Additive Transfer Free Energies of the Peptide Backbone Unit That Are Independent of the Model Compound and the Choice of Concentration Scale†

Matthew Auton‡ and D. Wayne Bolen*

University of Texas Medical Branch, 5.154 Medical Research Building, 301 University Boulevard, Galveston, Texas 77555

Received October 27, 2003

ABSTRACT: With knowledge of individual transfer free energies of chemical groups that become newly exposed on protein denaturation and assuming the group transfer free energy contributions are additive, it should be possible to predict the stability of a protein in the presence of denaturant. Unfortunately, several unresolved issues have seriously hampered quantitative development of this transfer model for protein folding/unfolding. These issues include the lack of adequate demonstration that group transfer free energies (ΔG_{tr}) are additive and independent of the choice of model compound, the problem arising from dependence of ΔG_{tr} on concentration scales, the lack of knowledge of activity coefficients, and the validity of the mathematical constructs used in obtaining ΔG_{tr} values. Regarding transfer from water to 1 M concentrations of the naturally occurring osmolytes, trimethylamine-*N*-oxide (TMAO), sarcosine, betaine, proline, glycerol, sorbitol, sucrose, trehalose, and urea, using cyclic glycylglycine, zwitterionic glycine peptides, and *N*-acetylglycine amide peptides as models for the peptide backbone of proteins, we set out to address these issues and obtain ΔG_{tr} values for the peptide backbone unit. We demonstrate experimental approaches that obviate the choice of concentration scale and demonstrate additivity in ΔG_{tr} of the peptide backbone unit for all solvent systems studied. Evidence is presented to show that the ΔG_{tr} values are independent of the chemical model studied, and experimental conditions are given to illustrate when the mathematical constructs are valid and when activity coefficients can be ignored. Resolution of the long-standing issues that have stymied development of the transfer model now make it possible to design transfer experiments that yield reliable and quantitative values for the interactions between osmolyte-containing solvents and native and unfolded protein.

Although Hofmeister's classic paper regarding the effects of salt on protein properties was published more than a century ago, interest in how solute molecules and ions interact with protein groups remains strong even today (*1*). A key development in the early part of the last century made it possible to measure propensities for certain types of weak interactions involving proteins and solvent components. Following the work of Brønsted (*2*) and Noyes (*3, 4*) on solubility and on the basis of the solubility limits of solute in two solvents, Lewis and Randall (*5*) noted the principle that at the solubility limit the chemical potentials of a solute in the crystal and in the saturated solution are equal. From this came the ability to evaluate the transfer free energy of a solute from one solvent to another. The principle was later applied to solute partitioning between immiscible solvents, as well as solubility-based transfer free energy measurements to identify and understand fundamental interactions responsible for the structure and function of proteins (*6*−*14*). Among other benefits, this body of work helped to establish a basis for developing and understanding concepts concerning the globular nature of soluble proteins, the hydrophobic effect, heat capacity changes on denaturation, the structure and properties of lipid bilayers, and protein denaturation by urea and guanidinium chloride (GdnHCl) (*5*−*10*, *12*−*16*). More recently, transfer free energy measurements were used to understand how naturally occurring osmolytes stabilize proteins in vivo (*17*−*20*). From these measurements, the solvophobic interaction of osmolyte with peptide backbone (the osmophobic effect) was identified as the driving force favoring protein folding in protecting osmolyte solutions (*21*). In terms of contributions, the wealth of conceptual ideas derived from the transfer free energy considerations form a substantial part of the foundation of our present day understanding of biological structure and function.

While transfer free energy considerations have been instrumental to the development of several fundamental concepts and ideas in molecular biophysics, its success as a tool in quantitatively describing such important processes as protein denaturation has been lacking. Some 40 years ago, Tanford and Nozaki began development of a quantitative model for explaining the basis of urea- or GdnHCl-induced denaturation of proteins, a model referred to as the transfer model of protein denaturation (*9*, *10*, *22*, *23*). Basically, the transfer model relates the free energy change for protein denaturation at a given concentration of denaturant (ΔG_D) to the free energy change accompanying denaturant interaction with peptide backbone and side chain groups newly exposed on denaturation ($\sum_i n_i \alpha_i \delta g_{tr,i}$) (*24*). An additional free energy term, $\Delta G_D^{H_2O}$, representing the denaturation free

† Support provided by NIH Grant GM49760 and Welch Foundation Grant H-1444.

* Reprint requests should be sent to D. Wayne Bolen. Telephone: (409) 772-0754. Fax: (409) 772-1790. Web address: http://www.hbcg.utmb.edu/faculty/bolen. E-mail address: wbolen@hbcg.utmb.edu.

‡ E-mail address: mtauton@utmb.edu.

10.1021/bi035908r CCC: $27.50
Published on Web 01/13/2004

energy change that occurs in the absence of denaturant is also an inherent part of the model, eq 1,

$$\Delta G_D = \Delta G_D^{H_2O} + \sum_i n_i \alpha_i \delta g_{tr,i} \quad (1)$$

where α_i is the average fractional change in degree of exposure to solvent of groups of type *i*, $\delta g_{tr,i}$ is the free energy of transfer of a group of type *i* from water to the specified concentration of denaturant, and n_i is the total number of groups of type *i* present in the protein (*24*). An inherent assumption implied by the summation term in eq 1 is that the free energy contributions of all parts of a protein are additive.

The linear extrapolation model (LEM), the denaturant binding model (DB), and the transfer model for protein denaturation are the principal models applied to denaturant-induced protein unfolding (*25*). In their present usage, the LEM and DB models provide means to evaluate $\Delta G_D^{H_2O}$, but they have little or no predictive value and provide little or no molecular-level detail or insight into denaturant-induced unfolding. By contrast, if the transfer model could be fully developed, it would have the potential for providing rich detail on the interaction of denaturants or protecting osmolytes with side chain and backbone units of the native and denatured states, as well as providing predictive capabilities regarding protein stabilities and relative solubilities of proteins in the presence of denaturants and protecting osmolytes.

The reason that little progress has been made in developing the transfer model over the past 40 years is that there are serious unsolved and unresolved issues dealing with practical matters of evaluating dependable transfer free energies of the various parts of proteins and in demonstrating that the individual transfer free energies are additive, as required of the summation shown in eq 1. For the model to be advanced, it is clear that these unsolved and unresolved issues must be fully addressed and evaluated. The issues can best be illustrated in the context of the original work of Nozaki and Tanford, who chose glycine, diglycine, and triglycine as models in evaluating the transfer free energy of a peptide backbone unit, the most numerous group in a protein (*10*, *23*). Solubilities of the glycine model compounds were determined in water and in urea solutions of known concentration, transfer free energies were calculated, and evidence of additivity was sought (*9*). Additivity in the transfer free energy of a peptide backbone unit in such a series of peptides requires the transfer free energy to change by the same $\delta g_{tr,i}$ increment with each successive glycine peptide unit in the chain. To obtain evidence of additivity for peptide backbone transfer free energy, Tanford and Nozaki used mathematical constructs devised originally by Cohn and Edsall (*6*). These constructs involve subtraction of transfer free energies of peptides differing in chain length and constructs that combined subtraction and division. The peptide backbone transfer free energies resulting from their studies and others provide no convincing evidence that the peptide backbone transfer free energies are additive (*7*, *9*, *10*, *18*, *22*, *23*, *26*−*29*). Much later, using chemically blocked N and C termini, Nandi and Robinson obtained evidence of additivity in peptide backbone transfer free energies from water to urea, but their numbers were significantly different from that of Nozaki and Tanford (*30*).

Three unresolved issues arise immediately from these studies: (1) Are peptide backbone transfer free energies additive as required of the transfer model? (2) Are the constructs of subtraction and subtraction/division combinations valid ways of evaluating transfer free energies? (3) The results of Nozaki and Tanford and Nandi and Robinson seem to show that different chemical models give different peptide backbone transfer free energies. Can the transfer free energy of the peptide backbone unit be shown to be independent of the chemical model used?

Adding to the list above, a fourth issue (4) arises from the fact that the absolute value of the transfer free energy is dependent on whether one chooses to use the molarity, the molality, or the mole fraction concentration scale to describe the solubility (*31*, *32*). The differences in magnitudes of the transfer free energy are particularly acute when the transfer is from one solvent whose density differs significantly from that of the other solvent and/or the solute is highly soluble. The issue then is to learn how to deal with the variability in transfer free energies brought on by the choice of the concentration scale. The final problem arises from the historical practice of ignoring activity coefficients in transfer free energy measurements. Thus, the fifth (5) and final issue is to know when it is and when it is not necessary to take activity coefficients into account.

In initiating the studies to be described, our goal was to take a comprehensive look at all of these issues with the objective of resolving them or reconciling why they cannot be resolved. Using cyclic glycylglyine and *N*-acetylated glycinamide peptides of different chain length, we show that if model studies are properly designed and carried out, a transfer free energy of the peptide backbone can be determined that exhibits additivity, is independent of the chemical model used, and is independent of the choice of concentration scale. In addition, we show conditions under which the subtraction, division, and subtraction/division combination constructs give valid results and when they do not. Overall, this study demonstrates that with proper experimental design issues encountered by previous workers can be overcome, thus paving the way for quantification of the transfer model for protein unfolding.

MATERIALS

Osmolytes. The osmolytes glycine, betaine, L-proline, glycerol, D-sorbitol, D-(+)-sucrose, and D-(+)-trehalose with purity above 99% were purchased from Sigma Chemical Co. Sarcosine was purchased from Fluka, and ultrapure urea was purchased from ICN. All purchased osmolytes were used without further purification. Trimethylamine-*N*-oxide (TMAO) was synthesized using a modification of the procedure of Hickinbottom (*33*) as described by Russo et al. (*34*).

Peptide Backbone Model Compounds. Glycine (Gly) and acetamide were purchased from Acros Organics. Diglycine (Gly_2), triglycine (Gly_3), tetraglycine (Gly_4), and cyclic glycylglycine (CGG, also known as glycine anhydride or diketopiperazine) were purchased from Sigma Chemical Co. All purchased model compounds were used without further purification. *N*-acetylglycinamide (NAGA), *N*-acetyldiglycinamide (NAG_2A), *N*-acetyltriglycinamide (NAG_3A), and *N*-acetyltetraglycinamide (NAG_4A) were synthesized as described below.

Organic Synthesis Materials. Glycinamide HCl, 98% purity, was purchased from Acros Organics, while diglycinamide HCl, triglycinamide HCl, and *N*-acetylglycine nitrophenylester were purchased from Bachem. ACS grade sodium hydroxide was purchased from Fisher Scientific, and Aldrich Chemical Co. was the supplier for acetic anhydride, 99+% purity, and *N*,*N*-dimethylformamide. Mixed bed ion-exchange resin AG 501-X8 was purchased from BioRad. All synthesis reagents were used without further purification.

