Local Chemical Potential Equalization Model for Cosolvent Effects on Biomolecular Equilibria

Paul E. Smith*

Department of Chemistry, 111 Willard Hall, Kansas State University, Manhattan, Kansas 66506-3701

Received: July 8, 2004; In Final Form: August 11, 2004

A simple model for describing the effects of cosolvents on biomolecular equilibria in solution is presented. The model is developed using the Kirkwood−Buff theory of solutions and relates changes in the chemical potential of the cosolvent, due to the presence of the solute, to changes in the cosolvent (and water) concentrations in the vicinity of the biomolecule. The model is then used to determine the dependence of changes in the free energy for protein denaturation on cosolvent concentration. The experimentally observed linearity for urea, and small deviation from linear behavior for guanidinium chloride denaturation, are reproduced. In addition, the model is also predicted to provide a good description of the helix-inducing effects of TFE solutions.

Introduction

Protein denaturation by additives (cosolvents) is a common tool that is used to determine protein stability. Cosolvents are particularly useful because the denaturation process can be studied at physiologically relevant temperature, pressure, and pH. The free energy for denaturation typically becomes more favorable in the presence of increasing concentrations of denaturing cosolvents such as urea and guanidinium chloride (GdmCl). However, whereas the free-energy change can be determined quite accurately in the region of the transition concentration, the corresponding free-energy change in the absence of a cosolvent requires an extrapolation of the free-energy difference against the cosolvent concentration curve to zero cosolvent concentration.[1] Because the exact dependence of the free-energy change on the cosolvent concentration is unknown, any extrapolation procedure can lead to significant errors in estimating protein stability and to differences between the estimates for the same protein using different cosolvents.

The extrapolation procedure could be improved by accurate models of protein denaturation that explain the experimental observations. The most commonly used approach is the linear extrapolation method, or *m*-value approach, which assumes a linear dependence of the free-energy change on cosolvent concentration.[2] This approach, although purely empirical, has been used extensively to characterize protein stability successfully. However, significant deviations from linearity have been observed, especially for denaturation by GdmCl.[3,4] More importantly, the model provides no insights into the denaturation mechanism.

Cosolvent binding and exchange models have been used to represent experimental data for denaturation by urea and GdmCl.[5−8] These models generally assume a set of identical and independent binding sites characterized by a single binding constant. Denaturation is a result of the presence of more binding sites in the denatured form of the protein. Both models predict near linear denaturation curves for urea, with deviations from linearity for GdmCl.[4,7] However, the models generate estimates of the number of bound cosolvent molecules that are sensitive to the exact analysis.[9] Furthermore, they provide representations of the data and not true thermodynamic properties, as emphasized by Timasheff.[10] This is a consequence of the weak binding of urea and GdmCl to proteins, which also means that it is difficult to locate or define the proposed binding sites.

The local-bulk domain model has also been used to explain denaturation free-energy curves.[11,12] In this model, the preferential interaction of the cosolvent with the protein is determined by assuming that the cosolvent concentration close to the protein surface is higher than that in the bulk cosolvent mixture. This increase is then characterized by a partition coefficient, and denaturation is favored by the larger surface area that is exposed in the denatured form. However, the model requires several parameters and appears to display a deviation from linearity for urea denaturation that is larger than, and in the opposite direction from, that for GdmCl denaturation.[12]

In this study, a general model for the effect of any cosolvent on the equilibrium of any infinitely dilute solute is presented. The model determines the change in cosolvent concentration around the solute that is required to counteract the perturbing effect of the solute on the cosolvent chemical potential. This is achieved by reference to the Kirkwood−Buff (KB) theory of solutions.[13,14] In principle, the model can be applied to any conformational transition induced by any cosolvent, although we will focus on protein denaturation as an example. Finally, the model is then compared with the *m* value, binding, and local-bulk domain models.

Background

Throughout this article, it is assumed that the equilibrium of interest involves just two major states of the biomolecule and that the total biomolecule concentration is so low that we can treat it as being infinitely dilute. This is not a limitation of KB theory, but it does make the final equations more tractable. Also, the common ion effect has been neglected, which is usually acceptable for conformational transitions of proteins. The notation of solute (s), cosolvent (c), and water (w) has been adopted for the application of KB theory because it is more descriptive than, albeit not as general as, the common notation of primary solvent (1), biomolecule (2), and additive (3).

* Tel: 785-532-5109. Fax: 785-532-6666. E-mail: pesmith@ksu.edu.

10.1021/jp046973t CCC: $27.50
Published on Web 09/21/2004

Kirkwood−Buff Theory for Binary Solutions. Because our solute is at infinite dilution, we will often refer to the properties of the cosolvent mixture alone. The KB theory can be used to describe the thermodynamics of solution mixtures and relates the radial distribution functions (rdfs) between the different solution species to the derivatives of the chemical potentials of the components and their partial molar volumes.[13,14] This is achieved via KB integrals that are defined as

$$G_{ij} = 4\pi\int_0^\infty [g_{ij}^{\mu VT}(r) - 1]r^2 \mathrm{d}r \tag{1}$$

where $g_{ij}^{\mu VT}(r)$ is the rdf between species i and j as a function of the center of mass distance r in the grand canonical (μVT) ensemble. The corresponding coordination numbers ($N_{ij} = \rho_j G_{ij}$) describe the excess number of j molecules around a central i molecule. For a binary mixture of cosolvent and water at a constant pressure (p) and temperature (T), the derivative of the cosolvent activity with respect to the cosolvent concentration is given by[14]

$$a_{cc} = \beta\left(\frac{\partial \mu_c}{\partial \ln \rho_c}\right)_{T,p} = \left(\frac{\partial \ln a_c}{\partial \ln \rho_c}\right)_{T,p} = \frac{1}{1 + \rho_c(G_{cc} - G_{cw})} \tag{2}$$

where a_c and D_c are the activity and number density (molar concentration) of the cosolvent, respectively, and $\beta = (RT)^{-1}$ where R is the gas constant. The above expression can be rewritten using the KB result for the partial molar volume (pmv) of the cosolvent, which is given by[15]

$$\bar{V}_c = RT\kappa_T - \rho_c \bar{V}_c G_{cc} - \rho_w \bar{V}_w G_{cw} \tag{3}$$

where $\overline{V_i}$ is the pmv of i and κ_T is the isothermal compressibility of the solution. Using eq 3 to eliminate G_{cw} from eq 2 one can show that[16]

$$a_{cc} = \frac{\phi_w}{1 + \rho_c(G_{cc} - RT\kappa_T)} \approx \frac{\phi_w}{1 + \rho_c G_{cc}} = \frac{\phi_w}{1 + N_{cc}} \tag{4}$$

where the approximation is usually valid for liquids (less than 1% error) due to their low compressibilities.

