used in Wintrode et al.,[39] wherein the folded state heat capacity C_{pn} is observed to obey a linear temperature-dependence, and the unfolded state heat capacity obeys a nonlinear temperature-dependence determined by the heat capacities of the amino acid constituents. Specifically, the heat capacity of the unfolded state C_{pu} is given by

$$C_{pu} = (N - N_{gly} - 1)C_p(\mathrm{bb}) + C_p(\mathrm{N/Cterm}) + \sum_{i=1}^{N} C_p(R_i) \quad (1)$$

Here N is the chain length, N_{gly} is the number of glycine residues in the polypeptide chain, $C_p(bb)$ is the heat capacity of the peptide backbone, C_p(N/C term) is the heat capacity of the N- and C-termini, and $C_p(R_i)$ is the heat capacity of the side chain corresponding to the ith amino acid (glycine is included in this sum). Values for $C_p(bb)$, C_p(N/C term), and $C_p(R_i)$ have been obtained by Makhatadze and Privalov[40] for temperatures of 5, 25, 50, 75, 100, and 125 °C. For the proteins and cosolutes we consider in Section 3.1, we use the values in ref 40 to interpolate $C_{pu}(T)$ from 5 to 125 °C with a cubic spline.

With $\Delta C_p(T)$ determined numerically, ΔH, ΔS, and ΔG can be calculated from

$$\Delta H = \Delta H_f + \int_{T_f}^{T} \Delta C_p(T')\mathrm{d}T' \quad (2a)$$

$$\Delta S = \Delta S_f + \int_{T_f}^{T} \frac{\Delta C_p(T')}{T'}\mathrm{d}T' \quad (2b)$$

$$\Delta G = \Delta H - T\Delta S \quad (2c)$$

The reference temperature T_f is taken to be the temperature at which the unfolding free energy is zero: $\Delta H(T_f) = T_f \Delta S(T_f)$. The unfolding heat capacity is given by $\Delta C_p(T) = C_{pu}(T) -$

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

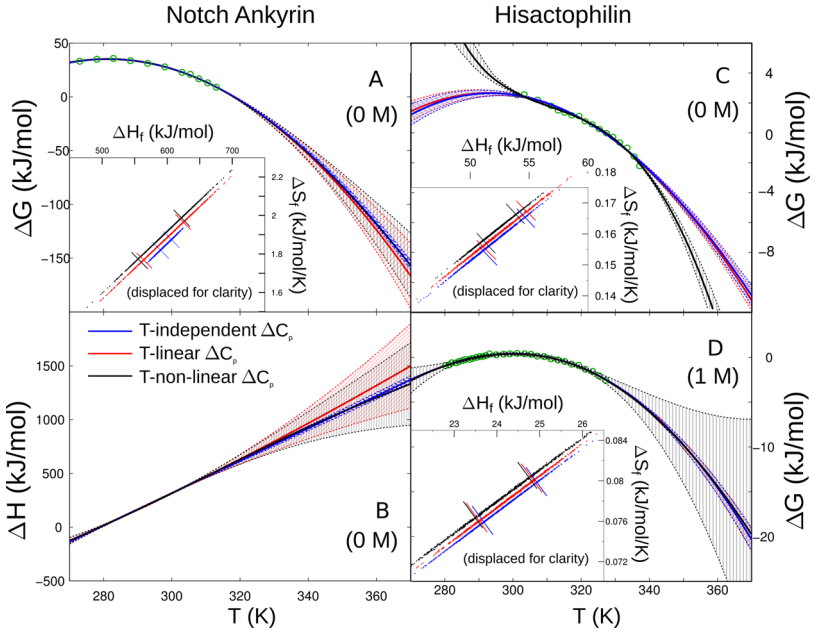


Figure 1. Stability and enthalpy as a function of temperature for notch ankyrin and hisactophilin. The green circles are experimental data from ref 41 for notch ankyrin (panels A and B) and ref 42 for hisactophilin (panels C and D). The blue lines are fits for the T-independent ΔC_p model, the red lines are fits for the linear T-dependent model, and the black lines are fits for the nonlinear T-dependent model (cf. Section 2.1). The solid lines arise from the best fit parameters for each model, while the dashed lines represent one standard deviation away, determined by the Monte Carlo procedure described in Section 3.1. (Panel A) Stability for notch ankyrin vs temperature in buffer. (Panel B) Enthalpy for notch ankyrin in buffer. (Panel C) Stability for hisactophilin in buffer. (Panel D) Stability for hisactophilin in buffer with 1 M urea. The insets in panels A, C, and D show the correlation of the midpoint enthalpy and entropy, in which the models are represented by the same colors as above. All data in the insets lies on top of the red scatter points; the blue and black points have been displaced for clarity. 1000 Monte Carlo instances have been generated for the inset plots. Bars indicate one standard deviation.

$C_{pn}(T)$, with $C_{pu}(T)$ and $C_{pn}(T)$ described above. The nonlinearity in the heat capacity is fixed in the model by the composition of the protein. The linear temperature-dependence of the native heat capacity is determined empirically from the fit to the data for each protein-cosolute system. Thus, when using this model to fit data, the free parameters in the heat capacity are the unfolding heat capacity at the transition temperature, ΔC_{pf}, and the linear coefficient to the temperature-dependence of the heat capacity of the native state, $\Delta C'_{pn}$. $\Delta C_p(T)$ is parametrized as $\Delta C_p = \Delta C_{pf} + C_{pu}(T) - C_{pu}(T_f) - \Delta C'_{pn}(T - T_f)$, where $C_{pu}(T)$ has the nonlinear T-dependence in eq 1.

In the approximation that ΔC_p is a linear function of temperature: $\Delta C_p = \Delta C_{pf} + \Delta C'_p(T-T_f)$, where $\Delta C'_p = \partial \Delta C_p/\partial T$, eqs 2a–2c become

$$\Delta H = \Delta H_f + \Delta C_{pf}(T - T_f) + \frac{\Delta C'_p}{2}(T - T_f)^2 \tag{3a}$$

$$\Delta S = \Delta S_f + \Delta C_{pf} \ln\left(\frac{T}{T_f}\right) + \Delta C'_p\left[T - T_f - T_f \ln\left(\frac{T}{T_f}\right)\right] \tag{3b}$$

$$\Delta G = \Delta H - T\Delta S \tag{3c}$$

The expressions for ΔH, ΔS, and ΔG in the limiting case of a T-independent unfolding heat capacity may be obtained by setting $\Delta C'_p = 0$ in (3a)–(3c); e.g. the unfolding free energy is

$$\Delta G = \left(1 - \frac{T}{T_f}\right)\Delta H_f + \left[T - T_f - T \ln\left(\frac{T}{T_f}\right)\right]\Delta C_{pf} \tag{4}$$

In Section 3.1, we examine the effect these approximations have on the parameters obtained from experimental data.

The probability p_u for the system to be unfolded in the two-state model is

$$p_u = 1/(1 + e^{\beta \Delta G}) \tag{5}$$

with ΔG given in eq 2c. Eq 5 can be equivalently written as $\Delta G = -k_B T \ln K_u$ where K_u is the unfolding equilibrium constant, given in the two-state model by $K_u = (1-p_u)/p_u$.

The total heat capacity in the two-state model is given (for example by differentiating $\langle H \rangle = H_n(1-p_u) + H_u p_u$ with respect to T) by

$$C_p(T) = \frac{\partial \langle H \rangle}{\partial T} = C_{pn} + \Delta C_p p_u + \Delta H \frac{\partial p_u}{\partial T}$$

Using the van't Hoff relation $\partial(\beta\Delta G)/\partial T = -\Delta H/(k_B T)^2$, we can write $\partial p_u/\partial T = p_u^2 e^{\beta \Delta G} \Delta H/(k_B T)^2$. Using the relation $\operatorname{sech}^2(x) = 2e^x/(1+e^{2x})$ we obtain finally

$$C_p(T) = C_{pu} - \Delta C_p + \frac{\Delta C_p}{1 + e^{\beta \Delta G}} + \frac{\Delta H^2}{4 k_B T^2} \operatorname{sech}^2\left(\frac{\Delta G}{2 k_B T}\right) \tag{6}$$

where C_{pu}, ΔC_p, ΔG, and ΔH are all temperature-dependent.

A plot of the Gibbs free energy vs temperature obtained from empirical data may be fit to eq 2c to obtain values of ΔH_f, ΔS_f, ΔC_{pf}, and $\Delta C'_p$. Similarly, a plot of the fraction of unfolded protein vs temperature may be fit to eq 5, or a plot of excess heat capacity may be fit to eq 6 to obtain values of these parameters. In all cases, once ΔH_f, ΔS_f, ΔC_{pf}, and $\Delta C'_p$ are obtained, eqs 2a and 2b can be used to obtain $\Delta H(T)$ and

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130–14144

Table 1. Thermodynamic Parameters for Several Proteins Used in Our Analysis and Comparison to Literature Data Where Available[e]

protein	model	ΔH_f (kJ/mol)	ΔS_f (kJ/mol/K)	ΔC_{pf} (kJ/mol/K)	$\Delta C_p'$ (kJ/mol/K^2)	$\Delta C_{pn}'$ (kJ/mol/K^2)
	T independent	297 ± 1	0.876 ± 0.003	4.38 ± 0.07	–[a]	–[a]
α-lactalbumin	T linear	304 ± 1	0.896 ± 0.004	2.88 ± 0.26	−0.101 ± 0.017	–[a]
	T nonlinear	304 ± 1	0.895 ± 0.004	2.93 ± 0.26	–[a]	0.065 ± 0.017
	reference value[38]	310	–[a]	5.3	−0.05	–[a]
	T independent	139 ± 3	0.454 ± 0.011	0.536 ± 0.5	–[a]	–[a]
arc repressor	T linear	118 ± 7	0.385 ± 0.02	1.45 ± 1.5	1.18 ± 0.44	–[a]
	T nonlinear	117 ± 7	0.384 ± 0.024	1.49 ± 1.4	–[a]	1.15 ± 0.4
creatine kinase[b]	T independent	780 ± 3	2.37 ± 0.01	25 ± 2	–[a]	–[a]
	T linear	733 ± 3	2.24 ± 0.01	66 ± 4	−73 ± 3	–[a]
	T independent	215 ± 7.4	0.659 ± 0.022	6.15 ± 0.75	–[a]	–[a]
hisactophilin	T linear	218 ± 5.9	0.669 ± 0.018	12.7 ± 1.7	−0.610 ± 0.17	–[a]
	T nonlinear	218 ± 17	0.667 ± 0.051	12.8 ± 3.3	–[a]	0.784 ± 0.28
	reference value[42]	226	–[a]	–[a]	–[a]	–[a]
	T independent	312 ± 2	0.929 ± 0.006	5.27 ± 0.7	–[a]	–[a]
HPr	T linear	314 ± 2	0.935 ± 0.006	4.41 ± 0.85	0.359 ± 0.150	–[a]
	T nonlinear	314 ± 2	0.935 ± 0.006	4.44 ± 0.80	–[a]	0.367 ± 0.14
	reference values[43]	316	0.941	6.0	–[a]	
	T independent	593 ± 9	1.86 ± 0.03	15.1 ± 0.5	–[a]	–[a]
notch ankyrin	T linear	602 ± 30	1.89 ± 0.1	16.2 ± 3.5	−0.045 ± 0.14	–[a]
	T nonlinear	601 ± 33	1.89 ± 0.1	15.7 ± 3.5	–[a]	0.114 ± 0.14
	reference value[41]	–[a]	–[a]	15.1	–[a]	–[a]
	T independent	496 ± 1	1.477 ± 0.002	14.3 ± 0.1	–[a]	–[a]
RNase A[b]	T linear	468 ± 1	1.396 ± 0.003	22.6 ± 0.2	0.89 ± 0.02	–[a]
	reference values[46c]	515	1.52	–[a]	–[a]	–[a]
	reference values[46d]	479	1.42	–[a]	–[a]	–[a]

[a]The symbol "–" represents not applicable for the respective model. [b]Literature data had background heat capacity subtracted for these proteins, so the nonlinear temperature-dependent model could not be applied. [c]Literature values obtained from differential scanning calorimetry. [d]Literature values obtained from spectroscopy measurements. [e]α-Lactalbumin, from heat capacity data in ref 38, arc repressor, from fraction of unfolded protein data in ref 44, creatine kinase from heat capacity data in ref 45, hisactophilin, from stability data in ref 42, histidine-containing phosphocarrier protein (HPr), from fraction of unfolded protein data in ref 43, notch ankyrin, from stability data in ref 41, and rNase A from heat capacity data in ref 46. Values obtained from the three models of the temperature-dependence of ΔC_p are given, as well as the values obtained from the appropriate reference where available. The reference value in ref 41 assumed a temperature-independent ΔC_p, the value of ΔH_f from ref 42 was obtained by integrating C_p up to T_f, and the values from ref 43 were obtained assuming a temperature-independent ΔC_p.