METHODS

Organic Synthesis

Acetylation. The hydrochloride salt of glycinamide (20 g, 181.4 mmol, 1 equiv), diglycinamide (10 g, 59.7 mmol, 1 equiv), or triglycinamide (1 g, 4.45 mmol, 1 equiv) was dissolved in water at 0 °C near the solubility limit in an ice bath while stirring. To generate the free reactive amine, the initial pH (~3.0) of the mixture was titrated to 10.2 using a cold (0 °C) saturated solution of sodium hydroxide in water. The temperature of the mixture increased during the titration and was adjusted to 0 °C by adding ice to the reaction mixture. Acetic anhydride stored in the freezer at −20 °C (25 mL, 226 mmol, 1.25 equiv for reaction with glycinamide; 7 mL, 74.5 mmol, 1.25 equiv for reaction with diglycinamide; or 1.6 mL, 16.7 mmol, 1.25 equiv for reaction with triglycinamide) was added to the stirring reaction mixture in one portion to facilitate the acetylation, which resulted in a significant increase in solution temperature upon reaction. The reaction mixture was allowed to stir in the ice bath until the temperature reached 0 °C.

Amino Acid Coupling. Two solutions were prepared. For solution 1, dissolution of the triglycinamide HCl (1 g, 4.45 mmol, 1 equiv) into 200 mL of *N*,*N*-dimethylformamide was facilitated by heating the solution to 80 °C. Dropwise titration of triethylamine into the mixture generated the free reactive amine of triglycinamide and the biproduct, triethylamine hydrochoride. For solution 2, *N*-acetylglycine nitrophenylester (1.06 g, 4.45 mmol, 1 equiv) was dissolved in approximately 10 mL of dimethylformamide, and the coupling of *N*-acetylglycine to triglycinamide was achieved by titrating solution 2 dropwise into solution 1 while stirring. During the titration, the mixture became bright yellow in color due to the generation of the *para*-nitrophenol biproduct. Also the solution slowly became cloudy indicating the precipitation of the product *N*-acetyltetraglycinamide. The mixture was allowed to stir and cool on ice to facilitate the complete precipitation of NAG_4A.

NAGA. To remove the salts, mixed bed ion-exchange resin AG 501-X8 was added to the reaction mixture until the solution conductivity was minimal. The mixture was filtered from the ion-exchange resin and evaporated to dryness under reduced pressure at 60 °C using a Büchi rotoevaporator system. The remaining solid product was rinsed with acetone and recrystallized from 100% ethanol to a constant melting point, 139−141 °C. Yield ≈ 70%. Elemental analysis from Galbraith Laboratories Inc.: C, 41.40%, H, 7.08%, O, 27.03%, N, 24.04%.

NAG_2A. The reaction mixture was sonicated in an ice bath at 0 °C, and crystals of the product began forming within 10 min. The crystals were filtered from the solvent and rinsed with cold acetone. The filtrate and the acetone washings were combined, and the remaining product in these washings was allowed to crystallize overnight at 4 °C to form a second crop. The crystals were combined, rinsed with acetone, and recrystallized in ethanol/ether to a constant melting point, 205−207 °C. Yield ≈ 74%. Elemental analysis: C, 41.58%, H, 6.56%, O, 26.14%, N, 24.18%.

NAG_3A. Upon reaction, a precipitate of the product formed and was filtered from the solvent and rinsed with cold (−20 °C) methanol. The filtrate and the methanol washings were combined, and the remaining product in these washings was allowed to crystallize overnight at 4 °C to form a second crop. The precipitate and the second crop crystals were rinsed with acetone, combined, and recrystallized from water/methanol to a constant melting point, 251−253 °C. Yield ≈ 74%. Elemental analysis: C, 41.29%, H, 6.27%, O, 26.29%, N, 23.98%.

NAG_4A. The precipitate of the product was filtered and rinsed with cold (4 °C) chloroform to remove excess *para*-nitrophenol and triethylamine hydrochloride. The product was recrystallized in water/methanol to a constant melting point, 273−274 °C. Yield ≈ 83%. Elemental analysis: C, 41.67%, H, 6.03%, O, 27.85%, N, 24.06%.

^{1}H NMR spectra were obtained for each of the peptides in d_6-dimethyl sulfoxide, and the proton chemical shifts and integrals were found to be consistent with the structure. The elemental analyses were consistent with the calculated atomic compositions of the peptides.

Solubility Measurements

Solubilities were determined as described by Liu and Bolen (*18*). To a minimum of eight glass vials, each containing a fixed weight of solvent, was added weighed amounts of a peptide backbone unit model compound to provide a series of mixtures of increasing composition of the peptide unit model compound. Each vial was capped with a Teflon-coated screw cap and the caps were wrapped with Parafilm to produce a watertight seal. Based on experience, the range of compositions, in units of (grams of peptide/100 g of solvent) calculated from eq 2, was selected so that approximately half of the solutions made were unsaturated and the remaining solutions saturated.

$$\text{Composition } (g_{\text{peptide}}/100\ g_{\text{solvent}}) = \left(\frac{\text{weight}_{\text{peptide}}\,(g)}{\text{weight}_{\text{solvent}}\,(g)}\right) \times 100 \quad (2)$$

The vials were immersed in a Tronac large capacity water bath equipped with a submerged reciprocating (30 cycles/min) shaker platform maintained at 25 ± 0.001 °C with a Tronac PTC-41 precision temperature controller. The solutions were allowed to equilibrate at 25 °C for 48 h before the solubility was determined. After equilibration, the supernatant of each solution was removed via syringe and filtered through glass fiber, and the density of the solution was determined using an Anton Paar DMA 602 density measuring cell and DMA 60 density meter. The temperature of the density measuring cell was maintained at 25 ± 0.001 °C with a Sodev Peltier programmable circulating water thermostat equipped with a Sodev PC-A small volume high flow rate positive displacement pump. The temperature of

the cell was monitored using an Anton Paar CKT100 precision thermometer. The solubilities of the peptide backbone unit model compounds, CGG, Gly, Gly_2, Gly_3, Gly_4, acetamide, NAGA, NAG_2A, NAG_3A, and NAG_4A were determined from plots of density vs composition in water and in 1 M concentrations of each of the nine osmolytes, trimethylamine-*N*-oxide (TMAO), sarcosine, betaine, proline, glycerol, sorbitol, sucrose, trehalose, and urea. The solubility limits expressed as (grams of peptide/100 g of solvent) are converted to molarity (moles of peptide/liter of solution), molality (moles of peptide/kg of solvent) and mole fraction (moles of peptide/total moles), and the transfer free energies were calculated as described below.

Theory and Design

Conversion of Composition to Molarity, Molality, and Mole Fraction. According to eq 2, the composition at the solubility limit gives the number of grams of peptide soluble in 100 g of solvent. This quantity is converted to the more appropriate units of molarity, molality, and mole fraction that are commonly used in determining transfer free energies. Equations 3, 4, and 5 convert the composition to the molarity scale, molality scale, and mole fraction scale, respectively, for a solution containing 100 g of solvent.

$$\text{molarity} = \frac{W_p \rho_{sn}}{M_p(W_p + W_{sv})} \tag{3}$$

$$\text{molality} = \frac{W_p}{M_p}\frac{1000}{W_{sv}} \tag{4}$$

$$\text{mole fraction} = \frac{W_p}{M_p}\Bigg/\left(\frac{W_p}{M_p} + \frac{W_{sv} - M_{os}\dfrac{c_{sv}W_{sv}}{\rho_{sv}}}{M_w} + \frac{c_{sv}W_{sv}}{\rho_{sv}}\right) \tag{5}$$

where W_p is the model compound weight in grams, W_{sv} = 100 g of solvent, ρ_{sn} is the density (g/L) of the saturated solution at the solubility limit determined experimentally, and M_p is the model compound molecular weight. In using the equations, we set the osmolyte molarity, c_{sv}, equal to zero when the solubility of the peptide in pure water is desired and equal to 1 when considering the solubility of the peptide in a 1 M osmolyte solution. ρ_{sv} represents the density (g/L) of the 1 M osmolyte solvent, M_w is the molecular weight of water, and M_{os} is the molecular weight of the osmolyte.

Equations 3, 4, and 5 can also be used to determine the molar, molal, and mole fraction of osmolyte in the presence of saturating amounts of the model peptide by simply substituting the number of moles of osmolyte, $c_{sv}W_{sv}/\rho_{sv}$, for the number of moles of model compound, W_p/M_p, in eqs 3 and 4 and in the numerator of eq 5. Because the solutions are made by weight, this substitution emphasizes that the osmolyte concentration is constant on the molality scale as a function of the composition of the peptide model compound. While it is conceivable to design the experiment such that the osmolyte concentration is constant on the molarity or the mole fraction scales, this is inherently more difficult because it requires prior knowledge of the density and composition of the saturated solution at the solubility limit of the peptide. For model peptides with high solubilities, the additional volume contributed by the soluble peptide will cause the concentration of the osmolyte on the molar and mole fraction scales to be significantly diminished. However, for model peptides with low solubilities, the molarity and mole fraction of the osmolyte will not be significantly altered.

Transfer Free Energy. The transfer free energy of a solute from a pure aqueous system to a mixed aqueous osmolyte system is based upon the solubility of the solute in each of the systems and gives the preference of the solute for one system over the other. Here, the solute refers to the chemical compound that is used to model the peptide backbone unit, and the aqueous osmolyte system is prepared to be unit molarity when the solute solubility is vanishingly small. At and above the solubility limit, the solute in the crystalline phase is in equilibrium with the solution phase, so the chemical potential of the solute in solution is equal to the chemical potential of the solute in the crystalline phase (*5*). Since the chemical potential of the solute in the crystal is constant and independent of the solvent with which it is in contact, the chemical potential of the solute at saturation in a pure aqueous system must be equal to the chemical potential of the solute at saturation in a mixed aqueous osmolyte system (*5*), eq 6.

$$\mu^\circ_{i,w} + RT\ln s_{i,w} + RT\ln \gamma_{i,w} = \mu^\circ_{i,os} + RT\ln s_{i,os} + RT\ln \gamma_{i,os} \tag{6}$$

where $\mu^\circ_{i,w}$ and $\mu^\circ_{i,os}$ are the standard chemical potentials assigned to the solute in water and in 1 M osmolyte, $s_{i,w}$ and $s_{i,os}$ are the solute solubilities in water and in 1 M osmolyte and are expressed either in molar, molal, or mole fraction units as calculated from eqs 3, 4, and 5, and $\gamma_{i,w}$ and $\gamma_{i,os}$ are the activity coefficients of the solute expressed in terms of molar, molal, or mole fraction in water and in 1 M osmolyte, respectively, at the solubility limit. The transfer free energy of the solute from water to 1 M osmolyte is equal to the difference between the standard state chemical potentials of the solute in each solution, eq 7.

$$\Delta\mu^\circ = \mu^\circ_{i,os} - \mu^\circ_{i,w} = RT\ln(s_{i,w}/s_{i,os}) + RT\ln(\gamma_{i,w}/\gamma_{i,os}) \tag{7}$$

Traditionally, the term containing the activity coefficients has been ignored because of the difficulty in obtaining these values for three component systems. When the activity coefficients are disregarded, part of the information describing the free energies of transfer is lost, and the total interaction free energy between the solute and the solvents is not known. What follows is an empirical approach that compares free energies of transfer calculated on the molar, molal, and mole fraction concentration scales to help identify when activity coefficients might be neglected.