The application of the KB theory to electrolyte solutions is complicated by the correlations between the KB integrals that are a consequence of electroneutrality.[17−19] To avoid this problem, we treated the anions and cations of salts as indistinguishable particles to apply the KB equations for a binary solution (water and cosolvent), rather than those for a ternary solution (water, cations, and anions). This approach has been discussed and used before.[15,20,21] Hence, we will distinguish between the usual molar salt concentration (C_c) and the concentration of indistinguishable ions ($\rho_c = n_\pm C_c$), where $n_\pm$ is the number of ions produced on dissociation of the salt.

Kirkwood−Buff Theory for Protein Denaturation. Applying the KB theory to protein denaturation, one finds that the derivative of the standard free-energy change for denaturation ($\Delta G°$) with respect to cosolvent concentration is given by[21,22]

$$-\beta\left(\frac{\partial \Delta G°}{\partial \rho_c}\right)_{T,p} = \left(\frac{\partial \ln K}{\partial \rho_c}\right)_{T,p} = \frac{\Delta\nu_{sc} a_{cc}}{\rho_c} \tag{5}$$

for an infinitely dilute protein undergoing a transition from a native (N) state to a denatured (D) state with an equilibrium constant of $K = \rho_D/\rho_N$. The above expression is equivalent to the general result

$$\left(\frac{\partial \ln K}{\partial \ln a_c}\right)_{T,p} = \Delta\nu_{sc} = \nu_{Dc} - \nu_{Nc} \tag{6}$$

which has been obtained by a number of approaches.[10,14,23−27] Here, $\Delta\nu_{sc}$ is the change in preferential interaction of the cosolvent between the native (ν_{Nc}) and denatured states (ν_{Dc}) of the protein. A higher preferential interaction of the cosolvent with the denatured state drives the equilibrium in favor of the denatured state. The preferential interaction of the cosolvent with the protein solute can be expressed as[10,14,23]

$$\nu_{sc} = \rho_c(G_{sc} - G_{sw}) = N_{sc} - \frac{\rho_c}{\rho_w} N_{sw} \tag{7}$$

where N_{sc} and N_{sw} are the excess number of cosolvent and water molecules associated with the protein, respectively. The values of N_{ij} correspond to the numbers of cosolvent and water molecules in the region that are perturbed by the presence of the protein. This includes molecules that are directly bound to the protein together with any changes in the molecular distribution over several solvation shells. Finally, we note that the change in cosolvent association with the protein (ΔN_{sc}) can be obtained from the following exact relationship[21]

$$\Delta N_{sc} = N_{Dc} - N_{Nc} = \phi_w \Delta\nu_{sc} - \rho_c \Delta\overline{V_s^\infty} \tag{8}$$

for which $\Delta\overline{V_s^\infty}$ is the change in pmv of the infinitely dilute protein upon denaturation.

Local Chemical-Potential Equalization (LCPE) Model

A model for the redistribution of the cosolvent and water molecules around a protein is developed here in an effort to describe the dependence of the denaturation equilibrium on cosolvent concentration. Let us consider a system of cosolvent and water molecules under NpT conditions at the desired cosolvent concentration. The LCPE model then involves two steps. In the first step, sufficient cosolvent and water molecules are removed to generate a hole large enough to insert the protein in the native or denatured form. At this stage, the protein does not interact with the cosolvent or water molecules. Hence, the free energy for formation of the hole is given by

$$G_H = -(n_c\mu_c + n_w\mu_w) + P_{V_s} \tag{9}$$

in which n_c and n_w are the number of cosolvent and water molecules that have been removed, μ_i is the chemical potential of i, and P_{V_s} is a factor that accounts for the probability that all of the removed cosolvent and water molecules were in the same volume (V_s). It will be assumed that the volumes of the native and denatured protein are essentially the same.[19,21] In this case, the exact result from eq 9 is not of interest because we will be considering the differences in hole formation for the denatured and native forms (ΔG_H), which will, therefore, be zero. We note that this process is not the same as cavity formation, as described by scaled particle theory,[28] primarily because cavity formation includes the contribution from the redistribution of cosolvent and solvent molecules.

In the second step, the cosolvent and water molecules are allowed to redistribute around the protein in response to the perturbing effect the protein has on the chemical potentials of the cosolvent and water molecules close to the protein surface. At this stage, one can consider the immediate region surrounding the protein, defined by the volume V^1, to be a μVT region surrounded by an infinite NpT bath as shown in Figure 1.