$\Delta S(T)$ at various temperatures. In Section 3.1 we will compare the best-fit values for the three models of ΔC_p described above, i.e. eqs 2a–2c, 3a–3c, and the temperature-independent ΔC_p model (*cf.* eq 4). This procedure can be performed at various cosolute concentrations, providing experimental data is available. Then the *changes* in unfolding enthalpy $\delta\Delta H(T, c)$ and entropy $\delta\Delta S(T, c)$ upon transfer from a solution of cosolute concentration 0 to one of concentration c at a given temperature T can be determined.

We also define the change in the midpoint parameters at each respective cosolute concentration, nonzero and zero, upon transfer:

$$\delta\Delta H_f(c) \equiv \Delta H(T_f(c), c) - \Delta H(T_f(0), 0) \quad (7a)$$

$$\delta\Delta S_f(c) \equiv \Delta S(T_f(c), c) - \Delta S(T_f(0), 0) \quad (7b)$$

$$\delta\Delta C_{pf}(c) \equiv \Delta C_p(T_f(c), c) - \Delta C_p(T_f(0), 0) \quad (7c)$$

In what follows we will often drop the explicit concentration-dependence in writing various thermodynamic equalities when it is unambiguous.

3. RESULTS

3.1. Monte Carlo Method To Determine Statistical Errors. To analyze the uncertainty involved in fitting the data, we perform a Monte Carlo procedure. We fit a given data set, such as $C_p(T)$, $\Delta G(T)$, or $p_u(T)$, to eqs 6, 2c, or 5, respectively. We then generate a large number of sample data sets by drawing each point from a normal distribution with a mean equal to the value of the best fit curve at that point and a standard deviation equal to the root-mean-square of the residual. Each of these generated sample data sets is then fitted and new fit parameters are obtained, thus generating a *distribution* of values for ΔH_f, ΔS_f, and, depending on the model, either ΔC_p, ΔC_{pf} and $\Delta C_p'$, or ΔC_{pf} and $\Delta C_{pn}'$. We can fit the different models for ΔC_p described in Section 2.1 to compare the parameters extracted. The uncertainty in the thermodynamic parameters could also be obtained by

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130–14144

examining the covariance matrix of the fit parameters, but the Monte Carlo method we use generates sets of data that may be bootstrapped to obtain statistical significance measurements without the assumptions inherent in keeping only the covariance and truncating all higher order moments. The bootstrapped distributions we find below are generally nongaussian.

As an example of fitting the stability $\Delta G(T)$, we have used experimental measurements by Zweifel and Barrick of the thermal denaturation of notch ankyrin in various concentrations of urea.[41] The best fit to the 0 M data in ref 41 (plotted as green circles in panel A of Figure 1) yields a root-mean-square of the residuals of 0.383 kJ/mol, so the square of this becomes the variance of the normal random distribution centered around the value of the best fit curve. From this we generate 1000 sample data sets, each of which is fit to either eq 2c, 3c, or 4, depending on the model. In Table 1, we compare the parameters obtained with the three different models of ΔC_p described above. We see that the parameters are consistent with each other and, for the case of notch ankyrin, with the tabulated value in ref 41. We analyze the variances and correlations of this data; this is reported in Table 2. We see that all parameters are generally strongly correlated or strongly anticorrelated. Figure 1 panels A,B show that the three models give similar curves even when extrapolating to high temperatures, though the variance in ΔG at high T is significantly smaller for the T-independent ΔC_p model than the other two models considered.

Table 2. Comparison of the Variance and Covariance of Fits to ΔG vs T Data for Notch Ankyrin from Ref 41, for the Three Models of the Temperature-Dependence of ΔC_p Discussed[a]

			ΔH_f	ΔS_f	ΔC_p
T-independent ΔC_p		ΔH_f	0.016	0.9997	0.989
		ΔS_f		0.017	0.989
		ΔC_p			0.033
		ΔH_f	ΔS_f	ΔC_{pf}	$\Delta C'_p$
T-linear ΔC_p	ΔH_f	0.055	0.99997	0.987	−0.960
	ΔS_f		0.056	0.988	−0.959
	ΔC_{pf}			0.22	−0.991
	$\Delta C'_p$				3.0
		ΔH_f	ΔS_f	ΔC_{pf}	$\Delta C'_{pn}$
T-nonlinear ΔC_p	ΔH_f	0.055	0.99997	0.987	−0.958
	ΔS_f		0.056	0.988	−0.957
	ΔC_{pf}			0.22	−0.990
	$\Delta C'_{pn}$				1.20

[a]For each model, a matrix is given in which the diagonal elements are the relative deviations for that quantity, and the off-diagonal elements are the correlation coefficients for the two quantities. Relative deviation for e.g. ΔH_f is defined as $\langle \Delta H_f^2 \rangle - \langle \Delta H_f \rangle^2)^{1/2}/\langle \Delta H_f \rangle$, where averages are over the Monte Carlo generated data. In all models the entropy and enthalpy of unfolding are highly correlated.

Figure 1 panels C and D plot data from ref 42 for the stability ΔG vs T for hisactophilin in 0 and 1 M urea. The data for 1 M urea includes both hot denaturation and cold denaturation regions. Comparing panels C and D of Figure 1 we can see that the *model* variance is much less for the 1 M urea data, in that all three models predict similar curves, presumably as a result of having a larger range of $\Delta G(T)$ data to fit. The large uncertainty in the nonlinear T-dependent model for the 1 M urea data is likely caused by fitting a more flexible model to a limited set of data.

Interestingly, there is a change in sign of the curvature at low temperatures in panel C of Figure 1 for the nonlinear T-dependent model. This effect is caused by a change in sign of ΔC_p at around 310 K. A similar effect is seen at high temperatures for some generated sets of data in the nonlinear T-dependent model at 1 M urea (Figure 1D).

We have performed the same analysis on data measuring the fraction of unfolded protein as a function of temperature. The data examined is for histidine-containing phosphocarrier protein (HPr), from ref 43, and for arc repressor, from ref 44. The midpoint parameters ΔH_f and ΔS_f are again in agreement between the three models (see Table 1), but the way that the different heat capacity models extrapolate quantities such as the stability and the enthalpy is markedly different, as it was for hisactophilin — see Figure 2 for HPr.

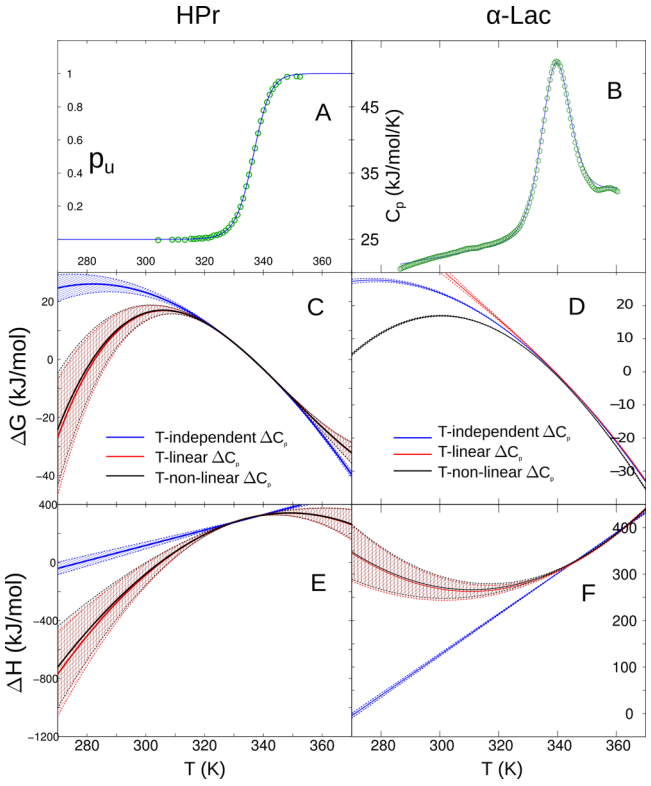


Figure 2. Analyzing unfolding fraction data and heat capacity data. Panels A and B show fraction of unfolded population data and heat capacity data from refs 43 and 38, respectively (green circles), along with best fit curves. Panels C and D show the corresponding stability as a function of temperature, and panels E and F show the corresponding unfolding enthalpy as a function of temperature. In all panels the blue lines are fits for the T-independent ΔC_p model, the red lines are fits for the linear T-dependent model, and the black lines are fits for the nonlinear T-dependent model. The solid lines arise from the best fit parameters for each model, while the dashed lines represent one standard deviation away, determined by the Monte Carlo procedure described in Section 3.1.