When activity coefficients are neglected, $\Delta\mu^\circ$ becomes an apparent transfer free energy, $\Delta\bar{G}^\circ_{app}$, that depends on the concentration scale used to express the solubilities of the solute in each solvent system. The apparent transfer free energy expressed in units of molarity, molality, and mole fraction is shown in eqs 8, 9, and 10, respectively.

$$\Delta\bar{G}^\circ_{app,c} = RT\ln(n_{i,w}/n_{i,os}) + RT\ln(V_{S,os}/V_{S,w}) \tag{8}$$

$$\Delta\bar{G}^{\circ}_{app,m} = RT\ln(n_{i,w}/n_{i,os}) + RT\ln(wt_{os}/wt_{w}) \quad (9)$$

$$\Delta\bar{G}^{\circ}_{app,x} = RT\ln(n_{i,w}/n_{i,os}) + RT\ln((n_{w,os} + n_{os} + n_{i,os})/(n_{w} + n_{i,w})) \quad (10)$$

In our experimental approach, $RT\ln(wt_{os}/wt_{w}) = 0$ in eq 9 because we have chosen to express the solubility on the basis of 100 g of solvent, regardless of whether the solvent is water or osmolyte solution. Here, $n_{i,w}$ and $n_{i,os}$ represent the moles of solute soluble in 100 g water and in 100 g of 1 M osmolyte; $V_{S,w}$ and $V_{S,os}$ are the total volumes of the aqueous solution and the osmolyte solution containing the saturating solute on the molarity scale; wt_{w} and wt_{os} are, respectively, the total mass of water and of 1 M osmolyte on the molality scale; n_{w} and $n_{w,os}$ are, respectively, the moles of water in the aqueous solution and the 1 M osmolyte solution, each containing the saturating solute, and for the mole fraction case, n_{os} is the actual moles of osmolyte in the 1 M osmolyte solution containing the saturating solute.

The first term on the right of eqs 8, 9, and 10 provides free energy of transfer information on the number of moles of solute soluble in 100 g of water and 100 g of 1 M osmolyte, while the second term contains all of the information that defines the particular concentration scale that could be used, that is, the total volumes, masses, and numbers of moles of the solutions. Taking the limit of $\Delta\bar{G}^{\circ}_{app}$ as the solubility of the solute becomes very low renders the second terms on the right of eqs 8, 9, and 10 as constants of known value, and one can factor out the concentration scale dependence of the apparent transfer free energy. To summarize, the limit of the first term corresponds to the intrinsic free energy of transfer of the solute, $\Delta\bar{G}^{\circ}_{int}$, from water to 1 M osmolyte (this limit is indeed reached with model compounds that have very low solubilities). As $n_i \rightarrow 0$, the second term approaches a constant that is defined by the properties of pure water and the 1 M osmolyte solution and is simply a concentration scale correction factor in eqs 8, 9, and 10 that can be numerically evaluated, as shown in eqs 11, 12, and 13, respectively.

$$\lim_{n_i \rightarrow 0}\Delta\bar{G}^{\circ}_{app,c} = \Delta\bar{G}^{\circ}_{int} + RT\ln(d_{w}/d_{os}) \quad (11)$$

$$\lim_{n_i \rightarrow 0}\Delta\bar{G}^{\circ}_{app,m} = \Delta\bar{G}^{\circ}_{int} \quad (12)$$

$$\lim_{n_i \rightarrow 0}\Delta\bar{G}^{\circ}_{app,x} = \Delta\bar{G}^{\circ}_{int} + RT\ln((n_{w,os} + n_{os})/n_{w}) \quad (13)$$

where d_{w} and d_{os} are the densities of water and 1 M osmolyte; n_{w} is 5.55 mol/100 g of pure water; $n_{w,os}$ and n_{os} are the number of moles of water and the number of moles of osmolyte in 100 g of a 1 M osmolyte solution. All this can be represented in a single equation for model compounds that have low solubilities in water and osmolyte solution, see eq 14.

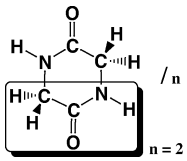
$$\Delta\bar{G}^{\circ}_{int} = \Delta\bar{G}^{\circ}_{app} - \Delta\bar{G}^{\circ}_{SC,\infty} = \Delta\mu^{\circ} - RT\ln(\gamma_{i,w}/\gamma_{i,os}) \quad (14)$$

where $\Delta\bar{G}^{\circ}_{SC,\infty}$ is the concentration scale correction factor defined above. Table 1 provides the numerical free energies of $\Delta\bar{G}^{\circ}_{SC,\infty}$ for each of the osmolytes on the molar and mole fraction concentration scales based on equal masses of the two solvents. At high solute solubilities, $\Delta\bar{G}^{\circ}_{SC,\infty}$ will not completely correct for differences between the transfer free energies evaluated on different concentration scales because the solute makes a significant contribution to the solution density and the mole fraction.

Table 1: The Concentration Scale Correction Free Energy, $\Delta\bar{G}^{\circ}_{int}$ (cal/mol), Corresponding to 100 g of Solvent Evaluated at Infinite Dilution of the Solute at 25 °C

1 M osmolyte	molarity $RT\ln(d_{w}/d_{os})$	mole fraction $RT\ln((N_{w,os} + N_{os})/N_{w})$
TMAO	−1.47	−34.84
sarcosine	−14.94	−42.69
betaine	−10.94	−60.84
proline	−19.17	−58.66
glycerol	−12.60	−44.74
sorbitol	−36.19	−99.70
sucrose	−71.58	−201.43
trehalose	−73.86	−200.51
urea	−9.25	−25.12

Scheme 1



Operationally, eq 14 suggests a means for identifying those model compounds having solubilities low enough that $\Delta\bar{G}^{\circ}_{int}$ is numerically the same in each of the concentration scales. For such model compounds, the solutions begin to approach ideality with respect to the solute when the ratio of the solute activity coefficients on different concentration scales approaches unity, eq 15, and the experimentally measured $\Delta\bar{G}^{\circ}_{int}$ becomes equal to $\Delta\mu^{\circ}$ in that limit.

$$\text{In the } \lim_{n_i \rightarrow 0} \quad (\gamma_{i,w}/\gamma_{i,os})_{c} = (\gamma_{i,w}/\gamma_{i,os})_{m} = (\gamma_{i,w}/\gamma_{i,os})_{x} = 1 \quad (15)$$

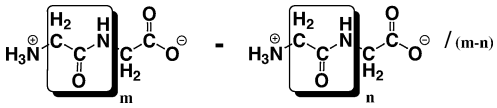
Model Systems and Conventions for Determining the Peptide Unit Transfer Free Energy

Cyclic Glycylglycine (CGG). CGG, first developed as a peptide model by Gill, has been the preferred model of the peptide backbone unit in our lab because it contains two peptide groups in a structure that exposes the peptide unit to the solvent (*35*). The transfer free energy of this compound from water to 1 M osmolyte is divided by 2 to obtain the contribution of one peptide unit as shown in Scheme 1.

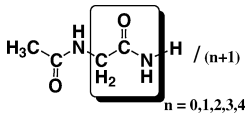
This divisional construct reports an average value for the transfer free energy of each of the peptide units in the model peptide. Because there are no end group effects with CGG, the peptide unit transfer free energy for this compound should be an accurate representation of the actual transfer free energy, though the transfer free energy will depend on which concentration scale is used.

Zwitterionic Glycine Series. The use of glycine peptides to ascertain the transfer free energy contribution of a single peptide unit was Nozaki and Tanford's model of choice (*9*, *10*, *22*, *23*). To evaluate the transfer free energy of the peptide backbone unit, these authors subtracted the transfer free energies of glycine peptides that differ in their length by $m - n$ peptide units, as shown in Scheme 2, where m and n are the number of internal peptide units in the glycine peptides such that $m \geq n + 1$. This subtraction provides a

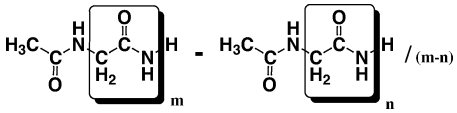
Scheme 2



Scheme 3



Scheme 4



mathematical means to remove the contribution of the end groups. Two types of mathematical constructs arise from such a scheme, simple subtractional constructs for which $m = n + 1$ and composite constructs for which $m > n + 1$. The simple subtractional constructs consist of subtracting the transfer free energies of two peptides, such as Gly_2 and Gly, that differ in length by one peptide unit, whereas the composite constructs consist of subtracting the transfer free energies of two peptides that differ in length by more that one peptide unit, such as Gly_4 and Gly, and then dividing the difference by the number of peptide units, three, remaining. The expressed purpose of the subtractive and composite constructs is to eliminate the transfer free energy contribution of end groups and evaluate the transfer free energy of an interior peptide backbone unit (*24*).

N-Acetylglycinamide Series: The *N*-acetylglycinamide series is analogous to the zwitterionic glycine series with the exception that the charged end groups have been blocked by an acetyl and an amide moiety. In these models, the end groups are chemically related to an internal peptide unit, so the transfer free energy of the peptide unit can be approximated by dividing by the number of internal units plus one, shown in Scheme 3, where *n* is the number of internal peptide units that in our experiments range from 0, for the transfer free energy of acetamide, to 4, for the transfer free energy of *N*-acetyltetraglycinamide. The divisional constructs provide an average transfer free energy of the peptide unit that contains contributions from the end groups.

Just as with the zwitterionic glycine series, the transfer of a single peptide unit from water to 1 M osmolyte can be evaluated as shown in Scheme 4 where $m \geq n + 1$, by first subtracting the transfer free energies of two peptides that differ in their length by $m - n$ peptide units and then dividing by $m - n$ pepide units.

All of these mathematical constructs for each of the model peptide systems provide a prediction for the peptide unit transfer free energy, but whether they provide reliable values of the peptide backbone transfer free energy remains to be tested.

