

- Milostone, L., Barton, D. E., Francke, U., & Broadus, A. E. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 597.
- McKee, R. L., Goldman, M. E., Caulfield, M. P., DeHaven, P. A., Levy, J. J., Nutt, R. F., & Rosenblatt, M. (1988) *Endocrinology* 122, 3008.
- McKee, R. L., Caulfield, M. P., & Rosenblatt, M. (1990) *Endocrinology* 127, 76.
- Merrifield, R. B. (1969) *Adv. Enzymol. Subj. Biochem.* 32, 221.
- Mosberg, H. I., Hurst, R., Hruby, V. J., Gee, K., Yamamura, H. I., Galligan, J. J., & Burks, T. F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 5871.
- Mosely, J. M., Kubota, M., Diefenbach-Jagger, H., Wettenthal, R. E. H., Kemp, B. E., Suva, L. J., Rodda, C. P., Ebeling, P. R., Hudson, P. J., Zajac, J. T., & Martin, T. J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 5048.
- Nissenson, R. A., Diep, D., & Strewler, G. J. (1988) *J. Biol. Chem.* 263, 12866.
- Nutt, R. F., Caulfield, M. P., Levy, J. J., Gibbons, S. W., Rosenblatt, M., & McKee, R. L. (1990) *Endocrinology* 127, 491.
- Pizurki, L., Rizzoli, R., Moseley, J., Martin, T. J., Caverzasio, J., & Bonjour, J.-P. (1988) *Am. J. Physiol.* 255, F957.
- Rabbani, S. A., Mitchell, J., Roy, D. R., Hendy, G. N., & Goltzman, D. (1988) *Endocrinology* 123, 2709.
- Rodan, S. B., Insogna, K. L., Vignery, A. M., Stewart, A. F., Broadus, A. E., D'Sousa, S. M., Bertolini, D. R., Mundy, G. R., & Rodan, G. A. (1983) *J. Clin. Invest.* 72, 1511.
- Rosenblatt, M., Segre, G. V., Tyler, G. A., Shepard, G. L., Nussbaum, S. R., & Potts, J. T., Jr. (1980) *Endocrinology* 107, 545.
- Shigeno, C., Yamamoto, I., Kitamura, N., Noda, T., Lee, K., Sone, T., Shiomi, K., Ohtaka, A., Fujii, N., Yajima, H., & Konishi, J. (1988) *J. Biol. Chem.* 263, 18369.
- Stewart, A. F., Insogna, K. L., Goltzman, D., & Broadus, A. E. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 1454.
- Stewart, A. F., Wu, T., Goumas, D., Burtis, W. J., & Broadus, A. E. (1987) *Biochem. Biophys. Res. Commun.* 146, 672.
- Strewler, G. J., Williams, R. D., & Nissenson, R. A. (1983) *J. Clin. Invest.* 71, 769.
- Strewler, G. J., Stern, P. H., Jacobs, J. W., Eveloff, J., Klein, R. F., Leung, S. C., Rosenblatt, M., & Nissenson, R. A. (1987) *J. Clin. Invest.* 80, 1803.
- Suva, L. J., Winslow, G. A., Wettenthal, R. E. H., Hammonds, R. G., Moseley, J. M., Diefenbach-Jagger, H., Rodda, C. P., Kemp, B. E., Rodriguez, H., Chen, E. Y., Hudson, P. J., Martin, T. J., & Wood, W. I. (1987) *Science* 237, 893.
- Thiede, M. A., Strewler, G. J., Nissenson, R. A., Rosenblatt, M., & Rodan, G. A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 4605.
- Veber, D. F., & Freidinger, R. M. (1985) *Trends. Neurosci.* 8, 392.
- Veber, D. F., Holly, F. W., Paleveda, W. J., Nutt, R. F., Bergstrand, S. J., Torchiana, M., Glitzer, M. S., Saperstein, R., & Hirschmann, R. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 2636.
- Veber, D. F., Saperstein, R., Nutt, R. F., Freidinger, R. M., Brady, S. F., Curley, P., Perlow, D. S., Paleveda, W. J., Colton, C. D., Zacchei, A. G., Tocco, D. J., Hoff, D. R., Vandlen, R. L., Gerich, J. E., Hall, L., Mandarin, L., Cordes, E. H., Anderson, P. S., & Hirschmann, R. (1984) *Life Sci.* 34, 1371.

Solvent Denaturation and Stabilization of Globular Proteins[†]

Darwin O. V. Alonso and Ken A. Dill^{*,†}

Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143

Received June 15, 1990; Revised Manuscript Received January 22, 1991

ABSTRACT: Statistical thermodynamic theory has recently been developed to account for the stabilities of globular proteins. Here we extend that work to predict the effects of solvents on protein stability. Folding is assumed to be driven by solvophobic interactions and opposed by conformational entropy. The solvent dependence of the solvophobic interactions is taken from transfer experiments of Nozaki and Tanford on amino acids into aqueous solutions of urea or guanidine hydrochloride (GuHCl). On the basis of the assumption of two pathways involving collapse and formation of a core, the theory predicts that increasing denaturant should lead to a two-state denaturation transition (i.e., there is a stable state along each path separated by a free energy barrier). The denaturation midpoint is predicted to occur at higher concentrations of urea than of GuHCl. At neutral pH, the radius of the solvent-denatured state should be much smaller than for a random-flight chain and increase with either denaturant concentration or number of polar residues in the chain. A question of interest is whether free energies of folding should depend linearly on denaturant, as is often assumed. The free energy is predicted to be linear for urea but to have some small curvature for GuHCl. Predicted slopes and exposed areas of the unfolded states are found to be in generally good agreement with experiments. We also discuss stabilizing solvents and compare thermal with solvent denaturation.

Proteins can be denatured in the presence of certain small-molecule solutes in high concentrations. Examples of

such solutes are urea and guanidine hydrochloride. Other solutes in high concentration stabilize proteins. Examples of these are sugars, glycerol, polyols, and poly(ethylene glycol) (Arakawa, 1982; Back, 1979; Gekko & Timasheff, 1981; Lee & Lee, 1981; Lee & Timasheff, 1981). One problem in understanding the molecular mechanism of nonspecific solute action on protein stability has been the absence of a molecular

[†] We thank the NIH, the Pew Foundation, and the DARPA URI program for financial support.

^{*} Address correspondence to this author.

[†] Present address: UCSF Laurel Heights Campus, 3333 California St., Room 102, San Francisco, CA 94118.

theory for the folded and unfolded states of the protein.

Thermodynamic treatments of this problem have been given by Schellman (1975, 1978, 1987) and Tanford (1970). They are premised on the idea that protein denaturation is a two-state process and that the relevant difference between the two states, native and unfolded, is in the increased area of exposure of hydrophobic residues to the solution (of water plus denaturant) upon denaturation. In the absence of a molecular model for the unfolded state, the thermodynamic models require that the exposed area of hydrophobic residues in the unfolded state be taken simply as an adjustable parameter. Consistency with experimental results has thereby implied that the unfolded state has considerably less area of exposure than a fully open chain molecule would have or that only certain hydrophobic amino acids such as the aromatics are involved (Ahmad & Bigelow, 1986; Tanford, 1970). A principal application of these thermodynamic theories has been toward the justification of the functional dependence of the free energy of folding on the denaturant concentration; experiments generally show this dependence to be linear (Ahmad & Bigelow, 1986; Bolen & Santoro, 1988a,b; Shortle et al., 1989; Thomson & Bigelow, 1986). This functional dependence is required to make the long extrapolation of experimental data to zero denaturant concentration in order to establish the free energy of folding in pure solvent ($\Delta F^\circ_{\text{fold}}$). However, because intrinsic limitations of current experimental methods preclude studies over wide ranges of denaturant concentration, experimental proof or disproof of linearity has not yet been possible. Within the thermodynamic model, Schellman (1975, 1978, 1987) has suggested that linearity can be justified at low denaturant concentration as the lowest order term, if the free energy of folding has a polynomial dependence on the denaturant concentration.

The purpose of the present work is to provide a molecular theory for nonspecific small-molecule solute action on protein stability. Our approach is based on a mean field lattice statistical mechanics theory of protein stability (Dill, 1985; Dill et al., 1989). This treatment assumes the premise that the stability of a folded protein is the result of a balance principally between the solvophobic interaction, on the one hand, which drives the molecule to fold, and the chain conformational entropy, on the other hand, which drives the molecule to unfold. This model has recently been shown to account satisfactorily for aspects of the temperature dependence of protein stability (Dill et al., 1989). Since the principal features of the theory are described elsewhere (Dill, 1985; Dill et al., 1989), only a brief summary is given here.

THEORY

We aim to calculate the free energy change upon folding of the protein as a function of the concentration of denaturant in the solution. Because only state functions are of interest, we are at liberty to choose fictitious pathways from the unfolded to the folded state. We construct two pathways (Dill, 1985; Dill et al., 1989): (I) a pathway from the unfolded state to a state of random condensation, wherein solvophobic and solvophilic residues are randomly distributed throughout a condensed globular structure; and (II) a process in which residues in the condensed state are reconfigured so that solvophilic residues largely surround a core of solvophobic residues. Along path I, the thermodynamic functions depend only on density, ρ , the mean volume fraction of space occupied by the chain segments. Along path II, they depend only on ordering, θ , described below, which accounts for the segregation of solvophilic and solvophobic segments into inside and outside

regions of the protein. ΔF_{fold} is the sum of these free energy changes over the two paths.