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

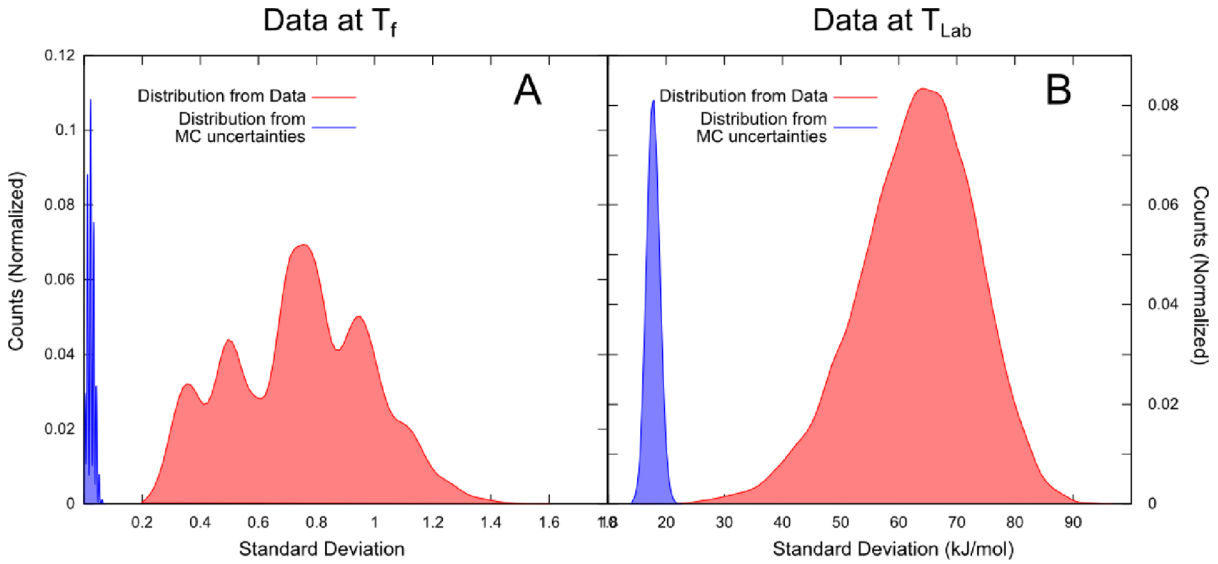


Figure 3. Distribution in standard deviation σ of the data arising from the bootstrapping method described in section 3.2.2. The overlap of the two distributions represents the statistical significance of entropy-enthalpy compensation in this case. (A) (Red) Distribution of σ obtained from bootstrapping the protein data at their corresponding melting temperatures; (Blue) Distribution of σ obtained from bootstrapping the Monte Carlo generated data representing the fit uncertainty. The probability of obtaining a σ from the null distribution (the statistical significance of S−H compensation) is smaller than $p < 1 \times 10^{-7}$. (B) Corresponding quantities at laboratory temperature. The statistical significance of S−H compensation is $p = 5 \times 10^{-5}$.

We have also performed the same analysis on heat capacity vs T data for α-lactalbumin, from ref 38. The midpoint parameters ΔH_f and ΔS_f are in agreement between the models (Table 1). Again, however, the way that the models extrapolate stability and enthalpy is very different (Figure 2, panels B, D, F).

In all the cases we have examined, the values of the unfolding entropy and enthalpy at the transition midpoint are robust across all three models. Further, the value ΔC_{pf} agrees within uncertainty for all the cases we looked at between the linear T-dependent ΔC_p model and the nonlinear T-dependent ΔC_p model. Fitting experimental data to a temperature-independent ΔC_p model will be sufficient, if only midpoint parameters are required and the accuracy of the unfolding heat capacity ΔC_{pf} is not particularly important. However, Figures 1 and 2 indicate that such data is prone to significant extrapolation errors.

3.2. Transfer Entropy and Enthalpy for Various Proteins and Solvents. *3.2.1. Transfer Entropy and Enthalpy at the Transition Midpoint.* The above analysis indicates that the thermodynamic parameters obtained by fitting experimental data are most accurately determined near the transition midpoint. We thus now examine the entropy and enthalpy of transfer for various proteins at their transition midpoints, from water to water plus various cosolutes, cf. eqs 7a and 7b.

For a number of proteins, thermodynamic data exists for more than one cosolute; as well, for a number of cosolutes thermodynamic data exists for more than one protein. These commonalities and differences enable useful comparisons.

Because $\Delta H_f = T_f\ \Delta S_f$, the folding temperature T_f is unchanged upon transfer to cosolute if $\delta\Delta H_f/\Delta H_f^0 = \delta\Delta S_f/\Delta S_f^0$ where ΔH_f^0 and ΔS_f^0 are the midpoint enthalpy and entropy of unfolding at cosolute concentration $c = 0$, respectively, and $\delta\Delta H_f$ and $\delta\Delta S_f$ are defined in eqs 7a and 7b. Thus, on a plot with $\delta\Delta H_f/\Delta H_f^0$ and $\delta\Delta S_f/\Delta S_f^0$ on the ordinate and abscissa, a line of slope unity would constitute no change in folding temperature due to the cosolute. Further, since

$$T_f^0 + \delta T_f = \frac{\Delta H_f^0 + \delta\Delta H_f}{\Delta S_f^0 + \delta\Delta S_f} \approx T_f^0 + T_f^0\left(\frac{\delta\Delta H_f}{\Delta H_f^0} - \frac{\delta\Delta S_f}{\Delta S_f^0}\right)$$

we conclude that the distance of each point from the line $y = x$ is, to first order, the relative change in the folding temperature, $\delta T_f/T_f^0$, as a result of the transfer.

We may further interpret the deviation from the line $y = x$ on a plot of $\delta\Delta S_f/\Delta S_0$ vs $\delta\Delta H_f/\Delta H_0$ as the free energy of transfer at fixed temperature T_f^0, $\delta\Delta G(T_f^0)$, divided by the unfolding enthalpy ΔH_f^0. That is, $\delta\Delta H_f/\Delta H_f^0 - \delta\Delta S_f/\Delta S_f^0 = \delta\Delta G(T_f^0)/\Delta H_f^0$.

In Figures 4 and 5, we divide both $\delta\Delta H_f/\Delta H_f$ and $\delta\Delta S_f/\Delta S_f$ by the concentration c of the cosolute, to compare cosolute solutions having different concentrations. Thus, the axes in Figures 4 and 5 can be thought of as a decomposition of m-values[47] into enthalpic and entropic components, each normalized by the corresponding unfolding enthalpy or entropy in the absence of cosolute.

Linear regression to the data in Figures 4 and 5, when taken together gives a slope of 0.99 ± 0.04. The statistical test outlined in Krugg et al.[19] requires that the slope of the best fit $\delta\Delta H_f/\Delta H_f^0$ vs $\delta\Delta S_f/\Delta S_f^0$ line be more than 2σ away from the harmonic mean of the temperatures at which the experiments were performed; thus for Figure 4 this requires that the slope be 2σ away from unity. Figure 4 fails this test. Nevertheless, we show that the results in Figure 4 are in fact statistically significant, given the magnitude of the experimental errors. We now describe a treatment of the statistical significance of entropy-enthalpy compensation that is valid when the slope of the ΔH-ΔS plot is near unity.

3.2.2. Statistical Significance of Entropy-Enthalpy Compensation. The Monte Carlo method in Section 3.1 can be applied to data at various concentrations of cosolutes to assess the significance of the linear relationship between enthalpy and entropy observed in Figures 4 and 5. Fitting each experimental data set to the appropriate equation as described in Section 3.1

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

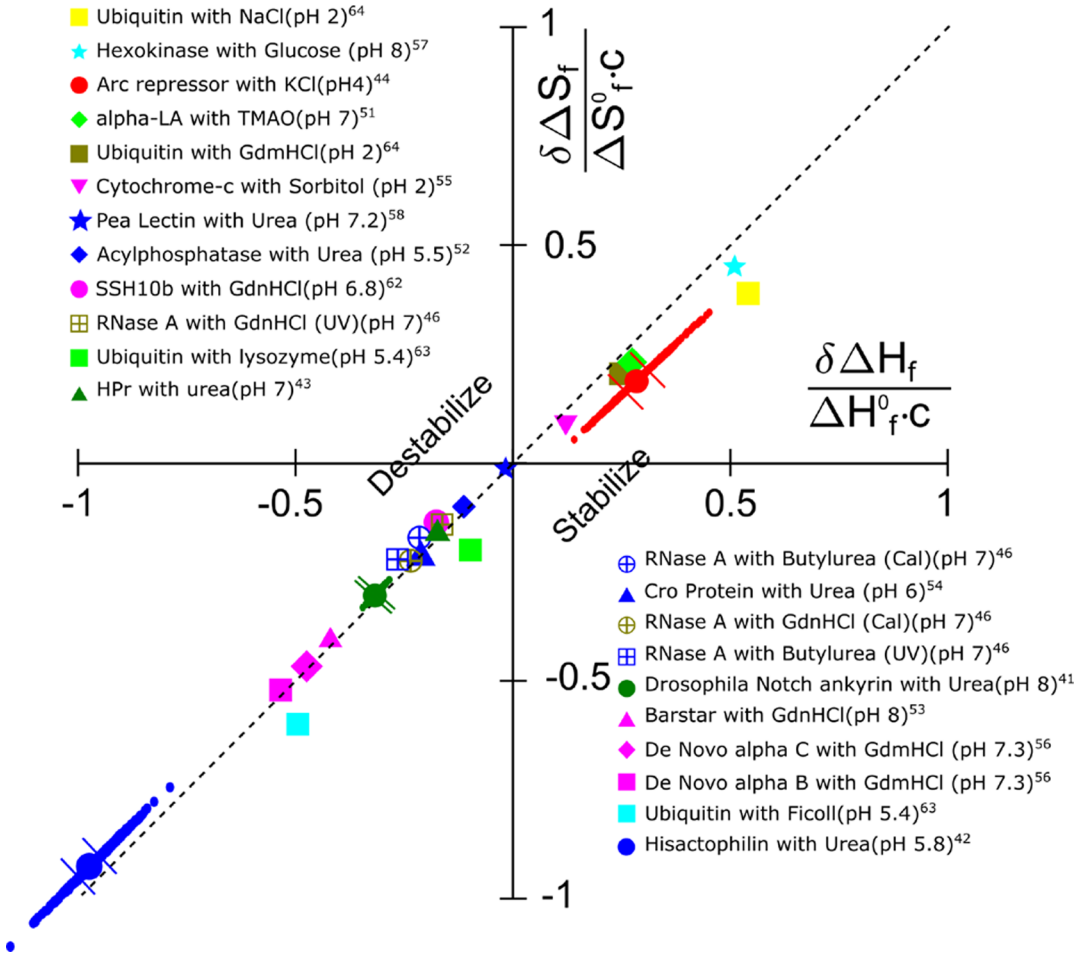


Figure 4. Entropy-enthalpy compensation for protein unfolding transfer enthalpy $\delta\Delta H_f/c$ and unfolding transfer entropy $\delta\Delta S_f/c$, both evaluated at the folding midpoint and suitably normalized as described below. The legend, listed from the upper right data point to the lower left data point, indicates the protein, cosolute, pH, and corresponding source of the experimental data. Cosolutes above and to the left of the diagonal are destabilizing as noted; cosolutes below and to the right of the diagonal are stabilizing. Abscissa/ordinate are the transfer enthalpy/entropy normalized by the unfolding enthalpy/entropy in the absence of solute, per 100 mg/mL of cosolute, i.e. $\delta\Delta S_f/(\Delta S_f c)$ vs $\delta\Delta H_f/(\Delta H_f c)$. Also plotted here are the Monte Carlo-generated scatter points for arc repressor in KCl (red circle), notch ankyrin in urea (green circle), and hisactophilin in urea (blue circle). Bars on each of these three points show the standard deviation in the direction of the scatter. The scatter here does not substantially reduce the significance of the linear compensating trend. See also Table 3, which gives thermodynamic parameters for the proteins we study here.

results in a set of best fit parameters, as well as a set of residuals from the best fit. For the following protein/cosolute systems — α-lactalbumin in ethanol,[38] arc repressor in KCl,[44] creatine kinase in glycerol,[45] hisactophilin in urea,[42] histidine containing phosphocarrier in urea,[43] notch ankyrin in urea,[41] and RNase A in urea[46] — 1000 data sets were generated as described in Section 3.1, for each of several concentrations of cosolute. Thus, 1000 values for $\delta\Delta H_f$ and $\delta\Delta S_f$ were generated. We plot the results of this procedure as scatter points in Figures 4–6.