RESULTS

Cyclic Glycylglycine. Figure 1 shows results for the transfer free energy of the peptide unit as evaluated using the CGG model. The data illustrate three important features

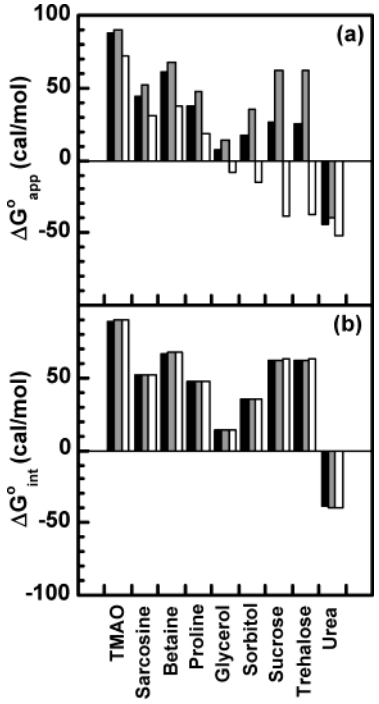


FIGURE 1: Transfer free energy of the peptide backbone unit (cal/mol) from water to 1 M of the indicated osmolytes defined by the divisional construct of cyclic glycylglycine, Scheme 1 (black = molarity scale, gray = molality scale, and white = mole fraction scale): (a) apparent transfer free energy, $\Delta\bar{G}^{\circ}_{app}$, before concentration scale correction for the solubility at infinite dilution; (b) intrinsic transfer free energy, $\Delta\bar{G}^{\circ}_{int}$, after the concentration scale correction for the solubility at infinite dilution.

concerning the peptide unit transfer free energy. First, panel a shows the apparent transfer free energy of the peptide unit before making the $\Delta\bar{G}^{\circ}_{SC,\infty}$ correction. These data demonstrate a marked difference between $\Delta\bar{G}^{\circ}_{app}$ calculated on different concentration scales. For TMAO, sarcosine, betaine, proline, and urea, there are significant differences in the magnitude of the free energy, while for the polyols, glycerol, sorbitol, sucrose, and trehalose, even the sign of the free energy is changed using the mole fraction scale. Second, as is shown in panel b, these differences in magnitude and sign of the peptide unit transfer free energy are completely accounted for upon applying the $\Delta\bar{G}^{\circ}_{SC,\infty}$ correction of Table 1 using eq 14, and all concentration scales conform to the same value. This shows that the differences in $\Delta\bar{G}^{\circ}_{app}$ are due only to differences in the definitions of the concentration scales. The fact that $\Delta\bar{G}^{\circ}_{int}$ on each concentration scale gives the same numerical result tells us that in water and the respective osmolyte solutions the solubilities of this model compound are so low that CGG contributes little to the solution density and the mole fraction. Third, the intrinsic transfer free energies of panel b show a pattern that indicates that all protecting osmolytes interact unfavorably ($\Delta G_{transfer}$ is positive) with the peptide backbone, while the nonprotecting osmolyte, urea, interacts favorably ($\Delta G_{transfer}$ is negative) with the peptide backbone. Being a cyclic dipeptide, CGG is unique in comparison with all of the linear chain models in that there are no end effects with which to be concerned. Accordingly, the divisional construct is a proper construct for CGG; this does not hold true for linear peptide models.

Zwitterionic Glycine Series. Figure 2 gives the intrinsic transfer free energy of the peptide unit obtained for the

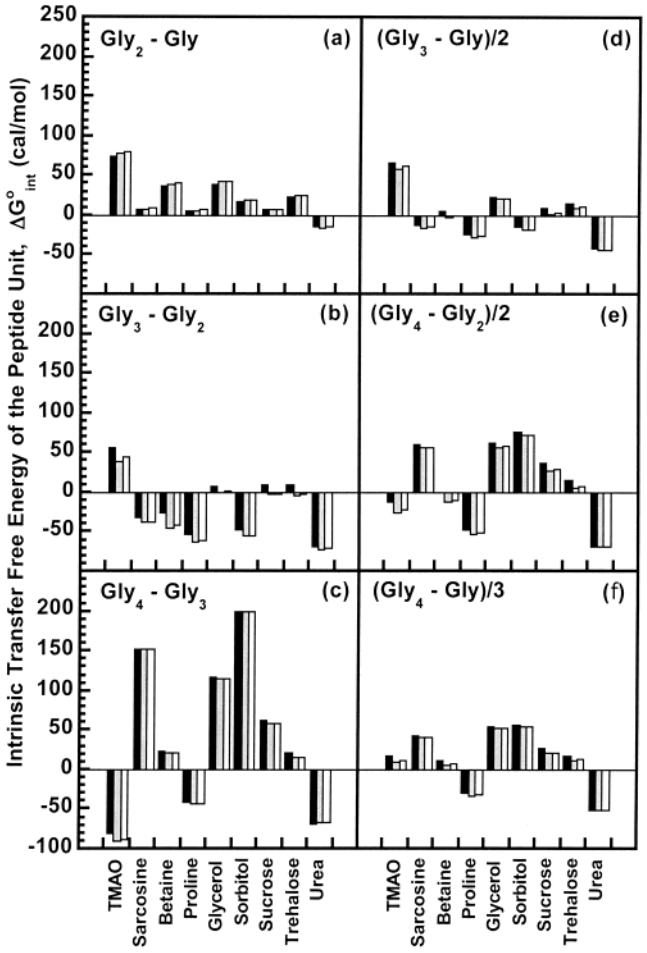


FIGURE 2: The intrinsic transfer free energy of the peptide backbone unit, $\Delta\bar{G}^{\circ}_{int}$ (cal/mol), from water to 1 M of the indicated osmolytes defined by the simple subtractional constructs (a,b,c) and the composite constructs (d,e,f) of the zwitterionic glycine series, Scheme 2 (black = molarity scale, gray = molality scale, and white = mole fraction scale).

zwitterionic glycine series of model peptides. Because these were the compounds used by Nozaki and Tanford in their studies, we will use the same mathematical constructs with our data. A major difference between the two studies is that Nozaki and Tanford focused on the transfer to urea, while we use a much larger variety of naturally occurring osmolytes. Shown in panels a, b, and c are the peptide unit transfer free energies obtained from the subtractional constructs, while the composite constructs are given in panels d, e, and f. For all of the protecting osmolytes, the peptide unit transfer free energy varies both in sign and in magnitude from one construct to another regardless of the concentration scale used to describe the system. Only for the nonprotecting osmolyte, urea, is the peptide unit transfer free energy consistently favorable ($\Delta G_{transfer}$ is negative) although significant differences in magnitude remain among the constructs. The results in Figure 2 show that the zwitterionic glycine series and mathematical constructs fail to give any clear pattern or basis for deciding which transfer free energies to believe. From these data alone, it is impossible to tell what is responsible for the observed variation in the data, whether it is due to the inadequacies of the mathematical constructs, the choice of concentration scale, the chain length not being long enough to eliminate the effects of charged end groups, lack of activity coefficient corrections, or some combination of effects. Near the end of the Results section, we will provide evidence that it is the high solubility, the lack of activity coefficient corrections, and the charged end group effects of zwitterionic glycine model compounds that are responsible for the erratic nature of the data shown in Figure 2.

N-Acetylglycinamide Series, Shown in Figure 3 are the intrinsic transfer free energies of the *N*-acetylglycinamide series obtained using the divisional constructs plotted as a function of the number of internal peptide units present in the model peptide. Here, we arbitrarily take the N-terminal acetamide, CH_3CONH, and C-terminal H as end groups and define CH_2CONH as the internal peptide unit (Scheme 3). Plotting the data as in Figure 3 illustrates three important points. First, for each of the osmolytes, $\Delta\bar{G}^{\circ}_{int}$ approaches a limiting value for the peptide unit as the chain length of the model peptide is increased, and second, this limiting value is independent of concentration scale at long chain length. For peptides with small chain lengths, acetamide and *N*-acetylglycinamide, the solubility is high, resulting in large differences in $\Delta\bar{G}^{\circ}_{int}$ between concentration scales, while at longer chain lengths, the lower solubility of the peptides causes $\Delta\bar{G}^{\circ}_{int}$ values to converge. This indicates that for peptide lengths longer than two internal units, the solubility is low enough that the peptides contribute negligibly to the solution density and the total number of moles in solution. The third and final point is that like the cyclic glycylglycine divisional construct of Figure 1, the $\Delta\bar{G}^{\circ}_{int}$ for protecting osmolytes all approach unfavorable transfer free energies whereas the nonprotecting osmolyte, urea, clearly gives a favorable transfer free energy. While the divisional construct is useful for providing insight into the relative adequacies of the model compounds as a function of chain length, it assumes that there are no end group effects on the transfer free energy. Thus, the absolute value of the transfer free energy even at the longest chain length in our study may not be the best value for the peptide unit.

The subtractional and composite constructs of Nozaki and Tanford were also applied to the *N*-acetylglycinamide series of model compounds. $\Delta\bar{G}^{\circ}_{int}$ values for all possible subtractional and composite constructs with the series are plotted in Figure 4 as a function of the number of internal units left after the subtraction and before the division for each of the indicated osmolytes. For example, at one internal unit, $\Delta\bar{G}^{\circ}_{int}$ values for each of the subtractive constructs are plotted for each concentration scale, along with composite constructs for two, three, and four internal units. These data illustrate two important points. First, the scatter in the transfer free energies of the peptide unit defined by these constructs is initially quite large and can range over 100 to ~300 cal/mol, spanning both positive and negative values. In large part, the variations are due to the high solubility of many of the participants and the resulting discrepancies due to choice of concentration scale. As the number of internal peptide units increases, this scatter becomes progressively smaller and convergence occurs because only one compound meets the requirement of having four internal peptide units. Variation between concentration scales is still evident at four internal peptide units because even at the highest chain length a low molecular weight (and highly soluble) model compound must be subtracted. For this reason, the converged