Path I: Random Condensation. Consider a protein with n residues, n_h of which are solvophobic and n_p of which are solvophilic. The fractional solvophobicity is given by $\Phi = n_h/n$. At all densities, ρ , the molecule is taken to be spherical with radius r , where ρ and r are related by

$$\rho = n/(\frac{4}{3}\pi r^3) \quad (1)$$

As ρ increases, there are fewer conformations available to the chain, largely because of the volume excluded by other segments, and this contributes to an entropy that disfavors collapse. However, as ρ increases there are more contacts among the solvophobic residues, and this interaction favors collapse. For the change to the randomly condensed state ($\rho = 1$) from the unfolded state ($\rho = \rho^*$), the solvophobic and conformational contributions to the free energy difference per residue are (Dill, 1985; Dill et al., 1989)

$$\frac{\Delta F_{h\phi I}}{nkT} = -\chi(c)\Phi^2(f_i + \sigma f_e)(1 - \rho^*) \quad (2)$$

$$\frac{\Delta F_{\text{conf I}}}{nkT} = \frac{-(1 - \rho^*)}{\rho^*} \ln(1 - \rho^*) + \frac{7}{2n}[\rho_0^{2/3} - (\rho_0/\rho^*)^{2/3}] - \frac{2}{n} \ln \rho^* \quad (3)$$

where kT is Boltzmann's constant multiplied by absolute temperature, and $f_i(r)$ and $f_e(r)$ are the fractions of the spherical volume in the unfolded state that correspond to interior and exterior sites, respectively (Dill et al., 1989). The reference state density is $\rho_0 = [19/(27n)]^{1/2}$ (Dill, 1985). The fractional exposure of surface residues to bulk solvent is $(1 - \sigma)$; hence σ is the fractional exposure of surface residues to contacts (solvent or residues) within the protein. The strength of the solvophobic interaction is characterized by $\chi(c)$. χ is the free energy of transfer divided by kT of an average hydrophobic residue from a pure medium of such residues to the aqueous solution of solvent plus denaturant. This free energy of transfer depends on the concentration, c , of denaturant in the aqueous solution. The functional dependence of $\chi(c)$ is described in detail in the following sections.

The change in free energy along path I is given by the sum of these two contributions

$$\Delta F_I = \Delta F_{h\phi I} + \Delta F_{\text{conf I}} \quad (4)$$

The equilibrium density, ρ^* , of the unfolded state is calculated by finding the minimum in free energy by solving the equation

$$0 = \left. \frac{\partial(F_I/nkT)}{\partial \rho} \right|_{\rho^*} = -\Phi^2\chi(c)(f_i + \sigma f_e) - \frac{7}{3n} \frac{\rho_0^{2/3}}{\rho^{5/3}} - \frac{1}{\rho^*} \left[\frac{\ln(1 - \rho^*)}{\rho^*} + 1 - \frac{2}{n} \right] \quad (5)$$

Path II: Chain Reconfiguration. From the randomly condensed state, an even greater number of favorable solvophobic contacts will be made if the chain is free to reconfigure so that solvophilic residues are largely on the outside of the molecule surrounding a core of solvophobic residues. However, this process must compete with the decrease in entropy due to ordering, which arises from segregating the two classes of residues into the two spatial regions of the molecule. The variables characterizing this process are θ , which is the fraction of surface sites occupied by solvophobic residues, and x , which

is the fraction of core sites occupied by solvophobic residues. The hydrophobic residues at the surface plus those in the core must equal the total number of hydrophobic residues; thus, θ and x are constrained by

$$\theta f_e + x f_i = \Phi \quad (6)$$

The contact and conformational contributions to the free energy change of reconfiguration are (Dill, 1985; Dill & Alonso, 1988)

$$\frac{\Delta F_{h\Phi II}}{nkT} = -\chi(c)[f_i(x^{*2} - \Phi^2) + \sigma f_e(\theta^{*2} - \Phi^2)] \quad (7)$$

$$\frac{\Delta F_{conf II}}{nkT} = f_i \left[x^* \ln \left(\frac{x^*}{\Phi} \right) + (1 - x^*) \ln \left(\frac{1 - x^*}{1 - \Phi} \right) \right] + f_e \left[\theta^* \ln \left(\frac{\theta^*}{\Phi} \right) + (1 - \theta^*) \ln \left(\frac{1 - \theta^*}{1 - \Phi} \right) \right] \quad (8)$$

and the change in free energy along path II is

$$\Delta F_{II} = \Delta F_{h\Phi II} + \Delta F_{conf II} \quad (9)$$

The equilibrium values of the order parameters, x^* and θ^* , are obtained by simultaneous solution of constraint eq 6 and

$$\left. \frac{\partial(F_{II}/kT)}{\partial \theta} \right|_{\theta^*} = \ln \left[\frac{x^*(1 - \theta^*)}{\theta^*(1 - x^*)} \right] - 2\chi(c)[x^* - \sigma \theta^*] = 0 \quad (10)$$

The total free energy of folding is the sum of the free energies from the two pathways, eqs 4 and 9

$$\Delta F_{fold}(n, \Phi, \chi(c)) = \Delta F_I + \Delta F_{II} \quad (11)$$

Denaturant Effects on Amino Acid Transfer: $\chi(c)$. Our aim is to calculate the free energy of folding of a protein as a function of denaturant concentration, $\Delta F_{fold}(c)$. We can do this using eq 11, provided χ , the elementary free energy of transfer of a hydrophobic amino acid side chain, is known as a function of c .

We obtain $\chi(c)$ for individual hydrophobic amino acid residues in denaturant solutions of either urea or GuHCl from experimental transfer measurements at 25 °C (McMeekin et al., 1936; Nozaki & Tanford, 1963, 1970, 1971). By measuring the free energy of transfer of a given amino acid from water to hydrocarbon and subtracting the measured free energy of transfer of glycine, Nozaki and Tanford obtained the free energy of transfer of a given side chain from hydrocarbon to water, g_h^w (Nozaki & Tanford, 1971). In other experiments (Nozaki & Tanford, 1963, 1970), they measured the free energy of transferring the side chain from water to the solution of water plus denaturant at concentration c , $g_w^s(c)$. In terms of these quantities, $\chi(c)$ is given as

$$\frac{kT}{1.4} \chi(c) = g_{tr}(c) = g_h^w + g_w^s(c) \quad (12)$$

where the factor of 1.4 accounts for the difference in dimensions between an amino acid residue and a cubic lattice segment required in the model (Dill, 1985). We neglect here the insignificant difference between Gibbs and Helmholtz free energies. The concentration dependences for individual amino acid side chains thus obtained from the Nozaki and Tanford experiments are shown in Figures 1 and 2. For consistency with our previous applications of this theory, we take as hydrophobic the Nozaki and Tanford list of eight amino acids: Ala, Val, Leu, Ile, Met, Tyr, Phe, and Trp. It is evident from the data shown in Figures 1 and 2 that the transfer of hydrophobic groups into aqueous solutions is unfavorable due

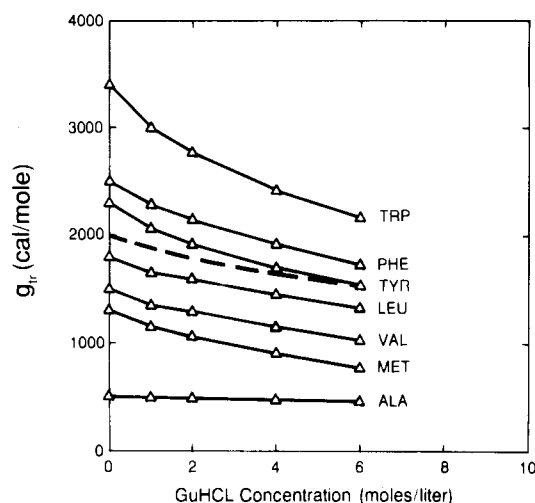


FIGURE 1: Experimental free energies of transfer of hydrophobic residues (solid lines) and their dependences on GuHCl concentration (eq 13) (Nozaki & Tanford, 1970, 1971). The heavy dashed line is the fit for $g_{tr}(c_{GuHCl})$ (eq 14).

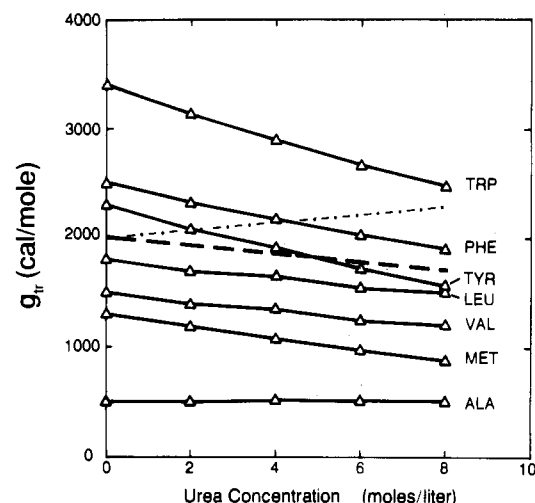


FIGURE 2: Experimental free energies of transfer of hydrophobic residues (solid lines) and their dependences on urea concentration (eq 13) (Nozaki & Tanford, 1963, 1971). The heavy dashed line is the fit for $g_{tr}(c_{urea})$ (eq 15). The dot-dashed line represents $g_{tr}(c_{stabilizer})$ used for the hypothetical stabilizer.

to the positive g_{tr} but becomes more favorable with increasing denaturant concentration.

We must establish an appropriate functional form for $\chi(c)$ from the data described above. This function must satisfy two requirements. First, it must represent in some sense a "typical" hydrophobic amino acid in its dependence on denaturant concentration ($g_w^s(c)$). Second, consistency requires that the transfer free energy in the limit of zero denaturant concentration must reduce to the standard experiments for the transfer to pure water (g_h^w). From transfer experiments reported in the literature, we have previously determined that the free energy of transfer of a typical hydrophobic amino acid at 25 °C in the absence of denaturant is 2.0 kcal/mol (Dill et al., 1989). Among the hydrophobic amino acids on which denaturant concentration experiments have been performed, we chose leucine as the most representative of denaturant transfer dependence for the following reasons. It is clear from Figures 1 and 2 that the curve for leucine represents the median for these eight amino acids. In addition, many transfer properties have been shown to depend linearly on the number of side-chain carbons (Prasad & Ahluwalia, 1976; Spink & Wadso, 1975; Tanford, 1980). For this set of eight amino

acids, the average number of side-chain carbons is 4.6. Leucine is the amino acid closest to this value. Alternatively, a mean value of g_w^s for the amino acids at each concentration could be used. The mean gives qualitatively similar results but a slightly steeper dependence of $g_w^s(c)$ on c .

We use a parameterized form of the data for $g_w^s(c)_{\text{Leu}}$ in order to interpolate between the urea or GuHCl concentrations at which $g_w^s(c)$ values were experimentally measured. A quadratic functional dependence on concentration of denaturant is justified on the basis of simple solution theories; that is

$$g_w^s(c) = g_1c + g_2c^2 \quad (13)$$

For ternary solutions in which the three components are randomly mixed, the free energy of transfer can readily be shown (Dill, 1987; Hildebrand & Scott, 1950) to obey a quadratic dependence on cosolvent (the denaturant in this case). The first-order contributions (g_1) account for the interaction of the solute (the hydrophobic amino acids, in this case) with either water or with the denaturant. The second-order contribution accounts for the interactions of the cosolvents (water and denaturant) with each other. The values g_1 and g_2 will depend on the chemical nature of the denaturant. In general, the second-order term will be small when the cosolvents are compatible with each other, and in that case, the free energy should be linear in denaturant concentration. On the basis of these arguments, we use quadratic functions to fit the data of Nozaki and Tanford for leucine.