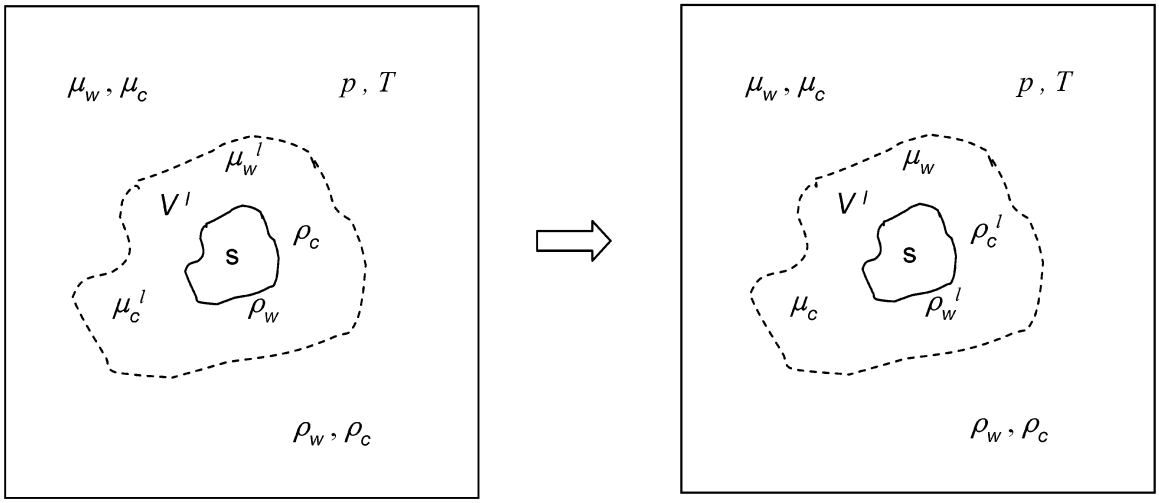

Figure 1. Step two for the LCPE model of protein denaturation. The protein solute perturbs the chemical potential of the cosolvent and water molecules in the local region defined by V^l. Molecules then move to and from the *NpT* bath to equalize the chemical potential between the local region and the particle bath.

Cosolvent and water molecules will move in or out of the μVT region in an effort to equalize the chemical potentials between the molecules that are close to the protein surface and those that are in the *NpT* bath. The present model attempts to derive an equation for this redistribution.

The change in chemical potential of the cosolvent that is required to compensate the change due to the presence of the protein is $\Delta\mu_c = \mu_c - \mu_c^l$ and is positive if a favorable interaction exists between the protein and cosolvent. It will be assumed that $\Delta\mu_c$ is the same for the native and denatured states and is independent of the cosolvent concentration. The validity of this assumption and the exact meaning of $\Delta\mu_c$ will be discussed later. The change in cosolvent concentration that is required to reestablish the equilibrium can be approximated by a Taylor series expansion for the μVT region

$$\rho_c^l = \rho_c + \Delta\mu_c\left(\frac{\partial\rho_c}{\partial\mu_c}\right)_{T,\mu_w} + \frac{\Delta\mu_c^2}{2}\left(\frac{\partial^2\rho_c}{\partial\mu_c^2}\right)_{T,\mu_w} + O(\Delta\mu_c^3) \quad (10)$$

Derivatives of the cosolvent concentration with respect to the water chemical potential are not included in the above expansion because μ_c and μ_w are dependent variables, being related by the Gibbs−Duhem equation ($n_c\, d\mu_c + n_w\, d\mu_w = 0$) for the constant pressure and temperature bath.[14,27] The above derivatives can be obtained from the KB theory and are given by[14]

$$\left(\frac{\partial\rho_c}{\partial\mu_c}\right)_{T,\mu_w} = \beta\rho_c(1 + N_{cc}) = \beta Z \quad (11)$$

and, therefore,

$$\left(\frac{\partial^2\rho_c}{\partial\mu_c^2}\right)_{T,\mu_w} = \beta\left(\frac{\partial\rho_c}{\partial\mu_c}\right)_{T,\mu_w}\left(\frac{\partial Z}{\partial\rho_c}\right)_{T,\mu_w} = \beta^2 Z\left(\frac{\partial Z}{\partial\rho_c}\right)_{T,\mu_w} \quad (12)$$

Hence, the change in the cosolvent concentration in the local region is then

$$\rho_c^l - \rho_c = \beta\Delta\mu_c Z\left[1 + \frac{\beta\Delta\mu_c}{2}\left(\frac{\partial Z}{\partial\rho_c}\right)_{T,\mu_w}\right] \quad (13)$$

To proceed further, we have to express the change in concentration in terms of a KB integral between the solute and the cosolvent. The concentration change can be written as

$$\rho_c^l - \rho_c = \frac{N^l - N^o}{V^l} \quad (14)$$

where N^o and N^l are the number of cosolvent molecules that are found in the local volume (V^l) around the protein before and after the redistribution, respectively. The number of cosolvent molecules in the local domain after redistribution can be written in terms of the corresponding rdf in the μVT ensemble

$$N^l = 4\pi\rho_c\int_{R_s}^{R_l} g_{sc}^{\mu VT}(r)r^2\,dr = 4\pi\rho_c\int_0^{R_l} g_{sc}^{\mu VT}(r)r^2\,dr \quad (15)$$

where we have assumed a spherical protein of radius (R_s) in a local spherical volume of radius (R_l) for simplicity of notation only. The corresponding number of cosolvent molecules before the redistribution is given by

$$N^o = 4\pi\rho_c\int_{R_s}^{R_l} r^2\,dr = 4\pi\rho_c\int_0^{R_l} r^2\,dr - 4\pi\rho_c\int_0^{R_s} r^2\,dr \quad (16)$$

Consequently, the cosolvent density change is given by

$$(\rho_c^l - \rho_c)V^l = 4\pi\rho_c\int_0^{\infty}[g_{sc}^{\mu VT}(r) - 1]r^2\,dr + \rho_c V_s^* = N_{sc} + \rho_c V_s^* \quad (17)$$

where N_{sc} is the required KB excess coordination number between the cosolvent and protein solute and V_s^* is the protein volume ($V_s^* = \overline{V_s^\infty} - RT\kappa_T$).[14] Combining eqs 13 and 17 gives

$$N_{sc} = V^l\beta\Delta\mu_c Z\left[1 + \frac{\beta\Delta\mu_c}{2}\left(\frac{\partial Z}{\partial\rho_c}\right)_{T,\mu_w}\right] - \rho_c V_s^* \quad (18)$$

and expresses the new cosolvent distribution in terms of the change in the cosolvent chemical potential and a KB integral between cosolvent molecules (G_{cc}) in the absence of the protein.