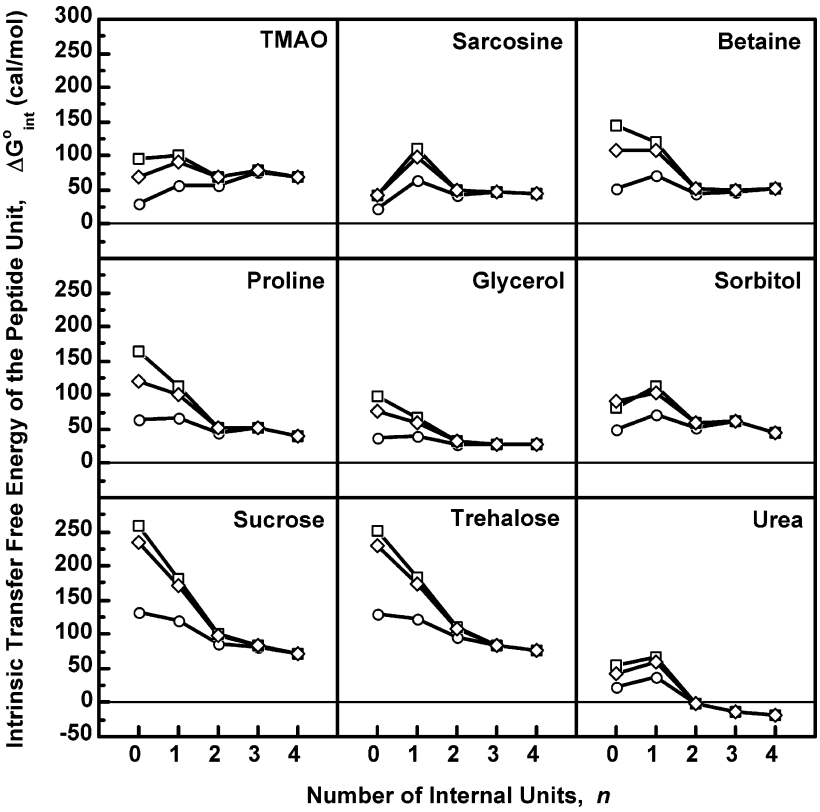


FIGURE 3: The average intrinsic transfer free energy of the peptide backbone unit, $\Delta\bar{G}^{\circ}_{int}$ (cal/mol), from water to 1 M of the indicated osmolytes defined by the divisional constructs of the *N*-acetylglycinamide series, Scheme 3, as a function of the number of internal peptide units, *n*: (○) molarity scale; (□) molality scale; (◇) mole fraction scale.

value will not be the best value for the peptide backbone transfer free energy. Second, despite these problems the composite constructs demonstrate that the protecting osmolytes approach unfavorable free energies for peptide backbone transfer, while the nonprotecting osmolyte, urea, approaches a favorable transfer free energy.

The Constant Increment Method. With model systems of increasing chain length, Nandi and Robinson used an additional construct (*30*), which we will call the constant increment method. This method determines the transfer free energy per peptide unit from the slope of a plot of the peptide transfer free energy versus the number of internal peptide units, and Figure 5 gives such data for the *N*-acetylglycinamide series of model peptides using the molar, molal, and mole fraction scales. According to the transfer model, the plots shown in Figure 5 may be represented by eq 16.

$$\Delta G_{tr,Peptide} = \Delta G_{EG} + N_{IU}(\partial\Delta G_{tr}/\partial N_{IU}) \qquad (16)$$

where $\Delta G_{tr,Peptide}$ is the intrinsic transfer free energy of the model peptide, ΔG_{EG} is the transfer free energy contribution of the end groups, N_{IU} is the number of internal peptide units in the model peptide, and $(\partial\Delta G_{tr}/\partial N_{IU})$ is the transfer free energy per internal peptide unit.

As expected from the previous figures, the shorter peptides, acetamide and *N*-acetylglycinamide, show large deviations due to the higher solubility of short peptides and because solvation end effects dominate short chain length transfer Gibbs energies. But for peptides longer than two internal units in length, the slope of the plot shows that the transfer free energy changes by the same increment for each peptide unit added in chain length. This constant increment in transfer free energy per peptide unit demonstrates that the peptide unit transfer free energy from water to 1 M of each of the osmolytes is additive and that this additivity is unfavorable, $(\partial\Delta G_{tr}/\partial N_{IU} = (+))$, for protecting osmolytes and favorable, $(\partial\Delta G_{tr}/\partial N_{IU} = (-))$, for urea.

Among all of the plots in Figure 5, urea, glycerol, betaine, and sarcosine show the least variation in $\Delta\bar{G}^{\circ}_{int}$ at two internal peptide units. For these osmolytes, we use the linear best fits of the data involving ≥2 internal units to obtain the slope. Upon extrapolation of the resulting lines to the ordinate, it is observed that the molar concentration scale data fit the respective lines better than do the other concentration scales. We take this as empirical evidence that of the three scales the molar scale best represents the $\Delta\bar{G}^{\circ}_{int}$ value at two internal peptide units. Accordingly, in all cases in Figure 5, we use the molar scale $\Delta\bar{G}^{\circ}_{int}$ values to obtain linear fits to data ≥2 internal peptide units and report the slopes, $(\partial\Delta G_{tr}/\partial N_{IU})$, as constant increment values.

Activity Coefficients and Charged End Group Effects. As shown in eq 14, $\Delta\bar{G}^{\circ}_{int}$ becomes equal to the change in standard state chemical potential for transfer, $\Delta_t\mu^{\circ}$, only when the activity coefficient ratio for the solute in the two solvents is unity. In the interest of knowing at what concentrations the activity coefficient ratio becomes unity, we turn to activity coefficient data presented in the literature.

While activity coefficient data on three component systems is scarce, fortunately their exists enough measurements involving the zwitterionic glycine series (Gly, Gly_2, and Gly_3) in water and in 1 M urea solution to apply to our study (*36*−*40*). Solubilities of the zwitterionic series along with that of the *N*-acetylglycinamide series shown in panel a of Figure

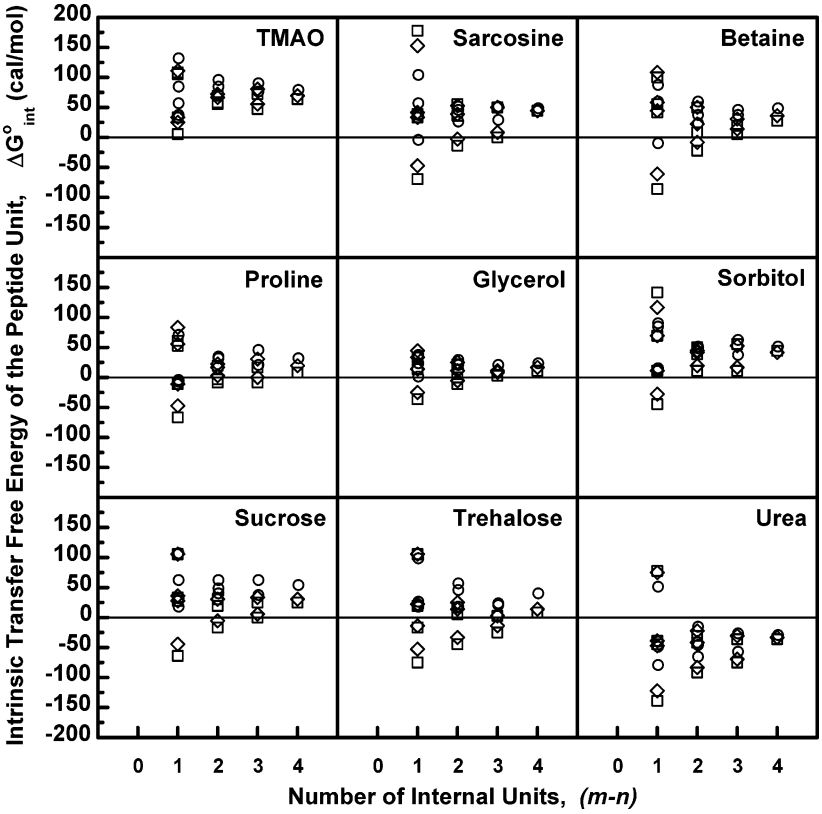


FIGURE 4: The intrinsic transfer free energy of the peptide backbone unit, $\Delta\bar{G}^{\circ}_{int}$ (cal/mol), from water to 1 M of the indicated osmolytes defined by the subtractional and composite constructs of the *N*-acetylglycinamide series, Scheme 4, as a function of the number of internal peptide units ($m - n$): (○) molarity scale; (□) molality scale; (◇) mole fraction scale. Shown are simple subtractional constructs, ($m - n$) = 1, and composite constructs, ($m - n$) > 1.

6 are all observed to decrease exponentially with chain length. As a function of molal concentration, panel b shows the activity coefficients of Gly, Gly_2, and Gly_3 in water (lines 5, 3, and 1, respectively) and in 1 M urea (lines, 6, 4, and 2 respectively). The activity coefficients for the zwitterionic model compounds approach unity only at very low concentrations, and they are much lower than 1.0 over the concentration range covered by the solubility limits. However, it must be recognized that it is the ratio of activity coefficients of the compounds in water and 1 M urea that is important and not the absolute value of the activity coefficients. The activity coefficient ratios in water and 1 M urea at the solubility limits of the compounds are calculated from the data (Figure 6b) to be very close but not exactly equal to unity. Knowing the ratio of activity coefficients for Gly, Gly_2, and Gly_3 at their solubility limits along with their $\Delta\bar{G}^{\circ}_{int}$ values, permits evaluation of $\Delta\mu^{\circ}$ for each of the compounds, using eq 14. These data are shown as open squares in Figure 7 along with the datum of Gly_4, the solubility of which is low enough that the activity coefficient ratio can legitimately be taken as 1.0.

It has long been considered that the zwitterionic glycine series may not be good models for obtaining the transfer free energy of the peptide backbone due to the long-range electrostatic attraction of the formal charges on the ends of the molecules (*6*). In the case of dipolar ions such as glycine, diglycine, etc., the electrostatic part of the free energy of their transfer from water to a medium of different dielectric constant can be approximated by use of the Scatchard and Kirkwood expression (*41*), eq 17,

$$\Delta G_{tr,dipole} = \frac{N_A z_1 z_2 e^2}{4\pi\epsilon_0}\left(\frac{1}{b} - \frac{1}{R}\right)\left(\frac{1}{D} - \frac{1}{D_0}\right) \qquad (17)$$

where N_A is Avogadro's number, ϵ_0 is the permittivity constant of a vacuum, e is the elementary charge of an electron, and z_1 and z_2 are the unit charges on the carboxylate, (−1), and amino groups, (+1). b is the ion radius, R is the distance separating the centers of the carboxylate and amino groups, and D_0 and D are the dielectric constants of water and 1 M urea, respectively (*42*). Following Cohn and Edsall, we take $b = 1.24$ Å (*43*) and evaluate R values for mono- through tetraglycine, from the dipole moments (μ_0) of these compounds using $\mu_0 = 0.79\sqrt{\mu\cdot\bar{\mu}}$(*44*), where $\sqrt{\mu\cdot\bar{\mu}}$ values for the compounds are taken from Table IV of Edsall and Wyman (*45*). The dipolar ion contributions to the transfer free energies of these compounds from water to 1 M urea are given as open circles in Figure 7. Subtracting these contributions from the overall transfer free energies (open squares) gives a linear dependence (open triangles) of the remaining transfer free energy of these model compounds as a function of the number of peptide units. The linear dependence is presumed to reflect the peptide backbone transfer free energy of the glycine series of compounds with removal of electrostatic effects on the transfer free energy due to the ionized end groups. The slope of this linear dependence gives −56 cal/mole for the transfer of the peptide backbone unit from water to 1 M urea.