As indicated above, consistency with previous work requires that g_h^w for an average nonpolar amino acid must equal 2.0 kcal/mol. Given this value for g_h^w and the quadratic fit for $g_w^s(c)_{\text{Leu}}$, the transfer of a hydrophobic side chain is described by

$$\frac{kT}{1.4}\chi(c_{\text{GuHCl}}) = g_{\text{tr}}(c_{\text{GuHCl}}) = 2.0 \times 10^3 - 111c_{\text{GuHCl}} + 5.76c_{\text{GuHCl}}^2 \quad (14)$$

$$\frac{kT}{1.4}\chi(c_{\text{urea}}) = g_{\text{tr}}(c_{\text{urea}}) = 2.0 \times 10^3 - 36.8c_{\text{urea}} \quad (15)$$

where the second-order term is small for GuHCl and negligible for urea. In eqs 14 and 15, concentration is in units of moles per liter, g_{tr} is in units of calories per mole of hydrophobic side chains, and χ is a dimensionless free energy. We plot these representative functions for $g_{\text{tr}}(c)$ for GuHCl and urea (eqs 14 and 15) as the heavy solid lines in Figures 1 and 2, respectively. It is clear from Figure 2, and has been noted before (Creighton, 1979), that 8 M urea or 6 M GuHCl does not fully abolish the hydrophobic interaction; they merely reduce it by less than 20%. To obtain the free energy of folding of the protein, we now substitute either eq 14 or eq 15 into eqs 2, 5, 7, and 10 and solve eq 11 for $\Delta F_{\text{fold}}(c)$.

RESULTS

The theoretically predicted denaturation profiles for proteins in GuHCl and urea are shown in Figure 3. As an example, we chose as representative of a small single-domain protein $n = 100$ (140 amino acids) and $\Phi = 0.45$ at $T = 25^\circ\text{C}$. From the free energy of folding given by the theory, we calculate the fraction of the protein molecules that are native, $f_n(c)$, as a function of denaturant concentration, c , using the expression

$$f_n(c) = \frac{\exp[-\Delta F_{\text{fold}}(\chi(c))/kT]}{1 + \exp[-\Delta F_{\text{fold}}(\chi(c))/kT]} \quad (16)$$

The theory predicts that the midpoint for denaturation is approximately 1.8 M GuHCl or 5.0 M urea for this example.

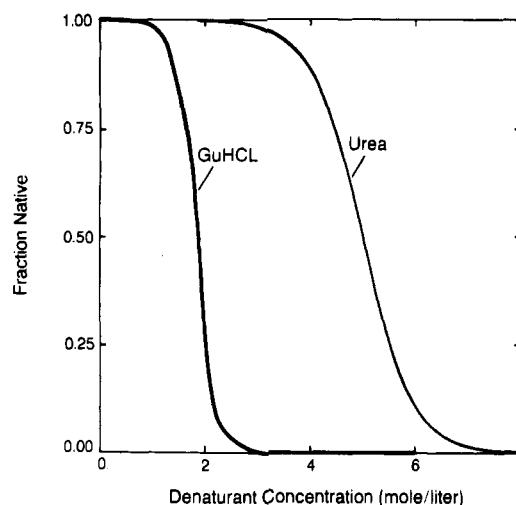


FIGURE 3: Theoretical denaturation curve plots of the fraction of protein molecules in the native form as a function of denaturant concentrations for a representative protein of $\Phi = 0.45$ and $n = 100$.

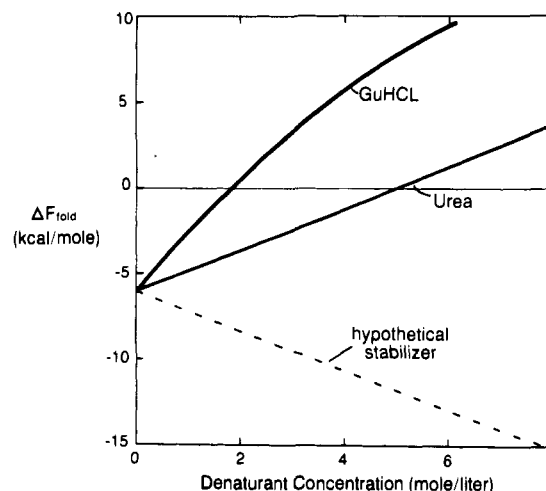


FIGURE 4: Free energy of folding as functions of solvent concentration for GuHCl, urea, and a hypothetical stabilizer for the same representative protein shown in Figure 3 ($\Phi = 0.45$ and $n = 100$). The c_m 's are the concentrations where $\Delta F_{\text{fold}}(\text{GuHCl})$ or $\Delta F_{\text{fold}}(\text{urea})$ equal zero.

For the purposes below, we define this midpoint denaturant concentration to be c_m .

Figure 4 shows the stability (ΔF_{fold}) vs denaturant concentration. The midpoint of denaturation occurs where $\Delta F_{\text{fold}} = 0$. A principal prediction, shown in Figure 4, is that the free energy of folding is linear function of urea over a wide concentration range. Linearity is found to be a worse approximation for GuHCl. These conclusions hold irrespective of chain length and over the composition range $\Phi = 0.38$ –1.0. Nonlinearity of $\Delta F_{\text{fold}}(\chi(c))$ with c could arise from curvature in either (i) $\chi(c)$ or (ii) $\Delta F_{\text{fold}}(\chi)$. The theory predicts that there is a very small degree of inherent curvature, concave upwards, of $\Delta F_{\text{fold}}(\chi)$. This inherent curvature is predicted to increase with hydrophobic composition, Φ . It is a result of the decreasing density of the unfolded state with decreasing χ . The principal curvature in $\Delta F_{\text{fold}}(c)$, however, arises from curvature in $\chi(c)$. For urea, the combined curvatures are very small; hence, $\Delta F_{\text{fold}}(c)$ is approximately linear with c . However the nonlinearity in $\chi(c)$ for GuHCl leads to nonlinearity in $\Delta F_{\text{fold}}(c)$, with curvature concave downwards. It should be noted that Figure 4 spans concentration and free energy ranges far beyond those accessible for a single protein with the common experimental spectrophotometric methods. Experimen-

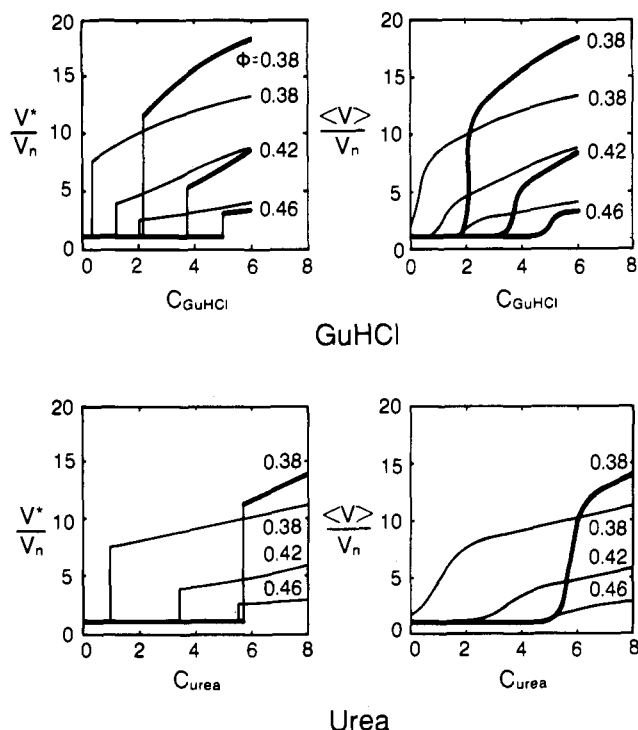


FIGURE 5: Volume of protein molecule relative to the native state as a function of denaturant concentrations. The left hand column shows v^* , the volume of the lowest free energy state, and the right hand column shows $\langle v \rangle$, the average volume (eq 17) (light lines, $n = 100$; heavy lines, $n = 200$). In two cases ($n = 200$, $\Phi = 0.42$ and $n = 200$, $\Phi = 0.46$), denaturation occurs only above 8 M urea.

tally, ΔF_{fold} can be measured over a range of roughly ± 1.4 kcal/mol around the denaturation point. When the theoretical GuHCl plot shown in Figure 4 is constrained to that range, it will appear to be essentially linear. The nonlinearities are sufficiently small that a linear extrapolation from the denaturation point (1.8 M GuHCl) to 0 M GuHCl would lead to an underestimate of the stability of only 0.5 kcal/mol (relative to $\Delta F_{\text{fold}} = -6$ kcal/mol) for this typical case.

We are aware of only one experiment (Pace & Vanderburg, 1979) that has attempted to span a wide concentration range to determine the possible linearity of ΔF_{fold} vs c . Pace and Vanderburg denatured myoglobin with GuHCl with different concentration ranges of GuHCl, and they obtained separate plots of $\Delta F_{\text{fold}}(c)$ at each pH. They then pieced together the separate data from each curve at differing pH's (pH = 6.0, 4.9, 4.7, and 4.6) into one curve corrected to pH = 7. The correction involves adding a different constant to the free energies for each pH set to compensate for the effect of the pH changes. These constants were chosen so that each data set superimposes on the data for the neighboring pH's. These constants were close to the values calculated from Puett's equation (Puett, 1973). They observed that $\partial \Delta F_{\text{fold}} / \partial c$ is positive but decreases with increasing c (i.e., concave downwards), in agreement with the theoretical prediction. The predicted curvature is slightly smaller than that of the experiments. It is not clear that more quantitative comparison is yet warranted in light of the inapplicability of the theory to a single protein and of the simplifications implicit in the experimental patching procedure (Pace et al., 1990).

Figure 5 shows two predictions of the theory: that unfolding is a two-state transition and that the unfolded state continues to expand for denaturant concentrations well above c_m . The figures on the left show the volume ($v = n/\rho \propto r^3$) of the lowest free energy state of the protein as a function of denaturant concentration.