The partial derivative on the right-hand side of eq 18 can be transformed to the *NpT* ensemble using the standard relationship between partial derivatives

$$\left(\frac{\partial Z}{\partial\rho_c}\right)_{T,\mu_w} = \left(\frac{\partial Z}{\partial\rho_c}\right)_{T,p} + \left(\frac{\partial Z}{\partial p}\right)_{T,\rho_c}\left(\frac{\partial p}{\partial\rho_c}\right)_{T,\mu_w} \quad (19)$$

TABLE 1: Fitting Parameters for N_{cc} Using Equation 25[a]

cosolvent	a_0 ($M^{-1/2}$)	a_1 (M^{-1})	a_2 (M^{-2})	a_3 (M^{-3})	a_4 (M^{-4})	rmsd
urea		−0.005	−0.0112	0.00061		0.002
GdmCl	0.4276	−0.233	−0.0104	0.00128		0.003
TFE		0.1780	−0.0442	0.02320	−0.00188	0.004
NaCl	0.3552	−0.4101	0.0443	−0.00318		0.004

[a] All data correspond to aqueous solutions at 298 K and 1 atm. The experimental data was generated as described previously.[15]

The second term on the right-hand side of eq 19 will be negligible because the compressibility of a liquid is small; therefore, $(\partial G_{cc}/\partial p)_{T,\rho_c} \approx 0$. Hence, the change in cosolvent association upon denaturation is given by

$$\Delta N_{sc} = \Delta V^{l}\beta\Delta\mu_c Z\left[1 + \frac{\beta\Delta\mu_c}{2}\left(\frac{\partial Z}{\partial \rho_c}\right)_{T,p}\right] - \rho_c\Delta\overline{V_s^{\infty}} \quad (20)$$

Combining eqs 5 and 8 yields

$$\left(\frac{\partial \ln K}{\partial \rho_c}\right)_{T,p} = \frac{(\Delta N_{sc} + \rho_c\Delta\overline{V_s^{\infty}})a_{cc}}{\rho_c\phi_w} \quad (21)$$

which leads, with use of the minor approximation introduced in eq 4, to the result

$$\left(\frac{\partial \ln K}{\partial \rho_c}\right)_{T,p} = \frac{1}{n_{\pm}}\left(\frac{\partial \ln K}{\partial C_c}\right)_{T,p} = \Delta V^{l}\beta\Delta\mu_c\left[1 + \frac{\beta\Delta\mu_c}{2}\left(\frac{\partial Z}{\partial \rho_c}\right)_{T,p}\right] \quad (22)$$

Integration of the above equation generates the final relationship describing the dependence of the denaturation equilibrium on the cosolvent concentration given by the LCPE model

$$\beta\Delta\Delta G = \beta[\Delta G°(C_c) - \Delta G°(0)] = -n_{\pm}C_cA[1 + BN_{cc}] \quad (23)$$

with

$$A = \Delta V^{l}\beta\Delta\mu_c\left(1 + \frac{1}{2}\beta\Delta\mu_c\right) \qquad B = \frac{\frac{1}{2}\beta\Delta\mu_c}{1 + \frac{1}{2}\beta\Delta\mu_c} \quad (24)$$

Hence, the LCPE model describes the change in denaturation free energy in terms of three unknowns, $\Delta G°(0)$, ΔV^{l}, and $\Delta\mu_c$, together with a concentration dependent property of the cosolvent solution in the absence of protein (N_{cc}).

Results and Discussion

Free Energy as a Function of Cosolvent Concentration. The shape of the denaturation plot ($\beta\Delta\Delta G$ against C_c) predicted by the LCPE model depends on the magnitude of B and the variation of N_{cc} with cosolvent concentration, which is a property of the cosolvent and water mixture alone for an infinitely dilute solute. The experimental values of N_{cc} at 298 K and 1 atm are known for many aqueous solutions such as urea, GdmCl, 2,2,2-trifluoroethanol (TFE), and NaCl.[15] A fit to the data is provided by the following expression

$$N_{cc} = \rho_c G_{cc} = a_0 C_c^{1/2} + \sum_{i=1}^{4} a_i C_c^{i} \quad (25)$$

and the results are presented in Table 1 and Figure 2. The $C_c^{1/2}$ term accounts for the Debye−Hueckel limiting law for salts.[15,18] From Figure 2, it can be seen that N_{cc} for urea and NaCl is negative at all but low (<1 M) cosolvent concentrations and is negative for GdmCl solutions above 2.9 M, whereas N_{cc} is large and positive for TFE. The consequences of this behavior are presented in Figure 3 where we display calculated free-energy curves for urea, GdmCl, and NaCl denaturation and helix formation by TFE, using estimated values ($\Delta V^{l} = 2\ M^{-1}$ and $\beta\Delta\mu_c = 0.5$; therefore, $A = 1.25\ M^{-1}$ and $B = 0.2$) for the two unknown parameters. These parameters are not necessarily representative of the experimental data on denaturation and helix formation but are used merely to emphasize the deviation from linearity that can be expected for the different cosolvents. The slope of the TFE curve becomes increasingly more negative with TFE concentration and is clearly nonlinear. The urea, GdmCl, and NaCl curves display larger slopes with a curvature leading to a decrease in slope magnitude at higher cosolvent concentrations.

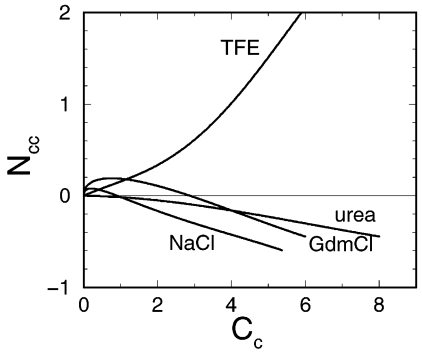

Figure 2. Experimentally observed variation of N_{cc} with molar cosolvent concentration obtained from eq 25 and the parameters in Table 1. For comparison, 6 M TFE corresponds to a TFE mole fraction of 0.16 or 43% v/v.