DISCUSSION

Concentration Scales. In performing transfer free energy measurements, one is immediately faced with the question

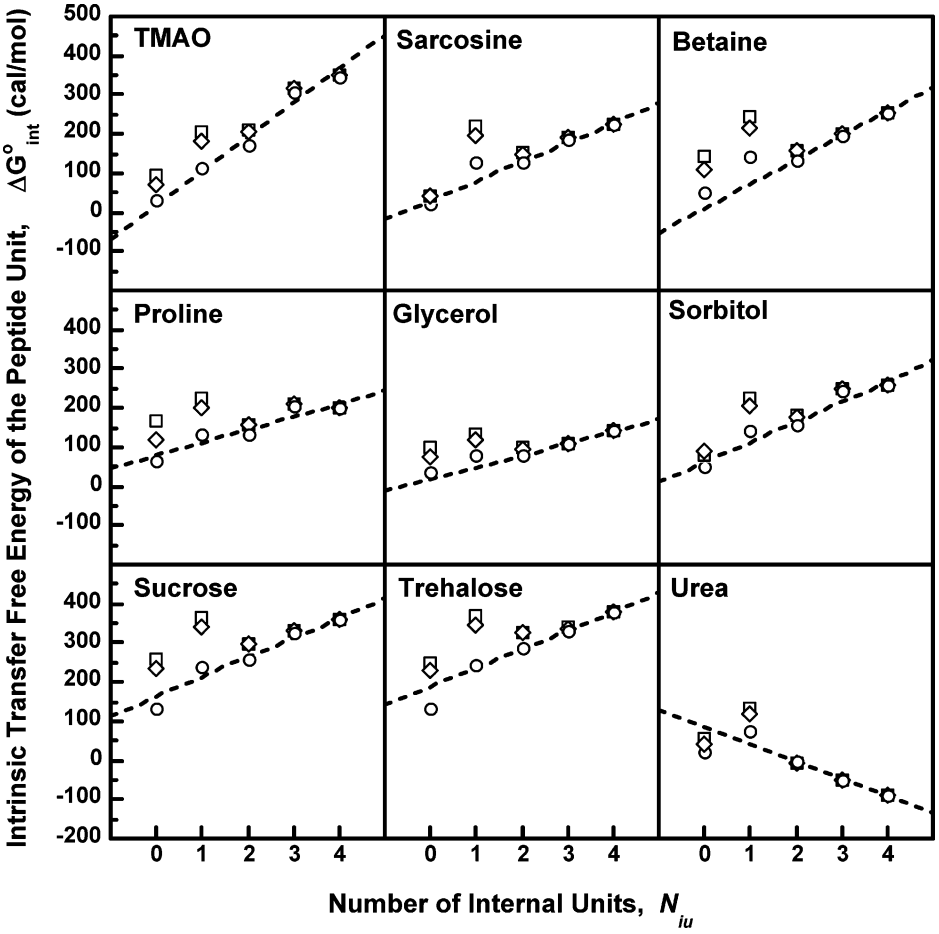


FIGURE 5: The intrinsic transfer free energy of the peptide backbone unit, $\Delta\bar{G}^{\circ}_{int}$ (cal/mol), from water to 1 M of the indicated osmolytes defined by the constant increment method of the *N*-acetylglycinamide series, eq 16, as a function of the number of internal peptide units, N_{IU}: (○) molarity scale; (□) molality scale; (◇) mole fraction scale. The slope of the linear fit through the data for two, three, and four internal units gives the transfer free energy per peptide unit. The data for 0 and 1 internal units were not used for the linear fit. The slopes, $(\partial\Delta G_{tr}/\partial N_{IU})$, are given in Table 2.

of choosing a concentration scale to use. Figure 1a shows that for transfer of CGG from water to 1 M concentrations of either glycerol, sorbitol, sucrose, or trehalose, the mole fraction scale gives a favorable transfer free energy while the molal and molar scales show the transfer is thermodynamically unfavorable! Others have noted the same disconcerting finding that for the same compound, different concentration scales can give transfer free energies that differ in sign (*32*). Cohn and Edsall, Kauzmann, Tanford, and others have argued the mole fraction scale as the most suitable for transfer free energies (*6*, *8*, *15*, *29*, *46*), while Robinson and Jencks promoted use of the molarity scale (*11*, *47*). On the basis of a statistical mechanics treatment of the problem, Ben-Naim has argued strongly in favor of the molarity scale and against the mole fraction scale (*31*).

There are two ways that the choice of concentration scale can affect values reported for transfer free energy; one is observed with compounds of low solubility, and the other can best be seen with highly soluble compounds. For a solute of low solubility, we mean low enough that the compound does not contribute substantively to the total moles of the solution or the density of the solution. For such solutes, the transfer free energy differences between the various concentration scales arise entirely from differences in the densities of the two solvents and differences in the total moles of species in the solvents (eqs 8, 9, and 10). Upon choosing a basis for comparison (e.g., equal masses or equal volumes of the solvents) and evaluating the mole ratio of the model compound in each solvent, one can readily make corrections for the differences between the two solvents, and the various concentration scales will all give the same transfer free energy for the basis selected when the model compound solubility is low in both solvents. Figure 3 shows that this low solubility condition is reasonably met for chain lengths of 2, 3, and 4 internal peptide units on transferring the *N*-acetylglycinamide series to a variety of 1 M osmolyte solutions. The transfer free energy results for compounds of low solubility in the model series are to be contrasted with the results for highly soluble model compounds in the series. The model compounds representing zero and one internal peptide units in the *N*-acetylglycinamide series (Figure 3) are soluble enough that they contribute significantly to the density and total moles of their respective saturated solutions, and though corrections for solvent differences have already been applied, the data show that concentration scale dependent contributions remain. These residual concentration scale effects are, no doubt, responsible for much of the lack of agreement of transfer free energy results between different labs dealing with highly soluble model compounds and perhaps different chemical models.

Nowhere is the influence of residual concentration effects more evident than the data shown in Figure 4 under the

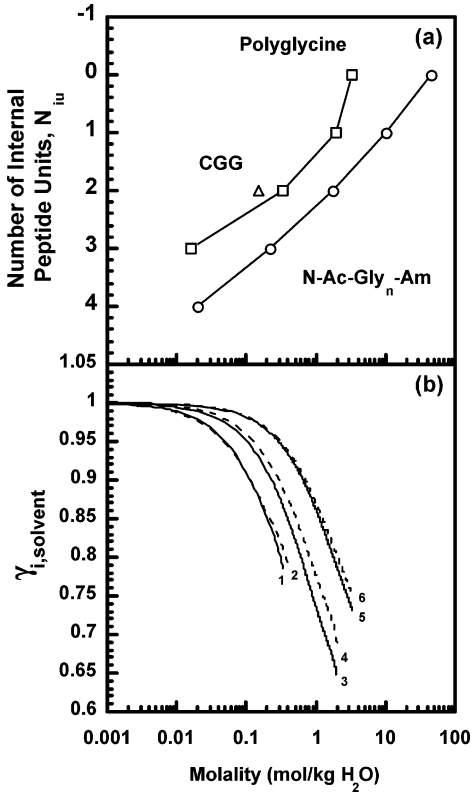


FIGURE 6: Comparison of model peptide solubilities on the molality per kilogram of water concentration scale (a) with molal activity coefficients of polyglycine−water−urea ternary systems reported in the literature (b): (a) solubilities of model peptides in water as a function of chain length for (△) cyclic glycylglycine, (□) zwitterionic polyglycine series, and (○) *N*-acetylglycinamide series; (b) activity coefficients for polyglycine−water−urea ternary systems on the molal per kilogram of water concentration scale: (1) $\gamma_{Gly_3,Water}$ (*36*, *37*); (2) $\gamma_{Gly_3,1M\ Urea}$ (*36*); (3) $\gamma_{Gly_2,Water}$ (*36*−*38*); (4) $\gamma_{Gly_2,1M\ Urea}$ (*36*, *39*); (5) $\gamma_{Gly,Water}$ (*38*, *40*); (6) $\gamma_{Gly,1M\ Urea}$ (*40*).

category of one internal unit for each of the nine osmolytes. The highly scattered data reported in this category were obtained by subtracting $\Delta\bar{G}^\circ_{int}$ of any one of the *N*-acetylglycinamide model compounds from $\Delta\bar{G}^\circ_{int}$ of the model compound one peptide unit longer. While a small number of the points represent transfer free energy differences between model compounds that fit in the low solubility regime, the majority of the points arise from compounds whose solubility is significant enough to cause changes in solution density and contribute a significant number of moles to the total solution. The high solubility of the relevant model compounds is what introduces this particular type of concentration scale dependence and generates a large range of values spanning both negative and positive $\Delta\bar{G}^\circ_{int}$ values. Given that these are the procedures used by many in reporting peptide backbone transfer free energies and that different concentration scales and chemical models were frequently used, it is not surprising that there can be significant lack of agreement between laboratories as to the transfer free energy of the peptide backbone unit.

Mathematical Constructs, Constant Increment Method, and Additivity. The example above demonstrates that when applied to compounds of high solubility, the subtractive construct provides transfer free energies that are not reliable. The same can be said in the use of the composite constructs if one of the compounds involved is highly soluble. The only possible conditions under which the subtractive and composite constructs can give reliable transfer free energies is when both compounds involved have low solubility.