The theory predicts a first-order transition, i.e., the maximum possible degree of cooperativity. That is, it predicts two stable states separated by a free energy barrier. We comment on two points regarding the nature of this two-state transition. First, these are not two fixed states. Although the properties of the native state in the model are found to be relatively independent of external conditions, those of the denatured state are not. In that regard, the two-state behavior predicted by the model is different than the two states described in a general $A \leftrightarrow B$ reaction or phase change. For example, increasing denaturant increases the volume of the denatured state and thereby reduces the number of nonpolar contacts, thus changing the thermodynamic state functions. The second point is to note that one assumption of the model is that the possible stable states may occur only along either of the two pathways: condensation or formation of a core in the compact chain. A full proof of the first-order nature of the transition would require demonstration that the barrier remains even if the stable states are not limited to these two paths. We have not yet undertaken this broader test. In essence, the present restriction is equivalent to the plausible assumption that no hydrophobic core is possible for chains less than fully compact. For these less compact chains, the model of course does allow the freedom for considerable hydrophobic clustering; it is simply that the distinction between inside and outside disappears for chains of less than maximal compactness.

The increase in volume upon denaturation can be large and is strongly dependent on the composition, Φ , the fraction of residues that are hydrophobic. For $\Phi = 0.46$ and $n = 100$, the volume increase upon denaturation relative to the native state is only 2-fold. For $\Phi = 0.38$ and $n = 100$, it is 8-fold.

Figure 5 also shows that the volume of the unfolded molecule continues to increase with denaturant concentration beyond the denaturation midpoint. Added denaturant further improves the solvation of the nonpolar residues. Although this further volume expansion is more gradual with c than the change at denaturation, it can be considerable: the volume increases by nearly an additional factor of 2 over the concentration range 0.3–6 M GuHCl for the chain with $\Phi = 0.38$.

The right-hand side of Figure 5 shows the average volume

$$\langle v \rangle = v_n f_n + v_d f_d \quad (17)$$

Subscripts n and d refer to native and denatured states, and v and f are the volumes and fractions of molecules in these states respectively ($f_d = 1 - f_n$, where f_n is given by eq 16). The volume that is accessible to experiment is $\langle v \rangle$, which can be obtained by viscosity, light scattering, or gel exclusion chromatography. This quantity shows a more gradual transition than v^* (where v^* is the volume of either the native or denatured state, whichever is stable at a particular denaturant concentration), because $\langle v \rangle$ is an average over both the native and denatured populations. It is interesting to note the effect of averaging: the case of $\Phi = 0.38$, $n = 100$ (right-hand side of Figure 5) is a protein that is native at zero denaturant concentration but is so marginally stable that a significant population of unfolded molecules contributes to the observed volume at zero denaturant concentration. Hence the sigmoidal nature of these denaturation curves is scarcely evident.

Figure 6 shows how the two contributions to the free energy—the contact interactions and the conformational entropy—are affected by denaturant concentration. We compare two hypothetical proteins, one with $\Phi = 0.38$ and the other with $\Phi = 0.45$. Although the denaturant directly influences the contact interactions, it also indirectly influences the conformational entropy difference, since the latter depends on the nature of the unfolded state. This is evident from Figure

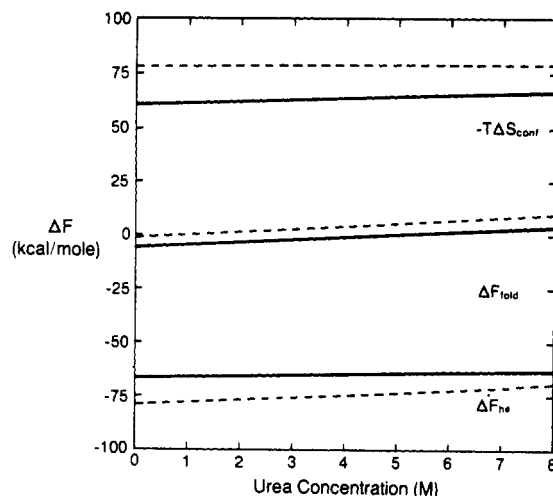


FIGURE 6: Conformational free energy ($-T\Delta S_{\text{conf}}$), contact contributions (ΔF_{hb}), and the total free energy of folding as functions of urea concentration for two different hypothetical proteins of $n = 100$ (dashed line, $\Phi = 0.38$; solid line, $\Phi = 0.45$).

6 for $\Phi = 0.45$ (see the curve for $-T\Delta S_{\text{conf}}$). However, for $\Phi = 0.38$, the unfolded conformations are nearly maximally exposed to the solvent. In that case ρ^* (the density of the unfolded state) decreases very little as χ decreases. Thus, for $\Phi = 0.38$ the conformational entropy difference upon folding is essentially independent of the denaturant concentration. For $\Phi = 0.45$, the contact free energy is relatively insensitive to the denaturant concentration because (i) the unfolded state becomes more exposed with increasing c but (ii) the contact interactions themselves weaken with increasing c . These two effects partially compensate.

Protein Stabilizers. It follows from the foregoing discussion that additives to the aqueous solution that increase, rather than decrease, the free energy of transfer of hydrophobic side chains from hydrocarbon to aqueous solution should have a stabilizing effect on proteins (i.e., $\partial\Delta F_{\text{fold}}/\partial c < 0$, if $\partial\chi/\partial c > 0$). That is, proteins should be stabilized by any solution that is an even less favorable environment for hydrophobic side chains than water itself. This may be the mechanism of glycerol and sucrose stabilization. We are unaware of experimental measurements of $\chi(c)$ for any stabilizer and hence cannot test the theory against experiments in cases where the effects of stabilizers on proteins were studied (Lee & Lee, 1981; Lee & Timasheff, 1981). However, shown in Figure 4 is a hypothetical stabilizer, equivalent in strength but opposite in sign to urea (dashed line in Figure 2), that is

$$\frac{kT}{1.4}\chi(c_{\text{stab}}) = g_{\text{tr}}(c_{\text{stab}}) = 2.0 \times 10^3 + 36.8c_{\text{stab}}$$

The effect of this hypothetical stabilizer on the free energy of folding for a representative protein ($n = 100$ and $\Phi = 0.45$) is shown as the dashed line in Figure 4.

Change in Solvent Accessibility upon Folding. There has been considerable interest in the physical interpretation of the slope of stability with denaturant concentration, $\partial\Delta F_{\text{fold}}/\partial c$, because it sheds light on the unfolding process and the states accessible to a protein at the denaturation point. Here we discuss the relationship between this slope and the exposed hydrophobic area in terms of the theoretical model. Combining the free energy contributions from the two pathways and separating the contributions involving (i) contact interactions (eqs 2 and 7) and (ii) conformational entropy (eqs 3 and 8) leads to

$$\Delta F_{\text{fold}}/kT = \chi\Delta A - \Delta S_{\text{conf}}/k \quad (18)$$

where

$$\Delta S_{\text{conf}}/k = \frac{n(1-\rho^*)}{\rho^*} \ln(1-\rho^*) + \frac{7}{2}[\rho_0^{2/3} - (\rho_0/\rho^*)^{2/3}] - 2 \ln \rho^* - n f_i \left[x^* \ln \left(\frac{x^*}{\Phi} \right) + (1-x^*) \ln \left(\frac{1-x^*}{1-\Phi} \right) \right] - n f_e \left[\theta^* \ln \left(\frac{\theta^*}{\Phi} \right) + (1-\theta^*) \ln \left(\frac{1-\theta^*}{1-\Phi} \right) \right] \quad (19)$$

and

$$\Delta A = A_{\text{folded}} - A_{\text{unfolded}} = -n[(f_i x^{*2} + \sigma f_e \theta^{*2}) - \Phi^2(f_i + \sigma f_e)\rho^*] \quad (20)$$

is the total change in nonpolar contact area upon folding (ranging from 0 to $-n_h$) in units of lattice site area or $\Delta A/1.4$ in units of area of the nonpolar amino acids. The fractional change in contact area is $\Delta f_A = \Delta A/n_h$. For a process in which every nonpolar group is fully exposed in the unfolded state and fully buried in the folded state, $\Delta f_A = -1$. The individual components of this change are f_A^N , the fractional area exposed in the native state, and f_A^D , the fractional area exposed in the denatured state, given by

$$f_A^N = 1 - \frac{1}{\Phi}[f_i(x^{*2}) + \sigma f_e(\theta^{*2})] \quad (21)$$

$$f_A^D = 1 - \Phi(f_i + \sigma f_e)\rho^* \quad (22)$$

We note an important distinction between two different interpretations of exposure. In the type described in eqs 21 and 22, it refers to the exposure of nonpolar residues to *anything other than nonpolar residues*, i.e., to solvent or polar residues. In the second instance, eqs 23 and 24, below, exposure is defined specifically as nonpolar contacts with solvent. The fractional exposures to the solvent are

$$f_S^N = (1-\sigma)f_e\theta/\Phi \quad (23)$$

and

$$f_S^D = 1 - (f_i + \sigma f_e)\rho^* \quad (24)$$

Since the interactions of nonpolar residues with solvent, on the one hand, or with polar residues, on the other, are taken to be approximately energetically equal in the present model, then it is the first definition of exposure (eqs 21 and 22) that is relevant to the folding thermodynamics. On the other hand, some spectroscopic experiments measure only nonpolar exposure to the solvent. These two quantities can be considerably different in general.