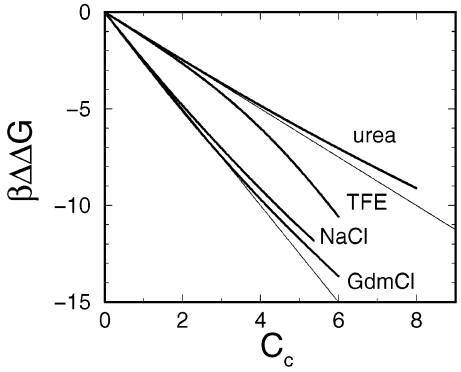

Figure 3. Free-energy difference curves obtained from Equations 23 and 25 for values of $\Delta V^{l} = 2\ M^{-1}$ and $\beta\Delta\mu_c = 0.5$. The parameters were chosen to illustrate the potential deviations from linearity and are not necessarily representative of real data on denaturation or helix induction. The thin lines are provided to emphasize the deviations from linearity and correspond to just the first term in the brackets from eq 23.

General Features of the Model. The LCPE model describes changes in the free-energy difference for denaturation in terms of two unknowns. The change in the chemical potential of the cosolvent determines whether the corresponding free-energy change for denaturation will become more favorable ($\Delta\mu_c$ is positive), more unfavorable, leading to increased protein stability ($\Delta\mu_c$ is negative), or indifferent ($\Delta\mu_c$ is zero). The value of $\Delta\mu_c$ represents the average change in the cosolvent chemical potential resulting from interactions with many different types of amino-acid side chains and is assumed to be the same for both the

native and denatured states. The average chemical-potential change includes a contribution from all cosolvent molecules in the local volume surrounding the protein. Cosolvent molecules close to the protein surface (first solvation shell) will be the most perturbed, but these will be balanced by an increasing number of molecules at larger distances which are only weakly perturbed; therefore, the average change in the cosolvent chemical potential will probably be small. It is almost certain that $\Delta\mu_c$ varies somewhat with protein composition and cosolvent concentration. However, we will assume that these variations are small for the majority of proteins. To some degree, this is inferred by the weak binding of many cosolvents, which suggests a series of small interactions with no major binding sites.

The second parameter represents the change in local volume between the native and denatured state. Because the denatured state has a significantly increased surface area over the native state, and because it has been assumed that $\Delta\mu_c$ is the same for both states, the LCPE model suggests that the main factor leading to protein denaturation is the larger local volume surrounding the denatured state, which will be different for different proteins. The size of the local volume surrounding both the native and denatured states is unknown. However, it is reasonable to assume that the distance from the protein surface over which the cosolvent distribution is perturbed ($R_P = R_l - R_s$) will be the same for both states. If this distance were known, then one could determine reasonable estimates of the accessible surface area of the denatured state from the denaturation curves of several proteins.

Several features of experimental denaturation curves have been observed and are reproduced by the LCPE model. First, the application of the linear extrapolation method to urea and GdmCl denaturation data indicates that GdmCl *m* values are typically twice the urea *m* values.[29] This is apparent from eq 23 and Figure 3, where for a 1:1 salt $n_\pm = 2$, the number of ions produced upon dissolving the salt in water. This term arises directly from our treatment of salts as a collection of indistinguishable ions to be able to use the KB theory for ionic solutions. Physically, this can be rationalized by assuming urea molecules and GdmCl cations interact with the protein in a similar fashion (the combination of $\Delta\mu_c$ and ΔV^l values are comparable), but because GdmCl is a salt, the electroneutrality dictates that an equivalent number of chloride anions must also be associated with the protein. Hence, the factor of 2 in the cosolvent preferential interaction and the corresponding *m* values for 1:1 salts.

The second observation concerning urea and GdmCl denaturation is that the *m* values are approximately proportional to the predicted changes in surface area on unfolding,[29] although the real changes in surface area on unfolding are unknown. This can be easily explained by the LCPE model if one assumes that the distance from the protein surface over which the cosolvent and water distributions are perturbed (R_P) is small compared to the effective radius of the protein. In this case, one can write $\Delta V^l = \Delta ASA R_P$, where ΔASA is the change in accessible solvent surface area on unfolding. Unfortunately, the value of R_P for urea and GdmCl denaturation is currently unknown, although it could be determined from accurate simulation data.[20,21,30,31] We note that the LCPE model predicts only a simple dependence on surface area changes if the value of R_P is small (one solvation shell).

Comparison with the *m*-Value Approach. The most common method used to quantify the effects of cosolvents on protein denaturation is the linear extrapolation method, or *m*-value approach.[3] In this case, one has

$$\beta\Delta\Delta G = -mC_c \tag{26}$$

where *m* is a constant. The *m*-value approach works well for urea denaturation but is less satisfactory for GdmCl denaturation.[4] This behavior is also predicted by the LCPE model where, for a given value of $\Delta\mu_c$, the GdmCl denaturation curves display more curvature than the urea curves (Figure 3). Linear behavior is predicted by the LCPE model when the value of $\Delta\mu_c$ is small or if N_{cc} is independent of cosolvent concentration. When the perturbation to the cosolvent chemical potential is small, the Taylor expansion can be truncated after the first derivative, leading to

$$\left(\frac{\partial \ln K}{\partial C_c}\right)_{T,p} = n_\pm \Delta V^l \beta \Delta\mu_c = m \tag{27}$$

which indicates a linear dependence of the denaturation equilibrium on the cosolvent concentration. Obviously in this situation, the LCPE model describes any linear behavior with two unknowns, compared to one for the *m*-value approach and is therefore less appealing. However, an advantage of the LCPE model is that the two unknowns convey a physical description of the denaturation process whereas the *m* value does not.