From the extent of the scatter for these 7 data points in Figures 4 and 5, we can assess the significance of the apparent linear relationship in the plot. The analysis rests on the assumption that the deviations from the line of slope unity, $\delta T_f = 0$, are much smaller than the deviations from zero of either $\delta\Delta H_f$ or $\delta\Delta S_f$, i.e., that the data are essentially distributed along the diagonal. We may then consider the data as transformed to a coordinate system that is rotated $\pi/4$ counterclockwise and translated so that the origin coincides with the mean of the data. Then the data consists approximately of points distributed along the abscissa all having zero ordinate. If the variance of this data is large compared to what would be expected from the experimental error as derived from the above Monte Carlo method, then the result of entropy-enthalpy compensation is statistically significant.

To assess the statistical significance, we use a *bootstrapping* method that avoids requiring any assumptions for the distribution of the data points along the $y = x$ line. From the 48 data points in Figures 4 and 5 we perform random sampling with replacement to generate new sets of 48 data points. We find the standard deviation of each of these generated sets and thus obtain a distribution for the standard deviations σ from the bootstrapped data. Then, we obtain another distribution for the values of σ that would result from sampling the Monte Carlo-generated deviations. That is, we had generated 1000 Monte Carlo points for each of the above 7 proteins. For each protein, we subtract the mean from the value of all 1000 data points, yielding 7000 values of deviations from the mean. From this set of 7000 total generated deviations, we repeatedly sample with replacement to obtain sets of 1000 deviations, and we calculate the standard deviation for each of these samples. This yields another distribution of standard deviations σ, now for the Monte Carlo-generated data consistent with the null

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130–14144

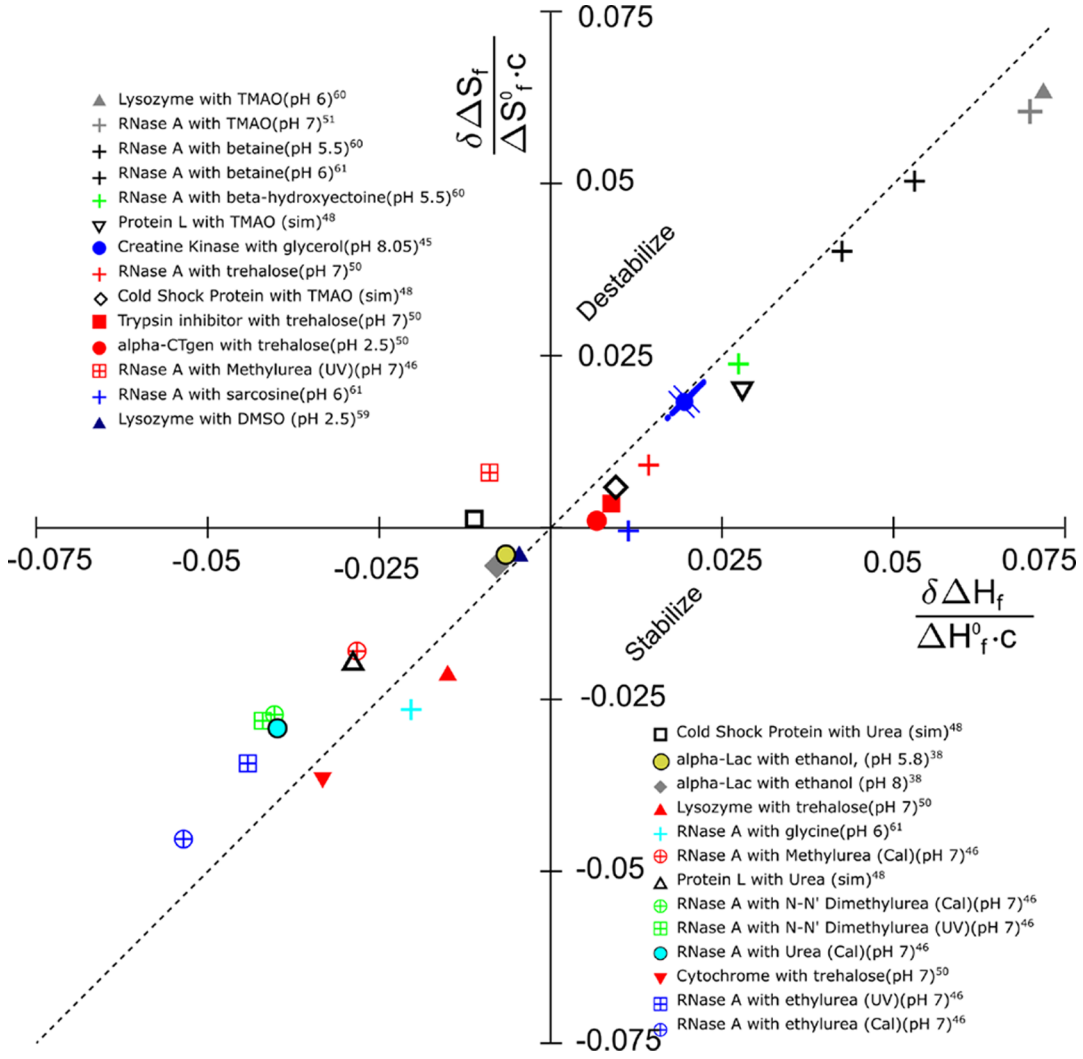


Figure 5. Further illustration of entropy-enthalpy compensation for various proteins and solvents. The notation here is the same as in Figure 4, but the scale of the plot is significantly smaller. Scatter as a result of uncertainty for creatine kinase in glycerol (blue circle) is shown, along with bars to indicate the standard deviation. Scatter was also calculated for RNase A in urea (cyan circle) and α-lac in ethanol (mustard circle), but the scatter is smaller than the data points appearing on this plot. Black open symbols correspond to simulation data using the Tanford transfer model taken from O'Brien et al. (ref 50). Legend labels are ordered from upper right data point to lower left data point.

hypothesis. Both distributions for σ are plotted in Figure 3A. The overlap of these two distributions then provides the statistical significance — the likelihood that the scatter in the experimental data points arises from the fit uncertainty of the thermodynamic parameters. For the data in Figures 4 and 5 together, we obtain $p < 10^{-7}$; i.e. in 10^7 samplings we did not observe any instances of overlap between the distributions. The data plotted in Figures 4 and 5 thus illustrates entropy-enthalpy compensation rather than experimental error.

3.2.3. Dominance of the Enthalpy in Protein Transfer. An interesting observation may be made regarding the data in Figures 4 and 5. For almost all the data points in these figures, the magnitude of the enthalpy change is larger than the magnitude of the entropy change. In fact, we did not observe any system with a destabilizing cosolute that had a positive change in entropy of unfolding larger than the change in enthalpy of unfolding (upper right quadrants of Figures 4 and 5), and only a few systems showed a stabilizing cosolute with a change in entropy of unfolding larger than the change in enthalpy of unfolding (lower left quadrants of Figures 4 and 5). Thus, for the most part, enthalpy can be thought of as driving the change in stability, while entropy tries to catch up and partially compensates. This finding is consistent with recent studies stressing the dominance of interactions over steric repulsion for crowding agents such as Ficoll, BSA, and lysozyme[48] and glucose, dextran, and poly(ethylene glycol).[49]

3.2.4. Tanford Model Predictions from Simulation Data. Figure 5 also shows the transfer enthalpy and entropy for simulation results by O'Brien et al. on Cold shock protein and protein L[50] (open black symbols in Figure 5). Here, thermodynamic parameters were extracted from fits to simulated heat capacity curves, for the transfer of the above proteins to either urea or TMAO. A surface-area based Tanford transfer model[51] was used to model the cosolute solution. We have not found experimental values for these thermodynamic parameters in the literature; the values in Figure 5 are thus predictions as a consequence of both the simulation method for generating unfolded ensembles and the Tanford transfer model, which are subject to experimental test.

3.2.5. Importance of the Cosolute- and Temperature-Dependence of the Unfolding Heat Capacity. The unfolding heat capacity $\Delta C_p(T, c)$ is generally both temperature- and

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

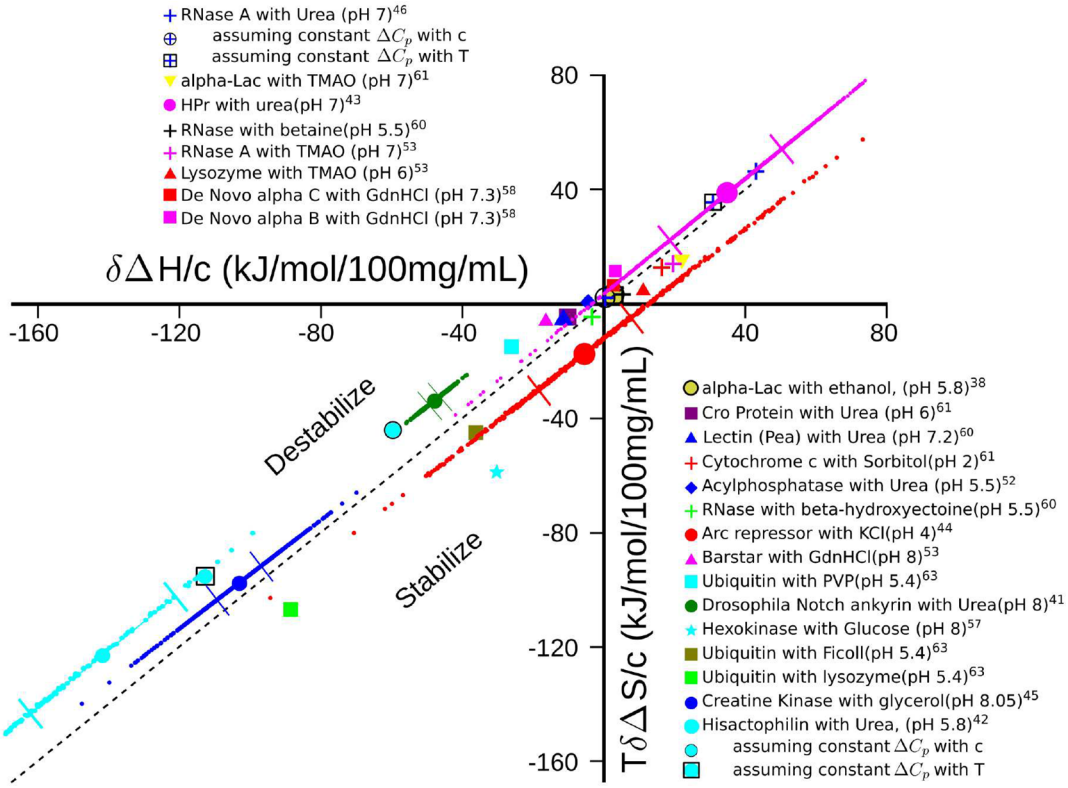


Figure 6. Entropy-enthalpy compensation for the transfer of various proteins to various solvents is also seen by plotting $\delta\Delta H$ vs $T\delta\Delta S$ at lab temperature (25 °C). Data are normalized by the concentration c = 100 mg/mL. The points cluster close to the $\delta\Delta H = T\delta\Delta S$ line; the deviation from that line (horizontal or vertical) represents the absolute change in stability upon transfer at 25 °C. Points above the line correspond to destabilizing cosolutes, points below the line correspond to stabilizing cosolutes. Scatter points representing the range of uncertainty in obtaining the enthalpy and entropy are also shown, as determined by the Monte Carlo method described in Section 3.1. The scatter is highly correlated with a magnitude that in some cases is large enough to change the sign of $\delta\Delta H$ and $T\delta\Delta S$. The compensation is statistically significant however—see Section 3.2.7. In the case of α-lactalbumin in ethanol, the scatter was smaller than the symbol. The cyan circle with black outline indicates the hisactophilin data assuming ΔC_p is independent of the concentration of urea but has nonlinear temperature-dependence obtained by fitting to eq 2c; the cyan circle with black square outline is the value obtained assuming ΔC_p is independent of temperature but still accounting for the concentration-dependence. These approximations both introduce significant error: ≈80 and 30 kJ/mol, respectively. Similarly, assuming a concentration-independent unfolding heat capacity introduces an error of ≈40 kJ/mol for RNase A in urea (circled blue cross), and a temperature-independent ΔC_p introduces an error of ≈10 kJ/mol. Legend labels are ordered approximately from upper right data point to lower left data point.

concentration-dependent. We define the change in unfolding heat capacity upon transfer, $\delta\Delta C_{pf}$, in eq 7c. While the quantities $\delta\Delta H_f$ and $\delta\Delta S_f$ in eqs 7a and 7b are independent of $\delta\Delta C_{pf}$, since they are always evaluated at the respective transition temperatures, the thermodynamics for transfer at fixed temperature (e.g., lab temperature) does depend on $\delta\Delta C_{pf}$.