Now taking the derivative of eq 18 leads to

$$\frac{\partial\Delta F_{\text{fold}}}{\partial c} = \frac{\partial\chi}{\partial c} \left[\Delta A + \left(\chi \frac{\partial\Delta A}{\partial\chi} - \frac{\partial\Delta S_{\text{conf}}/k}{\partial\chi} \right) \right] kT \quad (25)$$

The terms in the inner parentheses in eq 25 are of opposite sign and small and on balance add only about $+(5-10)\%$ to the ΔA term. Hence, a reasonable approximation to eq 25 is

$$\frac{\partial\Delta F_{\text{fold}}}{\partial c} \approx \Delta A \frac{\partial\chi}{\partial c} kT \quad (26)$$

and

$$\Delta A \approx \frac{1}{1.4} \frac{\partial\Delta F_{\text{fold}}}{\partial g_{\text{tr}}} = \frac{\partial\Delta F_{\text{fold}}/\partial c}{1.4\partial g_{\text{tr}}/\partial c} \quad (27)$$

The factor of 1.4 arises from the conversion between cubic lattice site units and amino acid residues, and the change in

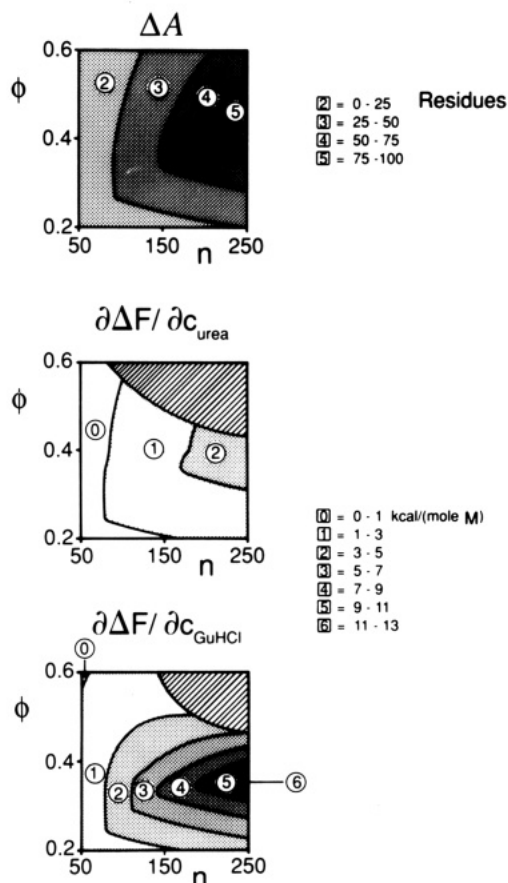


FIGURE 7: Contours of change in solvent-exposed hydrophobic area (ΔA) and denaturation slopes ($\partial\Delta F_{\text{fold}}/\partial c$) as functions of n and Φ . These are the values at the denaturation point or in the case where the folded state is not stable as $c \rightarrow 0$. The lined region in the upper right corner of the GuHCl and urea plots represents n and Φ compositions that are not denatured before 6 M GuHCl or 8 M urea.

area in units of average hydrophobic amino acids is $1.4\Delta A$. This approximate result justifies the general premise of the thermodynamic models that $\partial\Delta F_{\text{fold}}/\partial c$ is linearly dependent on the change in exposure (Ahmad & Bigelow, 1986; Schellman, 1975, 1978, 1987; Tanford, 1970). It is also interesting to note that $\partial\Delta F_{\text{fold}}/\partial c$ should be relatively independent of the nature of which hydrocarbon medium is used as a model for the transfer process to the protein interior, since that determines principally only the constant g_h^w in eq 13, which vanishes upon taking the derivative $\partial g_{\text{tr}}/\partial c$. The constant g_1 represents the process of exchanging a denaturant molecule at the surface of the nonpolar residue for an equivalent volume of water.

A drawback of thermodynamic theories of solvent denaturation is that they assume ΔA is a constant, independent of chain length and composition and independent of solvent and temperature. In those treatments, ΔA is simply taken to be an adjustable parameter. The present theory, on the other hand, predicts that these quantities will depend substantially on chain length and composition and on external thermodynamic quantities. The size and composition of the protein affect $\partial\Delta F_{\text{fold}}/\partial c$ because they affect the amount of hydrophobic surface area that can be buried upon folding, i.e., ΔA . The denaturation slope, $\partial\Delta F_{\text{fold}}/\partial c$, arises approximately as the product of two factors (eq 27): the slope of the elementary transfer free energy with denaturant concentration, $\partial g_{\text{tr}}/\partial c$, and ΔA . The transfer slope, $\partial g_{\text{tr}}/\partial c$, depends only on temperature and the chemical composition of the solution (i.e., the denaturant) and only insignificantly on the structure of the protein. Only ΔA depends appreciably on the structure

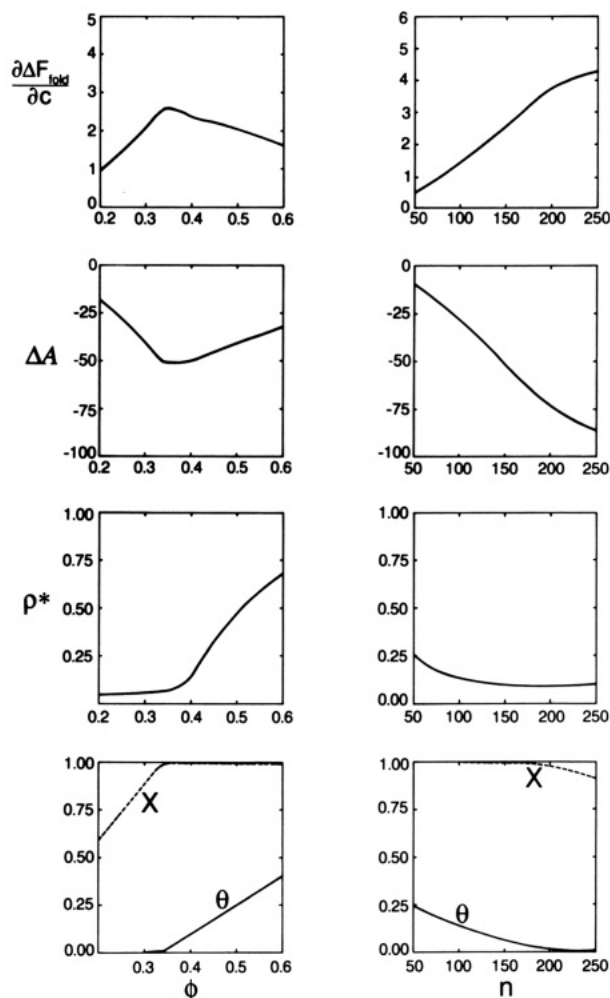


FIGURE 8: Cross-sections of Figure 7. The left column shows dependence on Φ at constant $n = 150$, and the right column shows dependence on n at constant $\Phi = 0.38$.

of the protein. We show the dependence of ΔA and the denaturation slopes (i.e., $\partial\Delta F_{\text{fold}}/\partial c$ for GuHCl and urea) on n and Φ in the form of contour graphs in Figure 7. These results are obtained from use of eqs 20 and 25; they cover the ranges of n and Φ values that include most small single-domain globular proteins. For some purposes, it is easier to read specific cross-sections of this contour diagram, so they are provided in Figure 8. Several properties characterizing the folded and unfolded states are shown vs. Φ in the left column and vs. n in the right column of Figure 8.

The principal observation shown in Figures 7 and 8 is that ΔA , the difference in exposed area in folded and unfolded states, strongly depends on chain length, n , and composition, Φ . This same dependence also arises in connection with thermal unfolding of proteins (Dill et al., 1989). There are two regimes in the dependence on composition of ΔA and $\partial\Delta F_{\text{fold}}/\partial c$. First, for small Φ , the slope $\partial\Delta F_{\text{fold}}/\partial c$ increases as nonpolar content, Φ , increases because ΔA becomes more negative with increasing Φ (note that $\partial g_{\text{tr}}/\partial c < 0$ for denaturants). An increase in Φ leads to greater difference in exposure upon folding, because with increasing hydrophobic content there is more hydrophobic surface area in the unfolded state that can be buried in the native core. Thus, for small Φ , $\partial\Delta F_{\text{fold}}/\partial c$ should become more positive as the nonpolar fraction of residues increases.

In the second regime, for higher Φ , the slope $\partial\Delta F_{\text{fold}}/\partial c$ decreases with increasing nonpolar content. This somewhat more puzzling behavior arises for two reasons. (i) Proteins

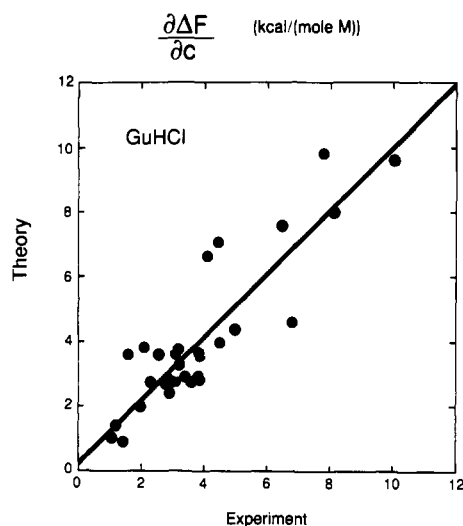


FIGURE 9: Plot of theoretical vs experimental values of GuHCl denaturation slopes; see also Table I. The linear least-squares best-fit line is also shown; it has slope = 0.97, intercept = 0.26, and correlation coefficient = 0.89.

with high contents of nonpolar residues will have dense unfolded states (Figure 8). Each additional nonpolar residue will further increase the density of the unfolded state. Hence, each added nonpolar residue causes less change in exposure (i.e., ΔA is less negative) upon folding, relative to molecules with fewer nonpolar groups. (ii) When the nonpolar content is large, the folded protein will have filled its core, which has limited size, and additional nonpolar residues will be added to surface sites. Surface nonpolar residues contribute less to $\partial\Delta F_{\text{fold}}/\partial c$ than buried residues.

The model can be compared with experiments that have estimated ΔA . However, only qualitative comparison is currently warranted, since the limited experimental data base is often obtained for proteins in acid or base solutions where the theory is not applicable. Solvent denaturation experiments generally show relatively small values of Δf_A , in the range of 0.14–0.41 (Ahmad & Bigelow, 1986; Tanford, 1970). Two factors appear to be responsible for such small Δf_A 's. First, proteins typically studied showed significant amounts of solvent exposure of the nonpolar amino acids in the native state (f_S^N). For the example of native ribonuclease A, Lee and Richards (1971) calculate $f_S^N = 0.06$ from the crystal structure coordinates and Schrier and Schrier (1976) experimentally found $f_S^N = 0.11$. Second, there is often considerable nonpolar burial in the denatured state (f_S^D); Schrier and Schrier report $f_S^D = 0.36$ for denatured ribonuclease A. In contrast, the experiments of Tanford (1968) show considerably greater nonpolar exposure in the denatured state. These experiments were performed in acid solutions at high GuHCl concentrations (>5.0 M GuHCl) and with reduced disulfides. In the high GuHCl experiments, Tanford estimates the solvent exposure is $f_S^D = 0.75$, on the basis of solvent perturbation of the spectra of tyrosine and tryptophan residues. Under those conditions, he also shows that $[\eta]M_0 \propto n^{2/3}$ for a series of different proteins, where $[\eta]$ is the intrinsic viscosity, n is the chain length, and M_0 is the mean residue weight. That is, the radius scales with chain length in the same way as expected for a random-flight chain in a good solvent. In agreement, the theory also predicts that the radius will be large and will depend principally on n and not significantly on Φ at those high GuHCl concentrations where g_{tr} is small.