Comparison with Binding Models. Cosolvent effects have also been quantified in terms of binding models. In this case, one has[5,6]

$$\beta\Delta\Delta G = -\Delta n \ln[1 + K_b C_c] \approx -\Delta n K_b C_c\left(1 - \frac{1}{2}K_b C_c\right) \tag{28}$$

where K_b is the binding constant of the cosolvent to a series of identical independent sites on the protein surface and Δn is the increase in the number of sites on going from the native to denatured states. The denaturation curve described by the binding model is initially linear in cosolvent concentration, but then the slope of the curve decreases in magnitude as the second term in the expansion becomes more important. Generally, the binding model fits the experimental data on urea and GdmCl denaturation very well and suggests greater curvature for GdmCl denaturation (K_b is larger).[4,6,9] Equation 23 indicates that the shape of the denaturation curve for the LCPE model is determined by the sign of N_{cc} because *A* and *B* are both positive constants. From Figure 2, one can see that N_{cc} is always small and negative for urea, whereas for GdmCl, it is initially positive but then is negative beyond 2.9 M GdmCl. Therefore, at concentrations where N_{cc} contributes to the change in denaturation free energy, the contribution is negative for both urea and GdmCl solutions. The effect is larger for GdmCl primarily because of the indistinguishable ion approximation ($\rho_c = n_\pm C_c$). Hence, the LCPE model reproduces the same curvature and properties that were predicted by the binding model.

Although the denaturation curves for the binding and LCPE models display the same trends with increasing cosolvent concentration, they arise from physically different origins. The simple binding approach essentially corresponds to a Langmuir binding isotherm where the binding is decreased as the number of unoccupied sites decreases. In contrast, the concentration dependence for the LCPE model arises from the properties of the cosolvent and water mixture (N_{cc}).

Helix Formation in TFE Solutions. The thermodynamics of helix formation in TFE solutions cannot be accurately explained by the simple *m*-value approach.[32] However, an alternative exchange model has been shown to fit the data adequately. In this case, the difference in the free energy for

helix formation is given by[32]

$$\beta\Delta\Delta G = -m_e\frac{\rho_c}{\rho_w} \approx -m_e\overline{V_w}\rho_c(1+\rho_c\overline{V_c}) \qquad (29)$$

where m_e is a constant, and the above approximation was generated by substituting the value of $\rho_w = (1-\phi_c)/\overline{V_w}$ and then expanding the denominator. The water pmv is reasonably independent of TFE concentration. It can be seen that the dependence of the free-energy change on the cosolvent concentration is different from that of urea and GdmCl denaturation for the quadratic term. This also suggests why one cannot fit the TFE data using a simple binding model (eq 28) where the quadratic term is negative. In contrast, the value of N_{cc} for aqueous TFE solutions is always positive and, therefore, mimics the exchange model (eq 29), as shown in Figure 3. Hence, the correct dependence of the free-energy change on TFE concentration is predicted by the LCPE model.

Comparison with the Local-Bulk Domain Model. The local-bulk domain model has been used to help explain the effects of denaturants in terms of the difference in the cosolvent concentration at the protein surface compared to that in the bulk solution. The model assumes that the difference between the local and bulk cosolvent concentrations is characterized by a partition coefficient $K_{s,c}$, which is the same for all proteins and is independent of the cosolvent concentration. The result for the denaturation equilibrium derivative is given by[12]

$$\left(\frac{\partial \ln K}{\partial C_c}\right)_{T,p} = \frac{n_\pm(K_{s,c}-1)b_w^0\Delta ASA a_{cc}}{m_w\phi_w(1+n_\pm K_{s,c}S_{w,c}m_c m_w^{-1})} \qquad (30)$$

where $K_{s,c}$, b_w^0, ΔASA, and $S_{w,c}$ are unknown positive constants with some physical significance and m_i is the molality of the i component. To compare with the LCPE model, we will use the KB result for a_{cc} (eq 4) in eq 30 and then expand the denominator, keeping terms up to ρ_c only. The result is

$$\left(\frac{\partial \ln K}{\partial C_c}\right)_{T,p} \approx \frac{n_\pm(K_{s,c}-1)b_w^0\Delta ASA}{m_w}\left[1-\rho_c\left(G_{cc}+\frac{K_{s,c}S_{w,c}}{m_w}\right)\right] \qquad (31)$$

The curvature in the denaturation plot will depend on the relative contributions of G_{cc} and $K_{s,c}\,S_{w,c}\,m_w^{-1}$, the latter of which is always positive. Because G_{cc} is negative for urea and GdmCl at concentrations where the ρ_c term becomes important, the different signs of the two terms help to explain the different curvatures that were predicted for urea and GdmCl denaturation by the local-bulk domain model.[12] In comparison, the corresponding result for the LCPE model can be written as

$$\left(\frac{\partial \ln K}{\partial C_c}\right)_{T,p} = n_\pm A\left[1+B\rho_c\left(G_{cc}+\frac{\partial N_{cc}}{\partial \rho_c}\right)\right] = m^{\mathrm{eff}} \qquad (32)$$

where both terms in the brackets on the right-hand side of eq 32 are always negative for urea and negative for GdmCl solutions above 2.9 M. Hence, the LCPE model predicts the same deviation from linearity for both urea and GdmCl at medium to high cosolvent concentrations.

The change in equilibrium with the cosolvent concentration for the exchange model that is used to describe helix formation in TFE solutions is given by

$$\left(\frac{\partial \ln K}{\partial C_c}\right)_{T,p} = m_e\frac{\overline{V_w}}{\phi_w^2} \approx m_e\overline{V_w}(1+2\rho_c\overline{V_c}) \qquad (33)$$

A comparison with eq 31, where the last two terms are positive for the local-bulk domain model, suggests that it is unlikely that the model can be applied with success to describe TFE effects. In contrast, all terms on the right-hand side of eq 32 are positive for the LCPE model and will, therefore, qualitatively reproduce the experimentally observed dependence of the equilibrium on TFE concentration.