A number of the works cited here obtained a concentration-independent ΔC_p however, by equating ΔC_p to the slope of ΔH_f vs T_f data for various cosolute concentrations. This assumes that ΔC_p is constant with varying cosolute concentration, and hence $\delta\Delta C_{pf} = 0$. These proteins/cosolutes are thus indicated in Table 3 by "0[g]" in the column for $\delta\Delta C_{pf}$. This assumption may be sufficient if $\delta\Delta H_f$, $\delta\Delta S_f$, or $\Delta C_p(c = 0)$ is the quantity of interest; however, when eqs 2a and 2b are used to evaluate thermodynamic parameters at lab temperature, this assumption produces unacceptably large errors. Examples of the change in thermodynamic values obtained by setting $\delta\Delta C_{pf} = 0$ are shown in Figure 6, for the transfer of both hisactophilin and RNase A to urea: neglecting $\delta\Delta C_p$ changes the resulting value of $\delta\Delta H$ by ≈80 kJ/mol for hisactophilin and ≈40 kJ/mol for RNase A. This is to be compared with the error introduced by neglecting the temperature-dependence of ΔC_p, which was ≈30 kJ/mol for hisactophilin and ≈10 kJ/mol for RNase A. Recent work by Senske et al. has shown that accounting for the temperature-dependence of the heat capacity can reverse the sign of the unfolding enthalpy and entropy upon transfer to some cosolutes such as glucose and dextran.[49]

In Table 3 we have made a note of where the $\delta\Delta C_p = 0$ assumption has been made, and we have not plotted the corresponding $\delta\Delta H_{lab}$ and $T\delta\Delta S_{lab}$ data in Figure 6. Note that the potential for large errors due to $\delta\Delta C_p$ is irrelevant when comparing data at the respective folding temperatures, i.e. for the quantities in eqs 7a and 7b and Figures 4 and 5.

Figure 7 plots the concentration-dependence of ΔC_{pf} for several protein-cosolute systems, obtained by using the nonlinear temperature-dependent model in eq 1. The values plotted do not change significantly if the linear temperature-dependent model is used (see for example Table 1, which shows that the values obtained from the two models are comparable). Some proteins have a ΔC_{pf} showing weak concentration-dependence (e.g., barstar in GdmHCl, acylphosphatase in urea), while for others, ΔC_{pf} shows significant concentration-dependence (RNase A in urea, α-lactalbumin in ethanol).

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

Table 3. Empirical Values for Thermodynamic Unfolding Parameters upon Transfer of Various Proteins to Various Solvents[p]

protein	cosolute	pH	ΔH_f^{0}[b]	ΔS_f^{0}	T_f	ΔC_{pf}	$\frac{\delta\Delta H_f}{\Delta H_f^{0}}$	$\frac{\delta\Delta S_f}{\Delta S_f^{0}}$	$\delta\Delta C_{pf}$	δT_f	$\delta\Delta H_{lab}$	$T\delta\Delta S_{lab}$	ref/Figure/Table[a]
α-CTgen[h]	trehalose	2.5	403	1.27	317	4.4	6.7e-3	9.7e-4	0[g]	1.8	2.69[g]	0.37[g]	57 Table 1
α-lactalbumin[o]	ethanol	8	310	0.916	338	5.3	−8.2e-3	−5.9e-3	0.10	19	1.11	1.67	38 Figure 3[e]
α-lactalbumin	TMAO	7	209	0.663	315	6.5	0.27	0.25	−0.133	5.0	22.5	17.4	58 Table 1
acylphosphatase	urea	5.5	351	1.06	331	6.2	−0.11	−0.10	−0.037	−3.7	−4.17	1.07	59 Table 2
arc repressor[o]	KCl	4	141	0.463	304	1.0	0.25	0.14	5.20	29	−89.1	−102	44 Figure 7[c]
barstar	GdnHCl	8	292	0.849	343	6.2	−0.43	−0.41	−1.9	−12	−16.2	−5.64	60 Table 2
Cro protein	urea	6	195	0.591	330	3.8	−0.22	−0.20	0.09	−8.2	−7.47	−3.45	61 Table 2
cytochrome c	sorbitol	2	226	0.740	305	5.2	0.12	0.09	−1.7	8.4	13.6	7.28	62 Table 1
cytochrome c	trehalose	7	161	0.502	321	N/A	−3.4e-2	−3.7e-2	0[g]	24	−5.38[g]	−4.78[g]	57 Table 1
creatine kinase[o]	glycerol	8.05	782	2.38	329	92	2.0e-2	1.9e-2	2.46	0.32	−93.1	−88.4	45 Figure 4[e]
de novo α B	GdnHCl	7.3	103	0.300	343	2.3	−0.53	−0.52	0.21	−7.2	0.564	4.52	63 Table 2
de novo α C	GdnHCl	7.3	153	0.441	347	2.7	−0.48	−0.47	−0.07	−6.5	3.94	10.5	63 Table 2
hisactophilin[o]	urea	5.8	215	0.658	327	6.1	−0.89	−0.85	3.63	−87	−142	−123	42 Figure 6[d]
hexokinase	glucose	8	700	2.19	320	30	0.51	0.46	1.5	11	−30.0	−58.1	64 Table 2
HPr[i,o]	urea	7	315	0.935	336	4.4	−0.16	−0.16	−0.0054	−27.1	36.4	39.7	43 Figure 2[c]
lectin (pea)	urea	7.2	1130	3.25	347	22	−1.6e-2	−1.1e-2	0.74	−1.8	−7.12	−3.40	65 Table 2
lysozyme	DMSO	2.5	535	1.58	339	7.8	−4.4e-3	−3.4e-3	0[g]	−034	1.00[g]	1.50[g]	66 Table 2
lysozyme	trehalose	7	397	1.20	331	N/A	−1.5e-2	−2.2e-2	0[g]	2.4	−5.99[g]	−7.09[g]	57 Table 1
lysozyme	TMAO	6	535	1.50	357	6.8	7.5e-2	6.5e-2	0.089	3.3	12.3	5.59	58 Table 1
notch ankyrin[o]	urea	8	592	1.86	318	15	−0.45	−0.43	−1.26	−11	−63.2	−43.8	41 Figure 4[d]
RNase A	β-hydrox[j]	5.5	364	1.09	334	4.4	2.8e-2	2.4e-2	0.194	1.3	−3.16	−4.36	67 Table 1
RNase A	betaine	5.5	364	1.09	334	4.4	4.2e-2	4.1e-2	0.214	0.32	5.39	3.42	67 Table 1
RNase A	betaine	6.0	364	1.09	334	0	5.3e-2	5.0e-2	0[g]	0.95	0[g]	−0.291[g]	68 Figures 2,4[f]
RNase A	trehalose	7	385	1.20	321	4.7	1.4e-2	9.1e-3	0[g]	1.6	5.51[g]	3.38[g]	57 Table 1
RNase A	glycine	6.0	364	1.09	334	0	−2.0e-2	−2.7e-2	0[g]	2.4	0[g]	−0.659[g]	68 Figures 2,4[f]
RNase A	sarcosine	6.0	364	1.09	334	0	1.1e-2	−4.0e-4	0[g]	3.8	0[g]	−1.20[g]	68 Figures 2,4[f]
RNase A	TMAO	7	490	1.46	336	5.2	7.3e-2	6.2e-2	0.022	3.5	16.1	9.35	58 Table 1
RNase A[m]	GdnHCl	7	515	1.53	337	11	−0.23	−0.22	0[g]	−4.3	−0.246[g]	10.8[g]	46 Table 1
RNase A[n]	GdnHCl	7	452	1.35	335	6.3	−0.16	−0.14	0[g]	−7.8	0.44[g]	12.3[g]	46 Table 1
RNase A[m]	methylurea	7	515	1.53	337	7.1	−5.3e-2	−4.5e-2	0[g]	−2.8	−0.508[g]	4.19[g]	46 Table 1
RNase A[n]	methylurea	7	452	1.35	335	4.8	−4.4e-2	−3.4e-2	0[g]	−3.5	−0.419[g]	4.28[g]	46 Table 1
RNase A[m]	dimethylurea[k]	7	515	1.53	337	3.7	−2.8e-2	−1.8e-2	0[g]	−3.4	−0.08[g]	5.04[g]	46 Table 1
RNase A[n]	dimethylurea[k]	7	452	1.35	335	1.2	−8.7e-2	3.2e-2	0[g]	−39	0.79[g]	5.64[g]	46 Table 1
RNase A[m]	ethylurea	7	515	1.53	337	3.6	−4.0e-2	−2.7e-2	0[g]	−4.5	0.13[g]	6.02[g]	46 Table 1
RNase A[n]	ethylurea	7	452	1.35	335	3.9	−4.2e-2	−2.7e-2	0[g]	−5.2	1.05[g]	8.25[g]	46 Table 1
RNase A[m]	butylurea	7	515	1.53	337	6.0	−0.21	−0.17	0[g]	−16	4.3[g]	23.4[g]	46 Table 1
RNase A[n]	butylurea	7	452	1.35	335	7.3	−0.26	−0.21	0[g]	−21	7.4[g]	30.3[g]	46 Table 1
RNase A[o]	urea	7	501	1.49	336	14.6	−4.0e-2	−2.9e-2	−1.5	−3.8	43.0	46.6	46 Figure 1[e]
Ssh10b	GdnHCl	6.8	307	0.840	365	N/A	−0.18	−0.14	0[g]	−17	−54.4[g]	−34.3[g]	69 Table 2
tryps inh[l]	trehalose	7	236	0.711	332	1.1	8.8e-3	3.4e-3	0[g]	1.8	2.08[g]	0.74[g]	57 Table 1
ubiquitin	PVP	5.4	100	0.265	377	1.5	−0.2	−0.2	−0.1	0	−26.0	−14.8	48 Table 3
ubiquitin	Ficoll	5.4	100	0.265	377	1.5	−0.5	−0.6	−0.5	94	−36.1	−44.5	48 Table 3
ubiquitin	BSA	5.4	100	0.265	377	1.5	0.5	0.5	1.6	0	−310	−295	48 Table 3
ubiquitin	lysozyme	5.4	100	0.265	377	1.5	−0.1	−0.2	0.1	47	−88.1	−107	48 Table 3
ubiquitin	NaCl	2	203	0.617	329	3.0	0.547	0.383	0[g]	39	111[g]	70.4[g]	70 Table 1
ubiquitin	$CaCl_2$	2	203	0.617	329	4.6	3.21	2.67	0[g]	48	652[g]	490[g]	70 Table 1
ubiquitin	$MgCl_2$	2	203	0.617	329	4.4	1.87	1.49	0[g]	50	379[g]	273[g]	70 Table 1
ubiquitin	GdmCl	2	203	0.617	329	4.7	0.248	0.201	0[g]	13	50.3[g]	37.0[g]	70 Table 1