In Figures 9 and 10 and Tables I and II, we compare the theoretical slope $\partial\Delta F_{\text{fold}}/\partial c$ (given by eq 25) with the experimental data base compiled by Hurle (M. Hurle, personal

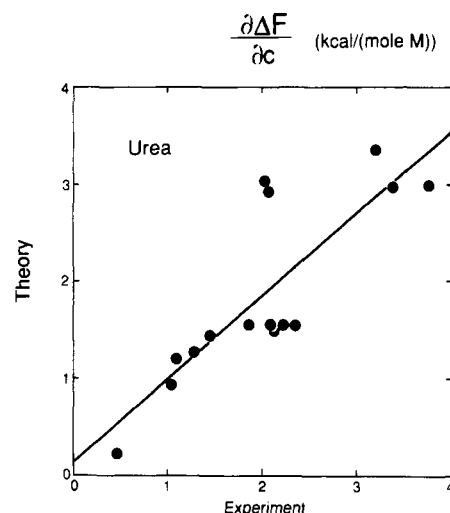


FIGURE 10: Plot of theoretical vs experimental values of urea denaturation slopes; see also Table I. The linear least-squares best-fit line is also shown; it has slope = 0.84, intercept = 0.15, and correlation coefficient = 0.83.

communication). Each theoretical prediction requires values for $\chi(c)$, T , n , and Φ . We use eqs 14 and 15 for $\chi(c)$, $T = 298.15$ K, and the known chain lengths and compositions of the proteins, where the nonpolar residues are as defined above. When experimental slopes were reported at different pH's, we used only the value for the pH of maximum stability, since the theory only applies to conditions where there are no electrostatic contributions to the free energy of folding. The most appropriate test of the theory is any experimental condition under which there is no electrostatic free energy, since the present theory does not include it. For any given protein, this will occur at some unknowable pH. The pH of zero electrostatics is the isoelectric point only if specific ion effects are negligible. Moreover, often the isoelectric pH is not known from the set of data, whereas the pH of maximum stability is readily obtained. Therefore, we use data taken at the pH of maximum stability, with the assumptions that (i) it is not very different from the isoelectric pH and (ii) within this pH range near neutral, electrostatics is not a dominant factor in protein stability. The main point is simply that for the present comparisons, it is advantageous to minimize the effects of electrostatics and an experimentally simple and tractable way to do so is to choose only data near the pH of maximum stability. We also eliminated slopes determined outside the temperature range 20–25 °C, because eqs 14 and 15 are based on experimental data determined at 25 °C. Figures 9 and 10 show that the predictions of the theory are in good agreement with both urea and guanidine experiments, at least as well as could be expected given the considerable scatter in the experimental data. The principal variations in experimental data are likely due to (i) differences in pH, and therefore of electrostatic contributions to the free energy; (ii) differences in the number of disulfide bonds; (iii) the presence of specific components such as heme groups; (iv) other specific interactions including hydrogen bonding and salt bridges; and (v) differences in temperature, since the effects of urea and GuHCl on nonpolar solubility vary with temperature (Wetlaufer et al., 1964).

The present model is nonspecific insofar as (i) it accounts only for effects of composition, and not of sequence; (ii) it is a mean-field model; and (iii) it treats only the interactions of the nonpolar groups with the solvent and not other types of interactions such as electrostatics. Therefore, this model is not suitable for predicting the effects of single-site mutations

Table I: GuHCl Denaturation Slopes and Changes in Exposed Hydrophobic Area (1.4Δ*A*)^a

protein	<i>n</i>	Φ	denaturation slopes, Δ <i>F</i> _{fold} /Δ <i>c</i> (kcal/mol M)		change in exposure, Δ <i>F</i> _{fold} /Δ <i>g</i> _{tr} (residues)		
			expt	theory	expt	theory	ss
kibyg ^{a,b}	415	0.424	24.00	9.22	-229.7	-144.9	0
kihug ^{a,c}	417	0.420	10.10	9.59	-98.2	-143.6	0
tseca ^{a,d}	268	0.478	8.20	7.95	-81.6	-80.7	0
pepg ^{a,e}	370	0.395	7.80	9.85	-78.6	-120.5	3
crbo2 ^{a,f}	259	0.374	7.80	8.01	-87.6	-89.6	0
ncsa1 ^g	149	0.362	6.80	4.61	-67.0	-42.0	0
pnsap ^{a,h}	257	0.381	6.55	7.59	-62.3	-88.1	0
myhu ⁱ	153	0.379	4.99	4.39	-56.6	-42.4	0
myho ^j	153	0.399	4.50	3.97	-49.1	-40.5	0
kyboa ^k	245	0.384	4.44	7.07	-54.0	-82.5	5
akyboa ^l	241	0.390	4.10	6.64	-46.0	-79.5	5
lggt ^m	162	0.438	3.87	3.57	-52.5	-41.9	2
k3hu ⁿ	106	0.330	3.87	2.83	-38.7	-25.5	1
lgboA ^o	162	0.432	3.83	3.66	-53.8	-42.2	2
ccck ^p	109	0.339	3.82	2.94	-42.8	-26.5	1
ccd ^q	104	0.317	3.60	2.75	-43.5	-24.8	1
ccby ^r	108	0.315	3.40	2.92	-34.5	-26.3	1
lahu ^s	123	0.374	3.23	3.40	-33.2	-30.7	4
mybo ^t	153	0.412	3.19	3.76	-38.5	-39.7	0
nrbo ^u	124	0.315	3.10	3.62	-40.6	-32.6	4
ccpg ^v	104	0.308	3.10	2.75	-39.4	-24.8	1
cchod ^w	104	0.308	3.10	2.75	-38.2	-24.8	1
ccrb ^x	104	0.327	3.00	2.75	-36.6	-24.8	1
ccho ^y	104	0.308	2.95	2.75	-35.9	-24.8	1
txec ^z	108	0.444	2.90	2.39	-35.3	-22.7	1
hybstII ^{aa}	111	0.387	2.90	2.82	-33.0	-25.4	0
ccbn ^{bb}	103	0.340	2.80	2.69	-34.8	-24.3	1
mywhp ^{cc}	153	0.425	2.60	3.58	-31.0	-39.1	0
ccch ^{dd}	104	0.308	2.30	2.75	-28.4	-24.8	1
lzch ^{ee}	129	0.357	2.12	3.77	-33.9	-34.0	4
hybstIIS ^{ff}	91	0.429	2.00	1.96	-23.1	-17.7	0
tseca ^{gg}	188	0.463	1.60	3.59	-17.6	-50.3	0
egmsg ^{hh}	53	0.302	1.44	0.89	-45.6	-8.1	3
tseca2 ⁱⁱ	80	0.512	1.20	1.38	-11.3	-12.7	0
tibo ^{jj}	58	0.345	1.1	1.02	-19.7	-9.2	2

^a The theoretical values for Δ*F*_{fold}/Δ*c*_{GuHCl} are calculated at *c*_m from eq 25 and the given *n* and Φ for the protein. The experimental Δ*A* is calculated from the experimental Δ*F*_{fold}/Δ*c*_{GuHCl}. The theoretical value of Δ*A* was calculated by use of eq 20. The column marked ss indicates number(s) of disulfide bonds. The experimental slopes marked with an asterisk (*) showed multiple transitions and were reported as the sum of the slopes of the two transitions. The abbreviations and amino acid composition data were taken from the Dayhoff Protein Data Base: ^b phosphoglycerate kinase (Saito & Wada, 1983); ^c phosphoglycerate kinase (Betton et al., 1984); ^d trp synthetase A (Miles et al., 1982); ^e pepsinogen (Saito & Wada, 1983); ^f carbonic anhydrase (Saito & Wada, 1983); ^g staphylococcal nuclease (Shortle & Meeker, 1986); ^h β-lactamase (Mitchinson & Pain, 1985); ⁱ myoglobin (Pace, 1975); ^j myoglobin (Pace & Vanderburg, 1979); ^k chymotrypsinogen A (Ahmad & Bigelow, 1986); ^l α-chymotrypsin (Greene & Pace, 1974); ^m β-lactoglobulin (Greene & Pace, 1974); ⁿ Ig light chain constant (Tsunegawa et al., 1987); ^o β-lactoglobulin A (Pace, 1975); ^p cyt c (Knapp & Pace, 1974); ^q cyt c (McLendon & Smith, 1978); ^r cyt c iso-1-ms (Ramdas et al., 1986); ^s α-lactalbumin (Nozaka et al., 1978); ^t myoglobin (Pace, 1975); ^u ribonuclease A (Greene & Pace, 1974); ^v cyt c (McLendon & Smith, 1978); ^w cyt c (McLendon & Smith, 1978); ^x cyt c (McLendon & Smith, 1978); ^y cyt c (McLendon & Smith, 1978); ^z thioredoxin (Kelley & Stellwagen, 1984); ^{aa} thermolysin FII (Vita et al., 1985); ^{bb} cyt c (McLendon & Smith, 1978); ^{cc} myoglobin (Pace, 1975); ^{dd} cyt c (McLendon & Smith, 1978); ^{ee} lysozyme (Ahmad & Bigelow, 1982); ^{ff} thermolysin FII (Vita et al., 1985); ^{gg} trp synthetase A FI (Miles et al., 1982); ^{hh} epidermal growth factor (Ahmad & Bigelow, 1986); ⁱⁱ trp synthetase A FII (Miles et al., 1982); ^{jj} pancreatic trypsin inhibitor (rcam 14-38) (Jullien & Baldwin, 1981).

on solvent stability of any given protein. Nevertheless, it is of interest to consider here some observations from single-site mutation experiments. Shortle and Meeker (1986) have observed that while many single-site mutations in staphylococcal nuclease cause very little change in Δ*F*_{fold}/Δ*c*, some can change the slope by as much as 30%, and some triple mutations can change the enthalpy of folding by as much as 35 kcal/mol (Shortle et al., 1988). They have interpreted this as implying that specific interactions can be important in the unfolded state. Such large changes are not predicted from a mean-field model such as the present one, but staphylococcal nuclease is particularly unstable and thus its stability may be more sensitive to small specific interactions than more stable proteins.