Estimates of ΔV^l and $\Delta\mu_c$ for Urea and GdmCl Denaturation. Although the previous sections suggest the LCPE model displays the correct deviation from linearity for the denaturation plot of GdmCl and the helix inducing behavior of TFE, the model will not provide an acceptable fit to the experimental data if the required values of ΔV^l and $\Delta\mu_c$ are physically unreasonable. Neither are known experimentally; however, one can estimate their values by making some simple approximations. To estimate the value of ΔV^l, we will assume a native protein that is represented by a prolate ellipsoid with semi axes of $a = 2.25$ nm and $b = c = 1.5$ nm. This corresponds to a protein with roughly 130 residues. If denaturation of the protein also results in a prolate ellipsoid of the same volume but with semi axes of $a = 4.5$ nm and $b = c = 1.06$ nm, then one can determine the change in local volume upon denaturation. The above change in shape corresponds to a change in the radius of gyration from 1.4 to 2.1 nm and is comparable to the experimentally observed changes for real proteins of this size.[33] Previous simulation studies have suggested that the value of R_P is approximately 1.5 nm for urea and 1.0 nm for GdmCl.[16,20,34,35] Using these values, the change in local volume is determined to be 14.1 M^{-1} for urea and 7.6 M^{-1} for GdmCl denaturation.

The change in the cosolvent chemical potential is more difficult to determine. However, it is generally expected that the local concentration (C_c^l) of both urea and GdmCl is higher in the local volume than the bulk solution. Hence, a simple approximation to the chemical potential change is given by

$$\beta\Delta\mu_c \approx \ln\left[\frac{a_c(C_c^l)}{a_c(C_c)}\right] \qquad (34)$$

where $a_c(C_c)$ is the cosolvent activity in the solution mixture at a cosolvent concentration of C_c. If we assume that the local-bulk domain model provides a reasonable estimate of the ratio of local to bulk cosolvent concentrations, then one can write $C_c^l = 1.12C_c$ for urea and $C_c^l = 1.48C_c$ for GdmCl.[12] Using recently suggested experimental fits to the experimental activity data,[36] eq 34 provides the results presented in Figure 4. The predicted values of $\beta\Delta\mu_c$ for urea are independent of urea concentration, as assumed during the derivation of the LCPE model. The change for GdmCl is larger and displays some variation with the cosolvent concentration, especially at low concentrations.

Using the average values of $\beta\Delta\mu_c = 0.1$ for urea and 0.25 for GdmCl, together with the estimated changes in the local volume, one finds m values for this model system to be ~1.4 M^{-1} for urea and ~4.0 M^{-1} for GdmCl denaturation. This is to be compared to typical corresponding experimental values of ~2.4 M^{-1} and ~4.8 M^{-1}, respectively.[29] Hence, the LCPE model predicts m values that are comparable to the experimentally observed data using physically reasonable estimates for C_c^l, $\Delta\mu_c$ and ΔV^l. This suggests that the LCPE model can be

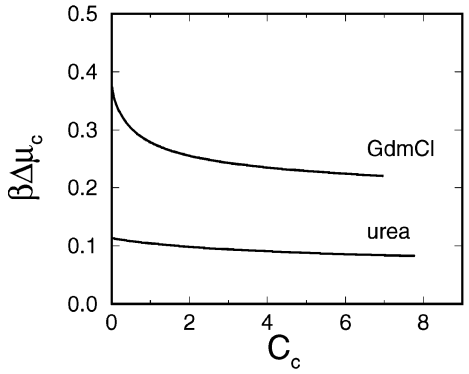

Figure 4. Change in cosolvent chemical potential as a function of molar cosolvent concentration as predicted by the experimental activities (298 K and 1 atm) and the approximation used in Equation 34. The local cosolvent concentration was taken to be $C_c^l = 1.12C_c$ for urea and $C_c^l = 1.48C_c$ for GdmCl.

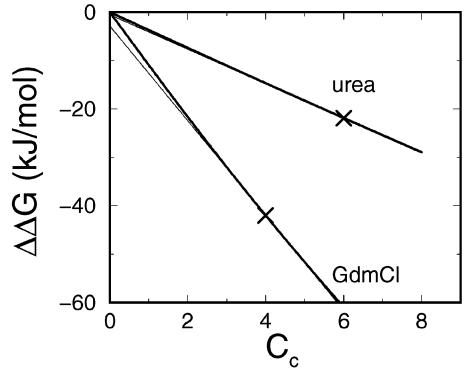

Figure 5. Difference in free energy of unfolding for a model protein at 298 K (see text). Urea denaturation is represented by values of $A = 1.48\ M^{-1}$ and $B = 0.05$. If the midpoint of the transition is at 6 M urea, then $m^{eff} = 1.4\ M^{-1}$ and linear extrapolation gives an error for $\Delta\Delta G$ at zero concentration of −0.5 kJ/mol. GdmCl denaturation is represented by values of $A = 2.14\ M^{-1}$ and $B = 0.11$. If the midpoint of the transition is at 4 M GdmCl, then $m^{eff} = 4.0\ M^{-1}$ and linear extrapolation gives an error for $\Delta\Delta G$ at zero concentration of −2.8 kJ/mol.

used with confidence to fit the raw experimental data for protein denaturation. The near linearity of the urea denaturation curve is related to the small value of $\beta\Delta\mu_c$ and the correspondingly small contribution of the second-order term in eq 10. The contribution of the second-order term is nonnegligible for GdmCl denaturation and leads to a noticeable departure from linear behavior.