[a]Literature reference and corresponding tabulated value or figure used to extract the values listed here. [b]Values listed for ΔH_f^{0} and ΔS_f^{0} are the unfolding enthalpy and entropy at T_f, 0 M cosolute concentration, and pH indicated in the corresponding column. Throughout this table, values for enthalpy, entropy, and heat capacity are given in kJ/mol, kJ/mol/K, and kJ/mol/K, respectively. [c]Fit to eqs 2c and 5. [d]Fit to eq 2c. [e]Fit to eq 6. [f]Used $\Delta H(T_f) = T_f\,\Delta S(T_f)$ to obtain unfolding entropy. N/A: ΔC_p data not available. [g]ΔC_p measured for one concentration and assumed constant with respect to concentration in these references. The values for $\delta\Delta H_{lab}$ and $T\delta\Delta S_{lab}$ for these systems are thus likely inaccurate. [h]α-Chymotrypsinogen. [i]Histidine-containing phosphocarrier protein. [j]β-Hydroxyectoine. [k]N,N'-Dimethylurea. [l]Trypsin inhibitor. [m]Measurements taken with calorimetry. [n]Measurements taken with UV absorption. [o]Monte Carlo error analysis is performed on this protein in Figures 4, 5, and 6. [p]For each protein-cosolute system, the values are listed at the concentration corresponding to (100 mg/mL). The literature reference from which the values were obtained, along with the corresponding figure or table in that reference, is listed in the last column. For systems in which the parameters were obtained through fitting a curve to the data in the reference work, the equation from this work used for the fitting is also referenced in the last column. Systems without an equation listed had ΔH_f and ΔS_f values directly available.

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

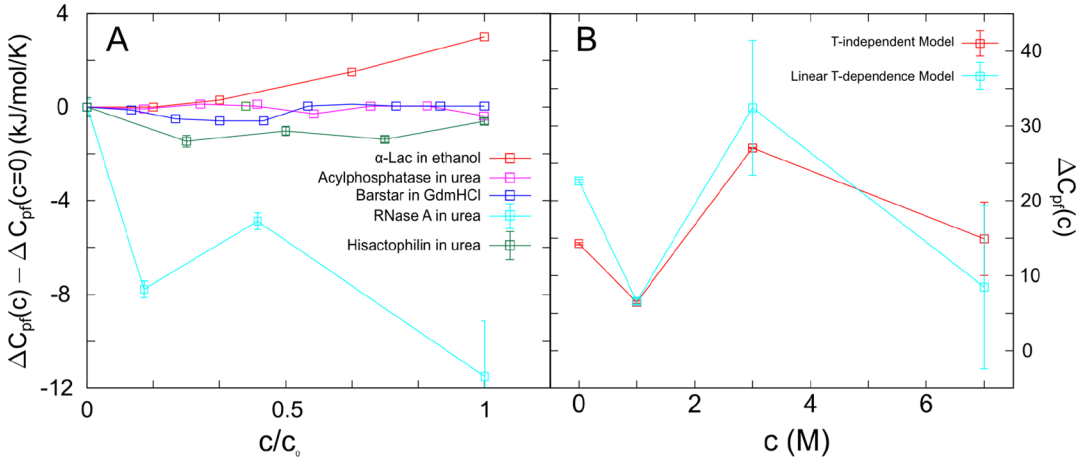


Figure 7. (A) Concentration-dependence of the heat capacity for several protein-cosolute systems. For RNase A in urea, hisactophilin in urea, and α-lactalbumin in ethanol, error bars were determined from the Monte Carlo method described in Section 3.1. Error bars are not present for acylphosphatase or barstar because the corresponding literature data were not available for application of the Monte Carlo method. The abscissa was normalized by the maximal concentration of cosolute in the experiment, to facilitate comparison across proteins. (B) Nonmonotonic-dependence of the unfolding heat capacity upon urea concentration is robust across models of the T-dependence of ΔC_p. Heat capacity vs concentration for the T-independent and linear T-dependent models of ΔC_p is plotted for RNase A in urea.

A concentration-independent heat capacity is often obtained from linear fits of the unfolding enthalpy vs melting temperature for various osmolyte concentrations. For the proteins in Figure 7 that have strong c-dependence, this would be a recipe prone to large errors.

3.2.6. The Cosolute-Dependence of the Unfolding Heat Capacity May Be Nonmonotonic. The unfolding heat capacity ΔC_{pf} may apparently even be nonmonotonic in c; the nonmonotonic behavior exhibited by RNase A in urea is well beyond what can be explained by the experimental uncertainty. Factoring in the intermodel uncertainty does not change this: panel B of Figure 7 shows the nonmonotonic behavior is present in both the T-independent and linear T-dependent models. We note as well that these measurements were made in the same lab.[46] Any systematic errors in the measurement would generally not be nonmonotonic and would be unlikely to explain this effect. A physical explanation, if the effect is indeed a real one, is an interesting topic of future work.

3.2.7. Transfer Entropy and Enthalpy at Lab Temperature. For 19 proteins and cosolutes that we had investigated, the concentration-dependence of ΔC_{pf} is known. For these proteins, we have obtained the transfer enthalpy of unfolding $\delta\Delta H(T = 25\ ^\circ\mathrm{C})$ and the transfer entropy of unfolding $\delta\Delta S(T = 25\ ^\circ\mathrm{C})$ at lab temperature; values are tabulated in Table 3. For 12 protein-cosolute systems, thermodynamic parameters at the folding temperatures corresponding to different cosolute concentrations were tabulated in the literature. For these systems, a temperature-independent ΔC_p model was invariably used to obtain the tabulated values. We thus had to also assume a temperature-independent ΔC_p model in order to extrapolate the thermodynamic values to lab temperature. We show below however that this procedure may be prone to large errors.

Seven references contained plotted data, which we had fitted to obtain thermodynamic parameters. For 5 of these protein-cosolute systems, the nonlinear temperature-dependent ΔC_p model was used to extrapolate to lab temperature. Two of these systems had baselines subtracted in the published data, so a linear temperature-dependent ΔC_p model was used to extrapolate to lab temperature. All of these 7 protein-cosolute systems show scatter due to our Monte Carlo procedure that is indicated in Figure 6 (though for α-lactalbumin in ethanol and RNase A in urea the scatter is small).

The extent of the scatter in Figure 6 makes it clear that for any of the three methods of obtaining $\delta\Delta H$ and $\delta\Delta S$ (heat capacity, fraction unfolded, or unfolding free energy, with the method for each protein indicated in Table 3), the uncertainties are highly correlated and can be quite large. There is very little scatter orthogonal to the lines of constant stability on the $\delta\Delta H$-$T\delta\Delta S$ plot. The scatter *along* the equistability line is significantly larger however: for arc repressor in particular, the scatter is large enough to render the sign of $\delta\Delta H$ and $T\delta\Delta S$ uncertain. The scatter in the data for RNase transfer to urea is quite small on the other hand, even though the scatter in the data for creatine kinase in glycerol, also from heat capacity measurements, is large. The average standard deviation along the diagonal was 21 kJ/mol, compared with an average of error of 0.32 kJ/mol perpendicular to the diagonal, corresponding to uncertainty in the change in the unfolding free energy upon transfer $\delta\Delta G$.

3.2.8. Statistical Significance Due to Fitting Uncertainty. We apply the same procedure described in Section 3.2.1 to evaluate the significance of the entropy-enthalpy compensation here. In this case there are 23 data points in Figure 6 (these proteins are also tabulated in Table 3), and the data themselves have a sample standard deviation of about $s \approx 65$ kJ/mol, whereas the mean Monte Carlo standard deviation applied to each data point is about $\sigma \approx 19$ kJ/mol. Bootstrapping with the same procedure described in section 3.2.2 rejects the hypothesis that the scatter arises from the uncertainty in fitting with a significance $p = 5 \times 10^{-5}$, see Figure 3B. The result in Figure 6 thus also illustrates entropy-enthalpy compensation rather than experimental error.

3.2.9. Statistical Significance Due to Model-to-Model Variance of $\Delta C_p(T)$. Different proteins in Figures 4, 5, and 6 have used different models for the temperature-dependence of ΔC_p, based on the availability of experimental data, and the methods used in various laboratories. For 6 proteins, however, data is available for all 3 temperature models described in Section 2.1. The model-to-model standard deviation in the rotated frame at the melting temperature for these proteins can

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

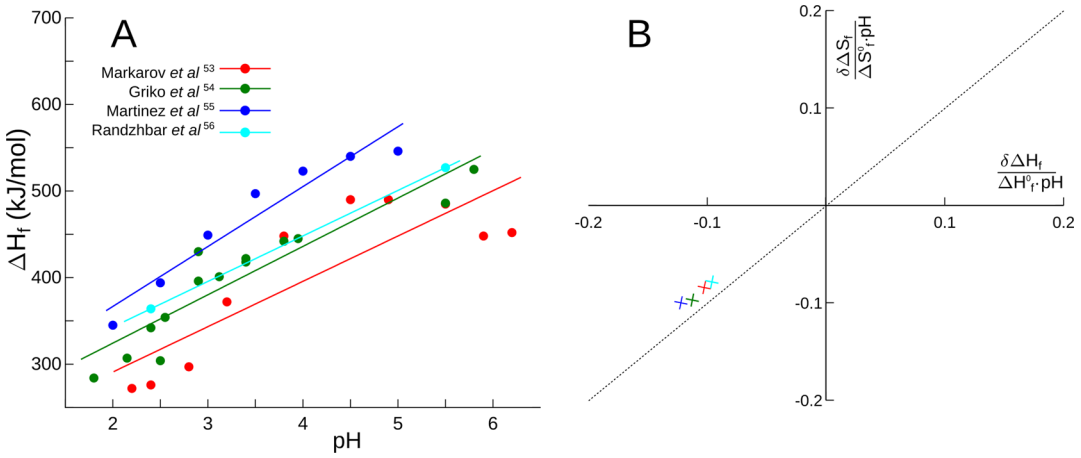


Figure 8. (A) Enthalpy of unfolding vs pH for barnase, measured for four different laboratories.[53−56] Each color corresponds to a different lab as indicated in the legend, and linear regression slopes are shown for each. (B) Change in entropy and enthalpy of unfolding, plotted as in Figures 4 and 5. The relative scatter here is about a factor of 4 larger than the Monte Carlo generated scatter in those figures; however, bootstrapping analysis (see text) shows that compensation is still significant with $p < 10^{-7}$.

be found from Table 1 and is 0.016/(100 mg/mL) on average. We thus found that entropy-enthalpy compensation is still significant at the $p < 1 \times 10^{-7}$ level. The standard deviations along the $y = x$ line at lab temperature for these proteins are as follows: arc repressor in KCl, 64.5 kJ/mol; creatine kinase in glycerol, 32.2 kJ/mol; notch ankyrin in urea, 7.45 kJ/mol; α-Lac in ethanol, 0.148 kJ/mol; hisactophilin in urea, 105.1 kJ/mol; HPr in urea, 55.3 kJ/mol. These numbers are on average about a factor of 2 larger than the fit uncertainty. We can bootstrap from these 6 values to obtain a distribution of 6^6 = 46,656 different possible values of σ, which may be used as the new null distribution. Because of the widely varying numbers above, this distribution is quite broad at lab temperature. Comparing this distribution to the one arising from bootstrapping the experimental lab temperature data points, we find that the significance drops to $p = 0.085$. Model-to-model variance is a significant enough source of error to call entropy-enthalpy into some question at lab temperature.