With these cautions in mind, nevertheless, the data of Shortle and Meeker (1986) can be used to show that stabilities cannot be linear in both urea and GuHCl. We compare the ratios of experimental denaturation slopes for GuHCl and urea with the ratios predicted by the slopes of the free energies of transfer of amino acid residues using eq 27 for each of the

different mutants investigated by Shortle and Meeker. The ratio of the denaturation slopes of the two denaturants predicted by eq 27 is

$$\frac{(\partial \Delta F_{\text{fold}} / \partial c_{\text{GuHCl}})}{(\partial \Delta F_{\text{fold}} / \partial c_{\text{urea}})} = \frac{\partial g_{\text{tr}} / \partial c_{\text{GuHCl}}}{\partial g_{\text{tr}} / \partial c_{\text{urea}}} \quad (28)$$

assuming that Δ*A*_{Δ*F*_{fold}=0}

 (i.e., Δ*A* at the denaturation point) for a given protein is independent of the denaturant. We substitute the concentration derivatives of eqs 14 and 15 for Δ*g*_{tr}/Δ*c* in the numerator and denominator in the right hand side of eq 28, with the result that

$$\frac{(\partial \Delta F_{\text{fold}} / \partial c_{\text{GuHCl}})}{(\partial \Delta F_{\text{fold}} / \partial c_{\text{urea}})} = \frac{g_1(\text{GuHCl})}{g_1(\text{urea})} + \frac{g_2(\text{GuHCl})}{g_1(\text{urea})} 2c_m \text{GuHCl} = 3.02 - 0.31c_m \text{GuHCl} \quad (29)$$

where *g*₁(GuHCl) and *g*₁(urea) are the coefficients to *c*¹ in eqs 14 and 15, respectively. Thus, the ratio of denaturant slopes for GuHCl and urea should depend on the GuHCl concentration at which the proteins denature. The ratio of denatu-

Table II: Urea Denaturation Slopes and Changes in Exposed Hydrophobic Area ($1.4\Delta A$)^a

protein	n	Φ	denaturation slopes, $\partial\Delta F_{\text{fold}}/\partial c$ (kcal/mol M)		change in exposure, $\partial\Delta F_{\text{fold}}/\partial g_{\text{tr}}$ (residues)		ss
			experiment	theory	experiment	theory	
tsecbat ^{a,b}	268	0.474	3.76	2.99	-102.2	-81.2	0
tseca ^{a,c}	268	0.478	3.38	2.97	-91.8	-80.7	0
pnsap ^d	257	0.381	3.20	3.24	-87.0	-88.1	0
ncsaf ^e	149	0.362	2.36	1.54	-64.1	-42.0	0
lgbo ^f	162	0.432	2.23	1.55	-60.6	-42.2	2
myho ^g	153	0.399	2.14	1.49	-58.2	-40.5	0
lggt ^h	162	0.438	2.10	1.54	-57.1	-41.9	2
akyboa ⁱ	241	0.390	2.07	2.92	-56.2	-79.5	5
kyboa ^j	245	0.384	2.03	3.03	-55.2	-82.5	5
lgboa ^k	162	0.432	1.87	1.55	-50.8	-42.2	2
mywhp ^l	153	0.425	1.46	1.44	-39.7	-39.1	0
lzch ^m	129	0.354	1.29	1.27	-35.1	-34.5	4
nrbo ⁿ	124	0.315	1.10	1.20	-29.9	-32.6	4
hybstII ^o	111	0.387	1.05	0.93	-28.5	-25.4	0
rpecl ^p	51	0.471	0.47	0.23	-12.8	-6.3	0

^aThe theoretical values for $\partial\Delta F_{\text{fold}}/\partial c_{\text{urea}}$ are calculated at c_m from eq 25 and the given n and Φ for the protein. The experimental ΔA is calculated from the experimental $\partial\Delta F_{\text{fold}}/\partial c_{\text{urea}}$. The theoretical value of ΔA was calculated by use of eq 20. The column marked ss indicates number(s) of disulfide bonds. The experimental slopes marked with an asterisk indicate multiple transitions and were reported as the sum of the slopes of the two transitions. The abbreviations and amino acid composition data were taken from the Dayhoff Protein Data Base: ^btrp synthetase A (Stackhouse et al., 1986); ^ctrp synthetase A (Stackhouse et al., 1986); ^d β -lactamase (Mitchinson & Pain, 1985); ^estaphylococcal nuclease (Shortle & Meeker, 1986); ^f β -lactoglobulin B (Cupo & Pace, 1983); ^gmyoglobin (Pace, 1975); ^h β -lactoglobulin (Greene & Pace, 1974); ⁱ α -chymotrypsin (Greene & Pace, 1974); ^jchymotrypsinogen A (Ahmad & Bigelow, 1986); ^k β -lactoglobulin A (Pace, 1975); ^lmyoglobin (Pace, 1975); ^mlysozyme (Ahmad & Bigelow, 1982); ⁿribonuclease A (Greene & Pace, 1974); ^othermolysin fII (Vita et al., 1985); ^plac repressor (Schnarr & Maurizot, 1981).

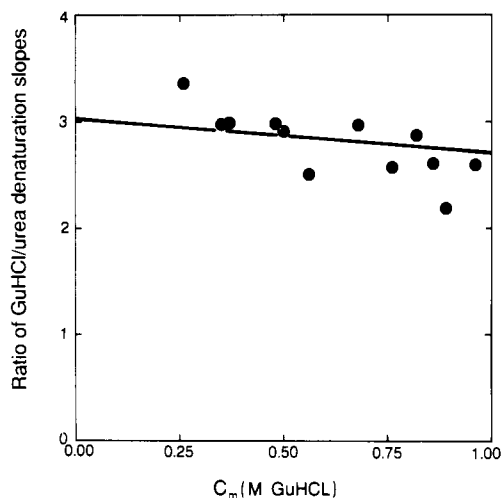


FIGURE 11: Ratio of experimental GuHCl/urea denaturation slopes $((\partial\Delta F_{\text{fold}}/\partial c_{\text{GuHCl}})/(\partial\Delta F_{\text{fold}}/\partial c_{\text{urea}}))$ plotted vs experimental c_m GuHCl for staphylococcal nuclease mutants (Shortle & Meeker, 1986).

ration slopes plotted against c_m GuHCl should give a straight line with slope equal to $2g(\text{GuHCl})_1/g(\text{urea})_1$ and intercept equal to $g(\text{GuHCl})_1/g(\text{urea})_1$. Figure 11 shows good agreement between the experimental ratio (points) for Shortle and Meeker's staphylococcal nuclease mutants and the solid line predicted by eq 29. That the ratios of the denaturation slopes depend on the concentration at which the proteins denature suggests that there is curvature of $\Delta F_{\text{fold}}(c)$ for either GuHCl or urea. The theory predicts that it is $\Delta F_{\text{fold}}(c_{\text{GuHCl}})$ that is curved. Figure 11 should not be viewed as a test of the theory, since eq 28 is independent of the theoretical prediction of ΔA . This treatment relies on the assumption that the denaturation slope is proportional to the change in exposed area multiplied by the slope of the concentration dependence of the free energy of transfer. It should be noted that for eq 28 to hold the denaturation slopes of urea and GuHCl must be measured under otherwise identical conditions. Changes in pH and salt concentration can strongly affect the solvent denaturation profiles of proteins (Pace & Grimsley, 1988; Pace et al., 1990), and they may change ΔA and the compactness of the unfolded state.

Comparison of Thermal and Solvent Denaturation. Are the denatured states that arise from heat, on the one hand, or denaturants, on the other hand, different (Pfeil, 1987; Privalov, 1979; Tanford, 1968)? The present model predicts that they are not essentially different if both types of experiments are performed near the isoelectric pH. It is useful to distinguish three sets of conditions: (i) denaturation at low pH by heating to intermediate temperatures, (ii) denaturation at neutral pH by high temperatures, and (iii) actions of denaturants. What has been referred to as "heat" or "thermal" denaturation in the past has generally been (i) at low pH. Neutral pH thermal denaturations (ii) are less common in order to avoid irreversible changes, including covalent modification (Tanford, 1968). Nevertheless, in our view, the more important characteristic of denaturation of type (i) is not that it is thermal but that it is done under acid conditions. We consider the effects of pH and ionic strength elsewhere (Stigter & Dill, 1989, 1990; Stigter et al., 1991).

The present theory predicts that the ensemble of conformations resulting from thermal denaturation (at neutral pH (ii)) or from denaturants (iii) should be very similar in the following respects. First, both are removed from the native state by a free energy barrier; i.e., the native-denatured transition in both cases should be first order. Second, the average radius at the midpoint of denaturation should be the same irrespective of whether denaturation is by a thermal process or by urea or GuHCl. This prediction of the theory arises because the single quantity that determines the denaturation midpoint is χ , characterizing the transfer free energy of the nonpolar monomers from water, irrespective of whether its particular value is determined by the temperature or the chemistry of the solvent. Third, the volume of the denatured molecules at the denaturation midpoint depends on the fractional hydrophobicity of the protein, Φ , and the volume is predicted to be much smaller than for random-flight chains for proteins of high Φ . The temperatures or denaturant concentrations required to denature globular proteins are considerably less than those required to achieve the Flory θ conditions that specify random-flight behavior. Fourth, denatured states should continue to expand in volume with further increase in temperature or denaturant above the de-

naturation midpoint, and this expansion will be more pronounced for proteins of high Φ . Therefore, for both types of denaturation, a highly open denatured state can ultimately be achieved by further increase in denaturing agent. This change is predicted to be gradual (not cooperative) and to involve no further transition (i.e., no free energy barrier). It is noted elsewhere, however, that multiple transitions between compact denatured states and highly unfolded states are predicted to occur when there are also electrostatic driving forces, as at the extremes of pH for proteins denatured in acids and bases (Stigter et al., 1991). Of course, beyond the midpoint of the transition, there is no direct comparison possible between a denatured state at some temperature and that at some denaturant concentration, because there is no simple interconversion of temperature and solvent concentration. If we draw a correspondence through use of the χ value, then we find that χ at 90 °C is approximately equal to χ in 3.6 M GuHCl at 25 °C, for neutral pH in both cases. The experiments of Pfeil (1987) show the similarity between temperature- and solvent-denatured α -lactalbumin compared to the acid-induced unfolded state.