Errors Introduced by Linear Extrapolation. The LCPE model can be used to provide an estimate of the possible errors involved in using the linear extrapolation procedure. Using the data presented in the previous section results in the denaturation plots shown in Figure 5. Assuming that the midpoint of the denaturation transition ($\Delta G°(C_c^{1/2}) = 0$) corresponds to 6 M urea and 4 M GdmCl, and using the slopes given by the LCPE model (eq 25 and 32), the linear extrapolation procedure produces intercepts ($\Delta\Delta G(0)$) of −0.5 and −2.8 kJ/mol for urea and GdmCl denaturation, respectively. In both cases, the stability of the native protein is underestimated. The error is negligible for urea, whereas for GdmCl, the error is predicted to be more significant. Obviously, the exact deviation from linearity will vary with the values of the unknown parameters. In addition, the deviation from linearity will also become smaller as the midpoint transition concentration decreases. Nevertheless, the above results still provide an indication of the behavior predicted by the LCPE model, which appear to agree with the experimental data.[4] Finally, we note that in the above example one could have set the values of the midpoint transition concentrations such that the values of $\Delta\Delta G$ were the same for both urea and GdmCl. This will only be true, however, if the denatured state is the same for both urea and GdmCl denaturation.

Conclusions

A general model for the effect of any cosolvent on any infinitely dilute equilibrium process has been presented. The model involves three unknown parameters ($\Delta\mu_c$, V^l, and $\Delta G°(0)$) and correctly describes the experimental deviations from linearity for GdmCl and TFE cosolvents. The LCPE model uses less parameters than the local-bulk domain model and provides more insights into the denaturation process than the *m*-value approach. Furthermore, it can be applied to a wide range of cosolvents and cosolvent effects. The determination of the three unknowns can be achieved by using the raw experimental denaturation data for a range of proteins. This will aid in the extrapolation process that is required to determine the stability of proteins in pure water. Alternatively, one could use computer simulations in an effort to determine V^l (or R_P), and possibly $\Delta\mu_c$, as a function of distance from a protein surface. Investigations in this direction are currently in progress. It appears that truncation of the Taylor expansion in eq 10 after including second order derivatives is sufficient for urea and probably reasonable for GdmCl. It is possible that additional terms will be required for TFE solutions. These terms can be obtained quite simply with the use of eq 25 and require no additional parameters. The model could also be improved by considering the concentration dependence of $\Delta\mu_c$.

Acknowledgment. This project was supported by the National Science Foundation. Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the ACS, for partial support of this research.

References and Notes

(1) Pace, C. N.; Shaw, K. L. *Proteins* **2000**, *S4*, 1−7.
(2) Greene Jr, R. F.; Pace, C. N. *J. Biol. Chem.* **1974**, *249*, 5388−5393.
(3) Pace, C. N. *Methods Enzymol.* **1986**, *131*, 266−280.
(4) Makhatadze, G. I. *J. Phys. Chem. B* **1999**, *103*, 4781−4785.
(5) Aune, K. C.; Tanford, C. *Biochemistry* **1969**, *8*, 4586−4590.
(6) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1992**, *226*, 491−505.
(7) Schellman, J. A. *Biopolymers* **1994**, *34*, 1015−1026.
(8) Poland, D. *J. Chem. Phys.* **2000**, *113*, 4774−4784.
(9) Schellman, J. A.; Gassner, N. C. *Biophys. Chem.* **1996**, *59*, 259−275.
(10) Timasheff, S. N. *Adv. Protein Chem.* **1998**, *51*, 355−432.
(11) Record, M. T., Jr.; Anderson, C. F. *Biophys. J.* **1995**, *68*, 786−794.
(12) Courtenay, E. S.; Capp, M. W.; Saecker, R. M.; Record, M. T., Jr. *Proteins* **2000**, *S4*, 72−85.
(13) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 774−777.
(14) Ben-Naim, A. *Statistical Thermodynamics for Chemists and Biochemists*; Plenum Press: New York, 1992.
(15) Chitra, R.; Smith, P. E. *J. Phys. Chem. B* **2002**, *106*, 1491−1500.
(16) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2003**, *118*, 5901−5910.
(17) Friedman, H.; Ramanathan, P. S. *J. Phys. Chem.* **1970**, *74*, 3756−3765.
(18) Kusalik, P. G.; Patey, G. N. *J. Chem. Phys.* **1987**, *86*, 5110−5116.
(19) Smith, P. E. *J. Phys. Chem. B*, submitted for publication, 2004.
(20) Chitra, R.; Smith, P. E. *J. Phys. Chem. B* **2001**, *105*, 11513−11522.
(21) Aburi, M.; Smith, P. E. *J. Phys. Chem. B* **2004**, *108*, 7382−7388.
(22) Ben-Naim A. *Advances in Thermodynamics: Fluctuation Theory of Mixtures*; Taylor and Francis: New York, 1990; pp 211−226.
(23) Tanford, C. *J. Mol. Biol.* **1969**, *39*, 539−544.
(24) Wyman Jr, J. *Adv. Protein Chem.* **1964**, *19*, 223−286.

(25) Casassa, E. F.; Eisenberg, H. *Adv. Protein Chem.* **1964**, *19*, 287−395.
(26) Anderson, C. F.; Record Jr, M. T. *J. Phys. Chem.* **1993**, *97*, 7116−7126.
(27) Parsegian, V. A.; Rand, R. P.; Rau, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3987−3992.
(28) Reiss, H. *Adv. Chem. Phys.* **1966**, *9*, 1−84.
(29) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Protein Sci.* **1995**, *4*, 2138−2148.
(30) Tang, K. E. S.; Bloomfield, V. A. *Biophys. J.* **2002**, *82*, 2876−2891.
(31) Baynes, B. M.; Trout, B. L. *J. Phys. Chem. B* **2003**, *107*, 14058−14067.
(32) Jasanoff, A.; Fersht, A. R. *Biochemistry* **1994**, *33*, 2129−2135.
(33) Chen, L.; Hodgson, K. O.; Doniach, S. *J. Mol. Biol.* **1996**, *261*, 658−671.
(34) Weerasinghe, S.; Smith, P. E. *J. Phys. Chem. B* **2003**, *107*, 3891−3898.
(35) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **121**, 2180−2186.
(36) Rosgen, J.; Pettitt, B. M.; Perkyns, J.; Bolen, D. W. *J. Phys. Chem. B* **2004**, *108*, 2048−2055.