3.2.10. Statistical Significance Due to Lab-to-Lab Variance. Assessing lab-to-lab variance for the transfer of protein to cosolute is challenging because of a lack of comparative data. One cosolute for which abundant data is available from several laboratories for the same protein is proton concentration or pH. However, pH is exceptional in that it is an extremely effective cosolute on either a per molar or per molar mass basis, for example even a 0.1 M solution of [H+] with a pH of 1 is an extreme solution with significant effects on protein stability. Nevertheless, we can compile values from the published literature for the midpoint enthalpy of unfolding and midpoint entropy of unfolding from various laboratories at various pH values for a given protein, for which we choose barnase. At least four different laboratories have measured the enthalpy and entropy of unfolding of barnase at more than one pH.[52−56] The enthalpy of unfolding is plotted vs pH in Figure 8A. From these data sets, a transfer enthalpy and entropy per pH may be obtained from the slope of the corresponding lines, see Figure 8B. We normalize these values by the unfolding enthalpy and entropy at maximum pH, to obtain $\delta\Delta H_f/(\Delta H_f^0 \cdot \mathrm{pH})$ and $\delta\Delta S_f/(\Delta S_f^0 \cdot \mathrm{pH})$ for each lab, analogous to the abscissa and ordinate in Figures 4 and 5. This gives a standard deviation of 0.016 pH^{-1} between the above 4 laboratories. To compare directly with the analysis in Section 3.2.2, we consider the relative uncertainty; that is, the standard deviation between the four laboratories divided by the mean distance of the points from the origin along the $y = x$ line: this number is 0.107. By comparison, the relative uncertainties of the protein/cosolute systems in section 3.2.2 ranged from 0.01 to 0.04. We then ask the question: if the protein/cosolute systems used in the Monte Carlo analysis in Section 3.2.2 had the same relative uncertainty as the barnase pH lab-to-lab data has, what would the significance of entropy-enthalpy compensation be? Using the same bootstrapping procedure as in Section 3.2.9, we found no instances of the entropy-enthalpy data in various cosolutes that could have been taken from the null distribution in 10^7 trials. I.e. even though the lab-to-lab error for pH is about a factor of 4 larger than the fitting error for various cosolutes, if we take this error as representitive for the null distribution, entropy-enthalpy compensation is still significant at the $p < 10^{-7}$ level.

Finally, we note that method-to-method error within the same lab may be comparable to the error across different laboratories. Figure 5 includes data for RNase A in methylurea, ethylurea, and *N*,*N*′-dimethylurea, as measured both by calorimetry and UV spectroscopy. The mean relative error here is 0.08, as compared to about 0.1 above for the effects of pH on barnase.

4. CONCLUDING REMARKS

In this paper we have observed and analyzed significant entropy-enthalpy compensation across both diverse proteins and diverse cosolute solutions, by performing a rigorous thermodynamic analysis of calorimetric and spectroscopic data, which included bootstrapping Monte Carlo error estimates and a comparison across different models of the temperature-dependence of the unfolding heat capacity. Uncertainties in enthalpy and entropy, while much larger than the uncertainty in free energetic stability, do not rule out significant entropy-enthalpy compensation as a general phenomenon in protein transfer. The accuracy of the temperature-dependence and concentration-dependence of the unfolding heat capacity is not important near the folding transition but is important if we are interested for example in the stability at lab temperature.

Early results by Ben-Naim,[71] Grunwald,[72] Karplus,[73] and Lee[74] have analyzed the invariable entropy-enthalpy compen-

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

sation that occurs during solvent reorganization around a solute due to solvent−solvent interactions. In these theories, cavity creation results in a singular solute−solvent potential and is noncompensating, the limiting case being the free energy of inserting a noninteracting, hard-sphere solute. This issue is unlikely to be a factor in the transfer scheme wherein a solute (protein) is transferred from pure buffer to solution containing cosolute: volume is indeed lost to buffer and cosolute upon transfer at constant pressure but is also gained to the pure buffer system. A systematic analysis of entropy-enthalpy compensation in protein transfer using density functional theory to capture the effects of solvation is an interesting topic for future work.[75,76]

AUTHOR INFORMATION

Corresponding Author

*Phone: 604-822-8813. E-mail: steve@phas.ubc.ca.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank George Makhatadze and Valerij Grinberg both for sharing their data and for helpful comments. We acknowledge funding from PrioNet Canada, the Alberta Prion Research Institute, ALS Canada, and the Natural Sciences and Engineering Research Council. This research was enabled in part by support provided by WestGrid and Compute Canada Calcul Canada.

REFERENCES

(1) Claussen, W.; Polglase, M. Solubilities and Structures in Aqueous Aliphatic Hydrocarbon Solutions. *J. Am. Chem. Soc.* **1952**, *74*, 4817−4819.

(2) Aveyard, R.; Lawrence, A. S. C. Calorimetric Studies on n-aliphatic Alcohol + Water and n-aliphatic Alcohol + Water Detergent Systems. *Trans. Faraday Soc.* **1964**, *60*, 2265−2278.

(3) Arnett, E. M., McKelvey, D. R. In *Solute-Solvent Interactions*; Coetzee, J. F., Ritchie, C. D., Eds.; Dekker: New York, Chapter 6, 1969; pp 344−395.

(4) Privalov, P. L.; Gill, S. J. Stability of Protein-Structure and Hydrophobic Interaction. *Adv. Protein Chem.* **1988**, *39*, 191−234.

(5) Lumry, R.; Rajender, S. Enthalpy-Entropy Compensation Phenomena in Water Solutions of Proteins and Small Molecules: A Ubiquitous Properly of Water. *Biopolymers* **1970**, *9*, 1125−1227.

(6) Dunitz, J. D. Win Some, Lose Some: Enthalpy-Entropy Compensation in Weak Intermolecular Interactions. *Chem. Biol.* **1995**, *2*, 709−712.

(7) Qian, H.; Hopfield, J. J. Entropy-Enthalpy Compensation: Perturbation and Relaxation in Thermodynamic Systems. *J. Chem. Phys.* **1996**, *105*, 9292−9298.

(8) Sharp, K. Entropy-Enthalpy Compensation: Fact or Artifact? *Protein Sci.* **2001**, *10*, 661−667.

(9) Fenley, A. T.; Muddana, H. S.; Gilson, M. K. Entropy-Enthalpy Transduction Caused by Conformational Shifts Can Obscure the Forces Driving Protein-Ligand Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20006−20011.

(10) Chodera, J. D.; Mobley, D. L. Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. *Annu. Rev. Biophys.* **2013**, *42*, 121−142. PMID: 23654303

(11) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Entropy-Enthalpy Compensation in Solvation and Ligand Binding Revisited. *J. Am. Chem. Soc.* **1998**, *120*, 4526−4527.

(12) Ben-Naim, A.; Marcus, Y. Solvation Thermodynamics of Nonionic Solutes. *J. Chem. Phys.* **1984**, *81*, 2016−2027.

(13) Perozzo, R.; Folkers, G.; Scapozza, L. Thermodynamics of Protein-Ligand Interactions: History, Presence, and Future Aspects. *J. Recept. Signal Transduction Res.* **2004**, *24*, 1−52. PMID: 15344878

(14) Lafont, V.; Armstrong, A. A.; Ohtaka, H.; Kiso, Y.; Mario Amzel, L.; Freire, E. Compensating Enthalpic and Entropic Changes Hinder Binding Affinity Optimization. *Chem. Biol. Drug Des.* **2007**, *69*, 413−422.

(15) Krishnamurthy, V. M.; Bohall, B. R.; Semetey, V.; Whitesides, G. M. The Paradoxical Thermodynamic Basis for the Interaction of Ethylene Glycol, Glycine, and Sarcosine Chains with Bovine Carbonic Anhydrase II: An Unexpected Manifestation of Enthalpy/Entropy Compensation. *J. Am. Chem. Soc.* **2006**, *128*, 5802−5812. PMID: 16637649

(16) Dee, D. R.; Horimoto, Y.; Yada, R. Y. Conserved Prosegment Residues Stabilize a Late-Stage Folding Transition State of Pepsin Independently of Ground States. *PLoS One* **2014**, *9*, e101339.

(17) Baldwin, A. J.; Knowles, T. P. J.; Tartaglia, G. G.; Fitzpatrick, A. W.; Devlin, G. L.; Shammas, S. L.; Waudby, C. A.; Mossuto, M. F.; Meehan, S.; Gras, S. L.; et al. Metastability of Native Proteins and the Phenomenon of Amyloid Formation. *J. Am. Chem. Soc.* **2011**, *133*, 14160−14163. PMID: 21650202

(18) Plotkin, S. S.; Onuchic, J. N. Understanding Protein Folding with Energy Landscape Theory I: Basic Concepts. *Q. Rev. Biophys.* **2002**, *35*, 111−167.

(19) Krug, R. R.; Hunter, W. G.; Grieger, R. A. Statistical Interpretation of Enthalpy-Entropy Compensation. *Nature* **1976**, *261*, 566−567.

(20) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Enthalpy-Entropy and Cavity Decomposition of Alkane Hydration Free Energies: Numerical Results and Implications for Theories of Hydrophobic Solvation. *J. Phys. Chem. B* **2000**, *104*, 6271−6285.

(21) Linhananta, A.; Hadizadeh, S.; Plotkin, S. S. An Effective Solvent Theory Connecting the Underlying Mechanisms of Osmolytes and Denaturants for Protein Stability. *Biophys. J.* **2011**, *100*, 459−468.

(22) Anfinsen, C. B.; Scheraga, H. A. Experimental and Theoretical Aspects of Protein Folding. *Advances in Protein Chemistry*; New York, 1975; pp 205−301.

(23) Jackson, S. E.; Fersht, A. R. Folding of Chymotrypsin Inhibitor 2. 1. Evidence for a Two State Transition. *Biochemistry* **1991**, *30*, 10428−10435.

(24) Makhatadze, G. I.; Privalov, P. L. Energetics of Protein Structure. *Adv. Protein Chem.* **1995**, *47*, 307−425.

(25) Privalov, P.; Tiktopulo, E.; Venyaminov, S.; Griko, Y.; Makhatadze, G.; Khechinashvili, N. Heat Capacity and Conformation of Proteins in the Denatured State. *J. Mol. Biol.* **1989**, *205*, 737−750.

(26) Dill, K. A.; Bromberg, S.; Yue, K.; Fiebig, K. M.; Yee, D. P.; Thomas, P. D.; Chan, H. S. Principles of Protein Folding-A Perspective From Simple Exact Models. *Protein Sci.* **1995**, *4*, 561−602.

(27) Plotkin, S. S.; Onuchic, J. N. Understanding Protein Folding with Energy Landscape Theory II: Quantitative Aspects. *Q. Rev. Biophys.* **2002**, *35*, 205−286.

(28) Snow, C. D.; Nguyen, H.; Pande, V. S.; Gruebele, M. Absolute Comparison of Simulated and Experimental Protein-Folding Dynamics. *Nature* **2002**, *420*, 102−106.