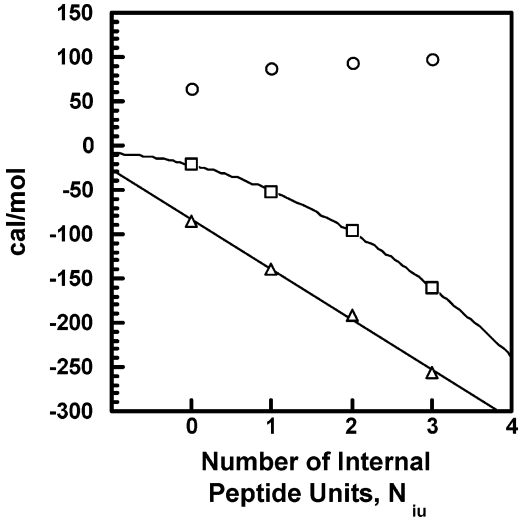


FIGURE 7: The transfer chemical potential, $\Delta\mu^\circ$, of the zwitterionic polyglycine series (cal/mol) from water to 1 M urea after accounting for model peptide solubilities and activity coefficients at the solubility limit (□) and the difference between electrostatic energies due to charged end group attraction within the peptides in water and in 1 M urea as evaluated using the Scatchard and Kirkwood expression, eq 17 (○). The distance, *R*, between charged end groups was obtained from the dipole moments listed in Table IV of Wyman and Edsall using $R = \mu_o(3.34 \times 10^{-20})/e$, where *e* is the elementary charge of an electron. $R_{Gly} = 2.75$ Å; $R_{Gly_2} = 4.71$ Å; $R_{Gly_3} = 6.13$ Å; $R_{Gly_4} = 7.02$ Å. The dielectric constant was taken to be that for the solvents water, 78.54, and 1 M urea, 81.36 (*42*), at 25 °C. Triangles are the difference between the squares and the circles, which approximately removes the electrostatic contributions of the charged end groups from the transfer chemical potential. The slope of the triangle data, −56 cal/mol, represents the transfer free energy of the peptide unit free from charged end group effects.

These two constructs, as well as the divisional construct, also have a bearing on the question of additivity. Robinson and Jencks (*11*) pointed out that because of their reciprocal relationship, subtracting transfer free energies of compounds whose structures differ by one additional functional group, implicitly assumes that the transfer free energies of the chemical groups are additive. The same is true of the divisional construct. Accordingly, if group additivity indeed occurs, it must be demonstrated by means of empirical observation. Through use of the constant increment method (Figure 5), such evidence is provided by the transfer free energies of *N*-acetylglycinamide compounds having ≥2 internal peptide units. The slope of the plot $(\partial\Delta G_{tr}/\partial N_{IU})$ illustrates two points. The first is that the transfer free energy increases by the same constant increment with each additional peptide unit, thus demonstrating additivity. The second is that the same slope will be obtained regardless of which concentration scale is selected as the base. That is, the constant increment method gives concentration scale independent evaluations of the transfer free energy of the chemical grouping of interest. When evaluated in the manner described, such transfer free energy changes are extremely important because they make the practical necessity of choosing a concentration scale irrelevant as a problem.

In effect, the constant increment method is a version of the subtractive and composite constructs. Although there will be larger error in doing so, one will get the same values for

Table 2: The Intrinsic Transfer Free Energy, $\Delta\bar{G}^{\circ}_{int}$ (cal/mol), of the Peptide Unit Defined by Cyclic Glycylglycine, CGG/2, and by the Constant Increment Method, $\partial\Delta G_{tr}/\partial N_{IU}$

1 M osmolyte	CGG/2[a]	$\partial\Delta G_{tr}/\partial N_{IU}$
TMAO	90	87
sarcosine	52	48
betaine	67	62
proline	48	33
glycerol	14	31
sorbitol	35	52
sucrose	62	50
trehalose	62	47
urea	−39	−43

[a] Calculations on an equal mass basis for both solvents.

the transfer free energy either using the constant increment method or applying the subtractive method to *N*-acetylglycinamide model compounds having ≥2 internal peptide units. By the same token, applying the composite construct to the transfer free energies of *N*-acetylglycinamide compounds having two and four internal peptide units and dividing by two give the same transfer free energy per peptide unit as that provided by the constant increment method. In summary, when applied to congener compounds of sufficient chain lengths such that the solubilities of both compounds are very low, the subtractive and composite constructs can give transfer free energy data that are valid and exhibit additivity.

Activity Coefficients and Transfer Free Energies from Different Chemical Models. Literature data on activity coefficients of the zwitterionic glycine series in water and 1 M urea are useful for evaluating the relative contributions of nonideality to transfer free energies. It is again worth emphasizing that it is the ratio of activity coefficients of the solute in the two solvents that is important in this connection and not the activity coefficients themselves. The corrections due to activity coefficient ratios for glycine, diglycine, and triglycine are −15.1, −30.0, and −5.1 cal/mol, respectively (evaluated from data given in Figure 6b). When one considers that the activity coefficient ratios become even smaller with increasing chain length, it is evident that activity coefficient contributions will be negligibly small for zwitterionic compounds >2 internal peptide units in length. Given these results, it is reasonable to assume that activity coefficient corrections will also be negligible for those members of the *N*-acetylglycinamide series that have solubilities in the same range as triglycine and tetraglycine and that the constant increment method is essentially free of activity coefficient effects within this range. Because CGG also falls within this concentration range (Figure 6), it too should be expected to have no activity coefficient contributions to its free energy of transfer.

Table 2 shows free energies of peptide unit transfer from water to 1 M concentrations of nine osmolytes, using CGG and constant increment values for the *N*-acetylglycinamide series as model systems. Values listed for CGG are single quantities obtained from solubility plots in which a line obtained from compositions below the solubility limit intersects with the line formed using compositions above the solubility limits (*18*, *20*). Consequently, we can only estimate the errors associated with the measurements. Fitting errors obtained for the point of intersection gives ±16 cal/mol of peptide unit for trehalose and ±10 cal/mol of peptide unit for urea and all other osmolytes. The transfer free energy of CGG from water to 1 M TMAO reported in Table 2 is comparable to that reported by Zou et. al., though the error we obtain is about two-thirds the error reported by them (*48*). Solubility plots for glycine obtained in this laboratory over a period of several years suggests that ±10 cal/mol of peptide unit can reasonably be expected as the range of experimental errors of transfer free energies obtained using our methods for determining solubility limits. Applying these error limits to the points shown in Figure 5 and propagating the errors to evaluate the slopes from the last three points in each plot give fitting errors for the data shown for ($\partial\Delta G_{tr}/\partial N_{IU}$) in Table 2. The fitting errors in ($\partial\Delta G_{tr}/\partial N_{IU}$) are well below 10 cal/mol of peptide backbone unit for all solutes except TMAO, proline, and sorbitol, for which the errors are 18, 11, and 20 cal/mol peptide unit, respectively.

Within the estimated error limits, the free energy values in Table 2 for CGG and the *N*-acetylglycinamide series are in excellent agreement for transfer of the peptide unit from water to 1 M TMAO or urea. The transfers are from water to the naturally occurring osmolytes that have the largest effects on proteins, TMAO being the most potent protein stabilizer and urea the most effective protein destabilizer. Though there is somewhat less agreement between the transfer free energy values involving some of the remaining osmolytes, the agreement is still quite good considering that the maximum transfer free energy difference between models is only 17 cal/mol of peptide unit. Given that a change of 1% in the mole ratio of the solubilities changes the transfer free energy by about 6 cal/mol/peptide unit, very small experimental errors can readily account for the observed differences. The two chemical models in Table 2 are chemically quite different, yet their peptide backbone transfer free energies to a diverse array of solvents are very much in agreement. We interpret this to mean that the peptide backbone transfer free energies in Table 2 are independent of the chemical model used for their determination. To our knowledge, this is the first broad-based demonstration of (chemical) model independent transfer free energies of the peptide backbone unit.

Which concentration scale should be used in evaluating transfer free energies is a question of long standing, and it can be addressed by comparing concentration scale dependent transfer free energies with those obtained by means of the concentration scale independent constant increment method. In fact, the molal concentration scale is the basis for the CGG transfer free energy results reported in Table 2, and different values are obtained if equal volumes of solvents (molar) or equal total moles of solvents (mole fraction) are chosen as the bases for the transfer free energy determinations. Data in Table 1 provide the factors to be used in converting from the CGG transfer free energies of Table 2, based on the molal scale, to the molar and mole fraction based CGG transfer free energies. To obtain the molar scale transfer free energies, values in column two of Table 1 are first divided by two (to account for CGG having two peptide units), and the result is added to the corresponding molal-based value in column 2 of Table 2. Using column 3 of Table 1, we perform similar operations as above in converting the molal-based values in Table 2 to mole fraction based quantities. Because many of the mole fraction based CGG transfer free energies are found to be opposite in sign from the values determined by means of the constant increment method, it is abundantly clear that

mole fraction is not a proper scale to use in transfer free energy measurements. By contrast, the molar-based CGG transfer free energies deviate less than 5 cal/mol of peptide unit from the constant increment method values in the cases of TMAO, sarcosine, betaine, proline, and urea but differ by ~23 cal/mol of peptide unit for glycerol, trehalose, and sucrose and 35 cal/mol of peptide unit for sorbitol. Empirically, the molal-based CGG peptide unit transfer free energies agree better with the constant increment values but not sufficiently enough to exclude the molar-based quantities. To rigorously test the question of preference regarding molal- and molar-based transfer free energies or whether neither is adequate in all cases, experiments performed in solvents of widely differing densities will be required.

The polyglycine model series originally used by Cohn and Edsall and by Nozaki and Tanford is the first model compound series to be used in obtaining transfer free energy data for the peptide backbone unit. With activity coefficient data on glycine, diglycine and triglycine in water and 1 M urea and apparent transfer free energy data from solubility measurements that we obtained for the glycine model compound series, it is possible to compare the peptide backbone unit transfer free energy values from water to 1 M urea obtained using the glycine series with those obtained using CGG and using the *N*-acetylglycinamide model compound series with the constant increment method. Figure 7 shows that the application of activity coefficient corrections to the zwitterionic glycine model compound transfer free energies remains nonlinear (open squares), demonstrating that this model compound series does not exhibit the group additivity required for valid application of the transfer model for protein unfolding (eq 1). However, if the transfer free energies are corrected for the dipolar ion contribution to the transfer by using the Scatchard and Kirkwood equation, the remaining parts of the transfer free energies are observed to become a linear function of the number of internal peptide backbone units, and the transfer free energy obtained from the slope is −56 cal/mol of peptide unit. That is, attempts to correct for the charged end group effects of the glycine zwitterionic series result in a peptide unit transfer free energy that exhibits additivity and is not too far different from the −39 and −43 cal/mol of peptide unit transfer free energy values obtained using CGG and the *N*-acetylglycinamide series shown in Table 2.

We have no way of evaluating the accuracy of the Scatchard and Kirkwood corrections, and we freely acknowledge that considerable work would be necessary to explore the many facets of the dumbbell model of Scatchard and Kirkwood, as well as other models that could be used to evaluate the differences in transfer free energies of dipolar species from water to 1 M urea. It is of interest, however, that the corrections due to end group effects are mathematically of the correct functional form to give what appears to be a linear dependence of peptide unit transfer free energy with the number of internal peptide units of the zwitterionic series, that the free energy becomes additive as a result of the corrections, and that the magnitude of the transfer free energy is near that of other model compounds. If the Scatchard and Kirkwood model gives a good representation of the dipolar ion effects on the transfer free energy, then the zwitterionic model compounds yield essentially the same results as those of the other two model systems, thus providing additional evidence that the peptide unit transfer free energy is independent of the choice of model compound.

CONCLUSIONS

The strengths of the constant increment method are that it makes no difference which of the concentration scales are used, additivity is a demonstrable requirement of its application, and it cleanly handles virtually all of the issues to date that have prevented the quantitative development of the transfer model. The work described here indicates that with careful design of model compounds and proper execution of experiments, there appears to be no serious impediment to quantitative development of the transfer model into a useful tool for describing major structural transitions involving proteins.

ACKNOWLEDGMENT

We thank Dr. Jörg Rösgen for discussions concerning the development of the transfer model. We also thank Erica Douglass for her contributions during the Summer Undergraduate Research Program of 2001 at UTMB. Support was provided by NIH Grant GM49760 and Welch Foundation Grant H-1444.

SUPPORTING INFORMATION AVAILABLE

Tables of peptide unit model compound solubilities and densities of the saturated solutions and solvents. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hofmeister, F. (1888) Zur Lehr von der Wirkung der Salze, *Arch. Exp. Pathol. Pharm. 24*, 247−260.
2. Brønsted, J. N. (1920) Studies on Solubility 1: The solubility of salts in salt solutions, *J. Am. Chem. Soc. 42*, 761−786.
3. Noyes, A. A., and Bray, W. C. (1911) The effect of salts on the solubility of other salts. I., *J. Am. Chem. Soc. 33*, 1643−1649.
4. Noyes, A. A., and Bray, W. C. (1911) The effect of salts on the solubility of other salts. II., *J. Am. Chem. Soc. 33*, 1650−1663.
5. Lewis, G. N., and Randall, M. (1923) The Activity of Electrolytes from Freezing Point Data, and Tables of Activity Coefficients, The Activity Coefficient in Mixed Electrolytes, from Solubility Measurements in *Thermodynamics and the Free Energy of Chemical Substances*, Chapter XXVII, p 369, McGraw-Hill Book Company, Inc., New York and London.
6. Cohn, E. J., and Edsall, J. T. (1943) *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions*, Reinhold Publishing Corporation, New York.
7. Edsall, J. T., and Wyman, J. (1958) *Biophysical Chemistry: Thermodynamics, electrostatics, and the biological significance of the properties of matter*, Vol. 1, Academic Press Inc., New York.
8. Kauzmann, W. (1959) Some Factors in the Interpretation of Protein Denaturation, *Adv. Prot. Chem. 14*, 1−63.
9. Nozaki, Y., and Tanford, C. (1963) The Solubility of Amino Acids and Related Compounds in Aqueous Urea Solutions, *J. Biol. Chem. 238*, 4074−4081.
10. Nozaki, Y., and Tanford, C. (1970) The solubility of amino acids, diglycine, and triglycine in aqueous guanidine hydrochloride solutions, *J. Biol. Chem. 245*, 1648−1652.
11. Robinson, D. R., and Jencks, W. P. (1965) The Effect of Compounds of the Urea−Guanidinium Class on the Activity Coefficient of Acetyltetraglycine Ethyl Ester and Related Compounds, *J. Am. Chem. Soc. 87*, 2462−2470.
12. Wimley, W. C., Creamer, T. P., and White, S. H. (1996) Solvation energies of amino acid side chains and backbone in a family of host−guest pentapeptides, *Biochemistry 35*, 5109−24.

13. Radzicka, A., Pedersen, L., and Wolfenden, R. (1988) Influences of solvent water on protein folding: free energies of solvation of cis and trans peptides are nearly identical, *Biochemistry 27*, 4538−4541.
14. Fauchere, J. L., and Pliska, V. (1983) Hydrophobic Parameters-Pi of Amino-Acid Side-Chains from the Partitioning of N-Acetyl-Amino-Acid Amides, *Eur. J. Med. Chem. 18*, 369−375.
15. Tanford, C. (1973) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, John Wiley and Sons, New York.
16. Brandts, J. F. (1969) Conformational Transitions of Proteins in Water and in Aqueous Mixtures in *Structure and Stability of Biological Macromolecules* (Timasheff, S. N., and Fasman, G. D., Eds.) Chapter 3, pp 213−290, Marcel Dekker, Inc., New York.
17. Bolen, D. W. (2001) Protein stabilization by naturally occurring osmolytes, *Methods Mol. Biol. 168*, 17−36.
18. Liu, Y., and Bolen, D. W. (1995) The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes, *Biochemistry 34*, 12884−12891.
19. Qu, Y., Bolen, C. L., and Bolen, D. W. (1998) Osmolyte-driven contraction of a random coil protein, *Proc. Natl. Acad. Sci. U.S.A. 95*, 9268−9273.
20. Wang, A., and Bolen, D. W. (1997) A naturally occurring protective system in urea-rich cells: mechanism of osmolyte protection of proteins against urea denaturation, *Biochemistry 36*, 9101−9108.
21. Bolen, D. W., and Baskakov, I. V. (2001) The osmophobic effect: natural selection of a thermodynamic force in protein folding, *J. Mol. Biol. 310*, 955−963.
22. Nozaki, Y., and Tanford, C. (1965) The solubility of amino acids and related compounds in aqueous ethylene glycol solutions, *J. Biol. Chem. 240*, 3568−3575.
23. Nozaki, Y., and Tanford, C. (1971) The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale, *J. Biol. Chem. 246*, 2211−2217.
24. Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation, *Adv. Protein. Chem. 24*, 1−95.
25. Pace, C. N. (1975) The Stability of Globular Proteins, *CRC Crit. Rev. Biochem. 3*, 1−43.
26. Lapanje, S., Skerjanc, J., Glavnik, S., and Zibret, S. (1978) Thermodynamic studies of the interactions of guanidinium chloride and urea with some oligoglycines and oligoleucines, *J. Chem. Thermodyn. 10*, 425−433.
27. Schrier, M. Y., and Schrier, E. E. (1976) Transfer Free Energies and Average Static Accessibilities for Ribonuclease A in Guanidinium Hydrochloride and Urea Solutions, *Biochemistry 15*, 2607−2612.
28. Tanford, C. (1964) Isothermal Unfolding of Globular Proteins in Aqueous Urea Solutions, *J. Am. Chem. Soc. 86*, 2050−2059.
29. Wetlaufer, D. B., Malik, S. K., Stoller, L., and Coffin, R. L. (1964) Nonpolar Group Participation in the Denaturation of Proteins by Urea and Guanidinium Salts. Model Compound Studies, *J. Am. Chem. Soc. 86*, 508−514.
30. Nandi, P. K., and Robinson, D. R. (1984) Effects of Urea and Guanidine-Hydrochloride on Peptide and Nonpolar Groups, *Biochemistry 23*, 6661−6668.
31. Ben-Naim, A. (1978) Standard Thermodynamics of Transfer. Uses and Misuses, *J. Phys. Chem. 82*, 792−803.
32. Arnett, E., and McKelvey, D. R. (1969) in *Solute−Solvent Interactions* (Coetzee, J. F., and Ritchey, C. D., Eds.) Chapter 6, Marcel Dekker, New York.
33. Hickinbottom, W. J. (1936) *Reactions of Organic Compounds*, Green and Co., London, New York.
34. Russo, A. T., Rösgen, J., and Bolen, D. W. (2003) Osmolyte effects on kinetics of FKBP12 C22A folding coupled with prolyl isomerization, *J. Mol. Biol. 330*, 851−866.
35. Gill, S. J., Hutson, J., Clopton, J. R., and Downing, M. (1961) Solubility of Diketopiperazine in Aqueous Solutions of Urea, *J. Phys. Chem. 65*, 1432−1435.
36. Schönert, H., and Stroth, L. (1981) Thermodynamic Interaction Between Urea and the Peptide Group in Aqueous Solutions at 25 °C, *Biopolymers 20*, 817−831.
37. Smith, E. R. B., and Smith, P. K. (1940) Thermodynamic Properties of Solutions of Amino Acids and Related Substances VI. The Activities of some Peptides in Aqueous Solution at Twenty-Five Degrees, *J. Biol. Chem. 135*, 273−279.
38. Ellerton, D. H., Reinfelds, G., Mulcahy, D. E., and Dunlop, P. J. (1964) Activity, Density and Relative Viscosity Data for Several Amino Acids, Lactamide, and Raffinose in Aqueous Solution at 25°, *J. Phys. Chem. 68*, 398−402.
39. Uedaira, H. (1972) Activity Coefficients for the System Glycyl-glycine-Urea-Water, *Bull. Chem. Soc. Jpn. 45*, 3068−3072.
40. Rafflenbeul, L., Pang, W.-M., Schönert, H., and Haberle, K. (1973) Thermodynamics of Hydrophobic Interaction in Systems Water+Glycine+Urea and Water+Alanine+Urea at 25 °C, *Z. Naturforsch. 28*, 533−554.
41. Cohn, E. J., and Edsall, J. T. (1943) Interactions between Amino Acids, Peptides and Related Substances, Activity Coefficients as a Function of Concentration in *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions*, Chapter 10, p 231, Reinhold Publishing Corporation, New York.
42. Wyman, J. (1933) Dielectric constants: Ethanol−Diethyl Ether and Urea−Water Solutions between 0 and 50°, *J. Am. Chem. Soc. 55*, 4116−4121.
43. Cohn, E. J., and Edsall, J. T. (1943) Interactions between Amino Acids, Peptides and Related Substances, Activity Coefficients as a Function of Concentration in *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions*, Chapter 10, p 232, Reinhold Publishing Corporation, New York.
44. Edsall, J. T., and Wyman, J. (1958) Dielectric Constants and Their Significance, Polar Liquids and Dipolar Ions in *Biophysical Chemistry: Thermodynamics, electrostatics, and the biological significance of the properties of matter*, Chapter 6, p 376, Academic Press Inc., New York.
45. Edsall, J. T., and Wyman, J. (1958) Dielectric Constants and Their Significance, Polar Liquids and Dipolar Ions in *Biophysical Chemistry: Thermodynamics, electrostatics, and the biological significance of the properties of matter*, Chapter 6, p 372, Academic Press Inc., New York.
46. Tanford, C. (1973) The Solubility of Hydrocarbons in Water in *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, Chapter 2, p 4, John Wiley and Sons, New York.
47. Robinson, D. R., and Jencks, W. P. (1963) Effect of Denaturing Agents of the Urea-Guanidinium Class on the Solubility of Acetyltetraglycine Ethyl Ester and Related Compounds, *J. Biol. Chem. 238*, 1558−1560.
48. Zou, Q., Bennion, B. J., Daggett, V., and Murphy, K. P. (2002) The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea, *J. Am. Chem. Soc. 124*, 1192−1202.

BI035908R