(29) Weikl, T. R.; Palassini, M.; Dill, K. A. Cooperativity in Two-State Protein Folding Kinetics. *Protein Sci.* **2004**, *13*, 822−829.

(30) Ejtehadi, M. R.; Avall, S. P.; Plotkin, S. S. Three-body Interactions Improve the Prediction of Rate and Mechanism in Protein Folding Models. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15088−15093.

(31) Rhoades, E.; Cohen, M.; Schuler, B.; Haran, G. Two-state Folding Observed in Individual Protein Molecules. *J. Am. Chem. Soc.* **2004**, *126*, 14686−14687.

(32) Ibarra-Molero, B., Sanchez-Ruiz, J. M. In Protein Folding, Misfolding and Aggregation: Classical Themes and Novel Approaches; Muñoz, V., Ed.; *RSC Biomolecular Sciences*; Royal Society of Chemistry: Cambridge, UK, 2008; pp 85−103.

(33) Poland, D. Empirical Protein Partition Functions. *J. Phys. Chem. B* **2012**, *116*, 6683−6693.

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

(34) Privalov, P. L. *Microcalorimetry of Macromolecules: The Physical Basis of Biological Structures*; John Wiley & Sons, Inc.: New York, 2012.
(35) Wafer, L. N. R.; Streicher, W. W.; Makhatadze, G. I. Thermodynamics of the Trp-cage Miniprotein Unfolding in Urea. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1376−1381.
(36) Fenley, A. T.; Muddana, H. S.; Gilson, M. K. Entropy-Enthalpy Transduction Caused by Conformational Shifts can Obscure the Forces Driving Protein-Ligand Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20006−20011.
(37) Becktel, W. J.; Schellman, J. A. Protein Stability Curves. *Biopolymers* **1987**, *26*, 1859−1877.
(38) Grinberg, V. Y.; Grinberg, N. V.; Burova, T. V.; Dalgalarrondo, M.; Haertlé, T. Ethanol-Induced Conformational Transitions in Holo-α-lactalbumin: Spectral and Calorimetric Studies. *Biopolymers* **1998**, *46*, 253−265.
(39) Wintrode, P. L.; Makhatadze, G. I.; Privalov, P. L. Thermodynamics of Ubiquitin Unfolding. *Proteins: Struct., Funct., Genet.* **1994**, *18*, 246−253.
(40) Makhatadze, G.; Privalov, P. Heat Capacity of Proteins: I. Partial Molar Heat Capacity of Individual Amino Acid Residues in Aqueous Solution: Hydration Effect. *J. Mol. Biol.* **1990**, *213*, 375−384.
(41) Zweifel, M. E.; Barrick, D. Relationships Between the Temperature Dependence of Solvent Denaturation and the Denaturant Dependence of Protein Stability Curves. *Biophys. Chem.* **2002**, *101*−*102*, 221−237.
(42) Liu, C.; Chu, D.; Wideman, R. D.; Houliston, R. S.; Wong, H. J.; Meiering, E. M. Thermodynamics of Denaturation of Hisactophilin, a β-Trefoil Protein. *Biochemistry* **2001**, *40*, 3817−3827. PMID: 11300762
(43) Nicholson, E. M.; Scholtz, J. M. Conformational Stability of the Escherichia coli HPr Protein:? Test of the Linear Extrapolation Method and a Thermodynamic Characterization of Cold Denaturation. *Biochemistry* **1996**, *35*, 11369−11378.
(44) Bowie, J. U.; Sauer, R. T. Equilibrium Dissociation and Unfolding of the Arc Repressor Dimer. *Biochemistry* **1989**, *28*, 7139−7143.
(45) Meng, F.-G.; Hong, Y.-K.; He, H.-W.; Lyubarev, A. E.; Kurganov, B. I.; Yan, Y.-B.; Zhou, H.-M. Osmophobic Effect of Glycerol on Irreversible Thermal Denaturation of Rabbit Creatine Kinase. *Biophys. J.* **2004**, *87*, 2247−2254.
(46) Poklar, N.; Petrovčič, N.; Oblak, M.; Vesnaver, G. Thermodynamic Stability of Ribonuclease A in Alkylurea Solutions and Preferential Solvation Changes Accompanying its Thermal Denaturation: A Calorimetric and Spectroscopic Study. *Protein Sci.* **1999**, *8*, 832−840.
(47) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, 3rd ed.; W. H. Freeman & Co.: 1998.
(48) Wang, Y.; Sarkar, M.; Smith, A. E.; Krois, A. S.; Pielak, G. J. Macromolecular Crowding and Protein Stability. *J. Am. Chem. Soc.* **2012**, *134*, 16614−16618.
(49) Senske, M.; Törk, L.; Born, B.; Havenith, M.; Herrmann, C.; Ebbinghaus, S. Protein Stabilization by Macromolecular Crowding through Enthalpy Rather Than Entropy. *J. Am. Chem. Soc.* **2014**, *136*, 9036−9041.
(50) O'Brien, E. P.; Ziv, G.; Haran, G.; Brooks, B. R.; Thirumalai, D. Effects of Denaturants and Osmolytes on Proteins are Accurately Predicted by the Molecular Transfer Model. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 13403−13408.
(51) Auton, M.; Bolen, D. W. Predicting the Energetics of Osmolyte-Induced Protein Folding/Unfolding. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15065−15068.
(52) Pfeil, W. *Protein Stability and Folding: A Collection of Thermodynamic Data*; Springer-Verlag: 1998.
(53) Makarov, A. A.; et al. Comparative Study of Thermostability and Structure of Close Homologues-Barnase and Binase. *J. Biomol. Struct. Dyn.* **1993**, *10*, 1047−1065.
(54) Griko, Y. V.; Makhatadze, G. I.; Privalov, P. L.; Hartley, R. W. *Protein Sci.* **1994**, *3*, 669−676.
(55) Martínez, J. C.; Harrous, M. E.; Filimonov, V. V.; Mateo, P. L.; Fersht, A. R. A Calorimetric Study of the Thermal Stability of Barnase and Its Interaction with 3′GMP. *Biochemistry* **1994**, *33*, 3919−3926.
(56) Randzhbar, B.; Protasevich, I. I.; Shul'ga, A. A.; Kurbanov, V. M.; Lobachev, V. M.; Kirpichnikov, M. P.; Makarov, A. A. Barnase, Binase, and a Hybrid of Them: Differences in Conformation and Heat Denaturation Parameters. *Molekuliarniaia Biologiia* **1997**, *31*, 492−499.
(57) Kaushik, J. K.; Bhat, R. Why Is Trehalose an Exceptional Protein Stabilizer?: An Analysis of the Thermal Stability of Proteins in the Presence of the Compatible Osmolyte Trehalose. *J. Biol. Chem.* **2003**, *278*, 26458−26465.
(58) Singh, R.; Haque, I.; Ahmad, F. Counteracting Osmolyte Trimethylamine N-Oxide Destabilizes Proteins at pH Below Its pKa: Measurements of Thermodynamic Parameters of Proteins in the Presence and Absence of Trimethylamine N-Oxide. *J. Biol. Chem.* **2005**, *280*, 11035−11042.
(59) Chiti, F.; van Nuland, N. A. J.; Taddei, N.; Magherini, F.; Stefani, M.; Ramponi, G.; Dobson, C. M. Conformational Stability of Muscle Acylphosphatase:? The Role of Temperature, Denaturant Concentration, and pH. *Biochemistry* **1998**, *37*, 1447−1455.
(60) Agashe, V. R.; Udgaonkar, J. B. Thermodynamics of Denaturation of Barstar: Evidence for Cold Denaturation and Evaluation of the Interaction with Guanidine Hydrochloride. *Biochemistry* **1995**, *34*, 3286−3299.
(61) Padmanabhan, S.; Laurents, D. V.; Fernández, A. M.; Elias-Arnanz, M.; Ruiz-Sanz, J.; Mateo, P. L.; Rico, M.; Filimonov, V. V. Thermodynamic Analysis of the Structural Stability of Phage 434 Cro Protein. *Biochemistry* **1999**, *38*, 15536−15547.
(62) Kamiyama, T.; Sadahide, Y.; Nogusa, Y.; Gekko, K. Polyol-induced molten globule of cytochrome c: an evidence for stabilization by hydrophobic interaction. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1999**, *1434*, 44−57.
(63) Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; Degrado, W. F. From Coiled Coils to Small Globular Proteins: Design of a Native-like Three-Helix Bundle. *Protein Sci.* **1998**, *7*, 1404−1414.
(64) Catanzano, F.; Gambuti, A.; Graziano, G.; Barone, G. Interaction with D-Glucose and Thermal Denaturation of Yeast Hexokinase B: A DSC Study. *J. Biochem.* **1997**, *121*, 568−577.
(65) Ahmad, N.; Srinivas, V. R.; Reddy, G. B.; Surolia, A. Thermodynamic Characterization of the Conformational Stability of the Homodimeric Protein, Pea Lectin. *Biochemistry* **1998**, *37*, 16765−16772.
(66) Kovrigin, E. L.; Potekhin, S. A. Microcalorimetric study of the effect of dimethylsulfoxide on the heat denaturation of lysozyme. *Biofizika* **1996**, *41*, 1201−1206.
(67) Knapp, S.; Ladenstein, R.; Galinski, E. A. Extrinsic Protein Stabilization by the Naturally Occurring Osmolytes β-hydroxyectoine and Betaine. *Extremophiles* **1999**, *3*, 191−198.
(68) Santoro, M. M.; Liu, Y.; Khan, S. M. A.; Hou, L. X.; Bolen, D. W. Increased thermal stability of proteins in the presence of naturally occurring osmolytes. *Biochemistry* **1992**, *31*, 5278−5283.
(69) Xu, S.; Qin, S.; Pan, X.-M. Thermal and Conformational Stability of Ssh10b Protein from Archaeon Sulfolobus Shibattae. *Biochem. J.* **2004**, *382* (2), 433−440.
(70) Makhatadze, G. I.; Lopez, M. M.; Richardson, J. M., III; Thomas, S. T. Anion Binding to the Ubiquitin Molecule. *Protein Sci.* **1998**, *7*, 689−697.
(71) Ben-Naim, A. Hydrophobic Interaction and Structural Changes in the Solvent. *Biopolymers* **1975**, *14*, 1337−1355.
(72) Grunwald, E. Thermodynamic Properties, Propensity Laws, and Solvent Models in Solutions in Self-associating Solvents. Application to Aqueous Alcohol Solutions. *J. Am. Chem. Soc.* **1984**, *106*, 5414−5420.
(73) Yu, H.; Karplus, M. A Thermodynamic Analysis of Solvation. *J. Chem. Phys.* **1988**, *89*, 2366−2379.
(74) Lee, B. Enthalpy-Entropy Compensation in the Thermodynamics of Hydrophobicity. *Biophys. Chem.* **1994**, *51*, 271−278.
(75) Mills, E. A.; Plotkin, S. S. Density Functional Theory for Protein Transfer Free Energy. *J. Phys. Chem. B* **2013**, *117*, 13278−13290.

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144

(76) Borgis, D.; Gendre, L.; Ramirez, R. Molecular Density Functional Theory: Application to Solvation and Electron-Transfer Thermodynamics in Polar Solvents. *J. Phys. Chem. B* **2012**, *116*, 2504−2512.

DOI: 10.1021/acs.jpcb.5b09219
J. Phys. Chem. B 2015, 119, 14130−14144