Finally, the theory is consistent with the observation (Pfeil, 1987; Privalov, 1979) that the enthalpy change between the highly unfolded and compact denatured states is small. For typical proteins, some 50–75% of the total nonpolar exposure (to other nonpolar groups and solvent) is predicted to be lost upon randomizing of the hydrophobic core with only negligible decrease in chain segment density, ρ^* . This represents the change from the native state to the random compact state. It is also predicted, however, that the change in exposure to only the solvent (rather than to the solvent plus polar groups) upon going from native to the random compact state will be much smaller than this (only about 20%).

Shortle et al. (Shortle et al., 1988) have noticed an interesting relationship between a property of thermal denaturation and a property of solution denaturation. They observe that the change in heat capacity upon unfolding is linearly related to the slope $\partial\Delta F_{\text{fold}}/\partial c$ for staphylococcal nuclease mutants. The existence of such a relationship can be rationalized as follows, irrespective of whether staphylococcal nuclease is representative of more stable proteins or not. The heat capacity change upon folding, $\Delta C_{p,\text{fold}}$, is given by (Dill et al., 1989)

$$\frac{\Delta C_{p,\text{fold}}}{1.4} = \Delta C_{p,\text{transfer}} \Delta A + \Delta H_{\text{transfer}}(T) \partial \Delta A / \partial T \quad (30)$$

At 25 °C, $\Delta H_{\text{transfer}}(T) = 0$; hence, eqs 30 and 27 give

$$\frac{\partial \Delta F_{\text{fold}} / \partial c_{\text{urea}}}{\Delta C_{p,\text{fold}}} = \frac{\partial g_{\text{tr}} / \partial c_{\text{urea}}}{\Delta C_{p,\text{transfer}}} = \frac{-36.8 \text{ cal}/(\text{mol M})}{55 \text{ cal}/(\text{mol K})} = -0.67 \text{ K/M} \quad (31)$$

where the value in the numerator is taken from the coefficient of the linear term in eq (15) and in the denominator is taken from Dill et al. (1989). Hence, this ratio is predicted to be simply a property of the denaturant and the transfer thermal properties of an average hydrophobic residue in the protein. According to eq 31, $\partial\Delta F_{\text{fold}}/\partial c$ plotted against $\Delta C_{p,\text{transfer}}$ should be linear with slope equal to -0.67 K/M and intercept equal to $0 \text{ cal}/(\text{mol M})$. Linear-regression best-fits of the data of Shortle et al. (Shortle & Meeker, 1986; Shortle et al., 1988) give a slope = -0.91 K/M , an intercept = $1.14 \text{ cal}/(\text{mol M})$, and a correlation coefficient = 0.90. The point of this comparison is to illustrate that there is experimental evidence for a relation between $\Delta C_{p,\text{fold}}$ and $\partial\Delta F_{\text{fold}}/\partial c$ and that there is a simple relation predicted by the theory. The quantitative

comparison should be viewed in light of the ambiguity in the representative $g_{\text{tr}}(c)$ and $\Delta C_{p,\text{tr}}$ as well as the uncertainties of the experimental van't Hoff determinations. In contrast to this data, when the data for $\Delta C_{p,\text{fold}}$ determined calorimetrically for other proteins (Privalov & Khechinashvili, 1974) are included, there appears to be little correlation between $\partial\Delta F_{\text{fold}}/\partial c$ and $\Delta C_{p,\text{fold}}$. In this larger set of proteins, where the solvent and calorimetric experiments were done under differing conditions (particularly pH and salt concentration), this relation is not expected to hold.

CONCLUSIONS

We have developed a statistical mechanical theory for the effects of denaturing and stabilizing solvents on protein stability. The three principal properties we calculate are the free energy of folding, the radius of the unfolded molecule, and the exposure of nonpolar residues to solvent or to other nonpolar groups. These properties depend on the length of the protein chain, the composition (i.e., the fraction of residues that are nonpolar), and the elementary free energy of transfer of a nonpolar residue into a nonpolar medium from the solution of solvent plus denaturant. The latter quantity is incorporated through a parameterization of the experimental data of Nozaki and Tanford (1963, 1970, 1971). The model makes several predictions. First, increasing concentrations of denaturants can cause sharp phase transitions of the native molecule to denatured states because the denaturant solutions favorably solvate the nonpolar groups in the denatured states. Guanidine hydrochloride should be an effective denaturant at lower concentrations than urea. Second, the model predicts that the free energy should depend approximately linearly on denaturant concentration for either GuHCl or urea over the typical experimentally accessible ranges but that there should be some small curvature for GuHCl over larger concentration ranges. The predicted slopes are found to be in good agreement with experimental values. Third, the model predicts that the ensemble-averaged radius of unfolded molecules and the corresponding nonpolar exposure should depend significantly on denaturant concentration and on protein properties such as the chain length and composition. We find that there is an important difference between two types of nonpolar exposure: (i) exposure to polar residues plus solvent, which dominates the thermodynamics, and (ii) exposure to solvent alone, which is accessible in some spectroscopic experiments. It is predicted that at neutral pH solvent denaturation and thermal denaturation should lead to the same relatively compact denatured conformations and that the denatured molecules will expand to more open configurations with further increases in denaturant or temperature. These changes in the unfolded state are predicted to be noncooperative, in contrast to states of denaturation caused by acids and bases.

ACKNOWLEDGMENTS

We thank David Shortle, Wayne Bolen, Marcelo Santoro, R. L. Baldwin, Ron Siegel, Gregg Fields, and Sarina Bromberg for helpful comments. We are grateful to Mark Hurle for providing us with his extensive compilation of experimental denaturation slopes.

Registry No. Urea, 57-13-6; guanidine hydrochloride, 50-01-1.

REFERENCES

- Ahmad, F., & Bigelow, C. C. (1982) *J. Biol. Chem.* **157**, 12935–12938.
- Ahmad, F., & Bigelow, C. C. (1986) *Biopolymers* **25**, 1623–1633.
- Arakawa (1982) *Biochemistry* **21**, 6536.

- Back (1979) *Biochemistry* 18, 5191.
- Betton, J. M., Desmadril, M., Mitraki, A., & Yon, J. M. (1984) *Biochemistry* 23, 6654-6661.
- Bolen, D. W., & Santoro, M. M. (1988a) *Biochemistry* 27, 8063-8068.
- Bolen, D. W., & Santoro, M. M. (1988b) *Biochemistry* 27, 8069-8074.
- Creighton, T. E. (1979) *J. Mol. Biol.* 129, 235-264.
- Cupo, J. F., & Pace, C. N. (1983) *Biochemistry* 22, 2654-2658.
- Dayhoff Protein Data Base, Protein Identification Resource (1990) National Biomedical Research Foundation, Washington, DC, 20007.
- Dill, K. A. (1985) *Biochemistry* 24, 1501-1509.
- Dill, K. A. (1987) *J. Phys. Chem.* 91, 1980-1988.
- Dill, K. A., & Alonso, D. O. V. (1988) in *Protein Structure and Protein Engineering* (Huber, T., & Winnacker, E. L., Eds.) Colloquium-Mosbach, Vol. 39, pp 51-58, Springer-Verlag, Berlin.
- Dill, K. A., Alonso, D. O. V., & Hutchinson, K. A. (1989) *Biochemistry* 28, 5439-5449.
- Gekko & Timasheff (1981) *Biochemistry* 20, 4667.
- Greene, R. F., & Pace, C. N. (1974) *J. Biol. Chem.* 249, 5388-5393.
- Hildebrand, J. H., & Scott, R. L. (1950) *The Solubility of Nonelectrolytes*, ACS Monograph Series, Vol. 17, Reinhold Publishing Corp.
- Jullien, M., & Baldwin, R. L. (1981) *J. Mol. Biol.* 145, 265-280.
- Kelley, R. F., & Stellwagen, E. (1984) *Biochemistry* 23, 5095-5102.
- Knapp, J. A., & Pace, C. N. (1974) *Biochemistry* 13, 1289-1294.
- Lee, B., & Richards, F. M. (1971) *J. Mol. Biol.* 55, 379-400.
- Lee, J. C., & Timasheff, S. N. (1981) *J. Biol. Chem.* 256, 7193-7201.
- Lee, L. L. Y., & Lee, J. C. (1981) *J. Biol. Chem.* 256, 625-631.
- McLendon, G., & Smith, M. (1978) *J. Biol. Chem.* 253, 4004-4008.
- McMeekin, T. L., Cohn, E. J., & Weare, J. H. (1936) *J. Am. Chem. Soc.* 58, 2173-2181.
- Miles, E. W., Uutani, K., & Ogasahara, K. (1982) *Biochemistry* 21, 2586-2592.
- Mitchinson, C., & Pain, R. H. (1985) *J. Mol. Biol.* 184, 331-342.
- Nozaki, M., Kuwajima, K., Nitta, K., & Sugai, S. (1978) *Biochemistry* 17, 3753-3758.
- Nozaki, Y., & Tanford, C. (1963) *J. Biol. Chem.* 238, 4074-4081.
- Nozaki, Y., & Tanford, C. (1970) *J. Biol. Chem.* 245, 1648-1652.
- Nozaki, Y., & Tanford, C. (1971) *J. Biol. Chem.* 246, 2211-2217.
- Pace, C. N. (1975) *CRC Crit. Rev. Biochem.* 3, 1-43.
- Pace, C. N., & Grimsley, G. R. (1988) *Biochemistry* 27, 3242-3246.
- Pace, C. N., & Vanderburg, K. E. (1979) *Biochemistry* 18, 288-291.
- Pace, C. N., Laurents, D. V., & Thomson, J. A. (1990) *Biochemistry* 29, 2564-2572.
- Pfeil, W. (1987) *Biochim. Biophys. Acta* 911, 114-116.
- Prasad, K. P., & Ahluwalia, J. C. (1976) *J. Solution Chem.* 5, 491-507.
- Privalov, P. L. (1979) *Adv. Protein Chem.* 33, 167-241.
- Privalov, P. L., & Khechinashvili, N. N. (1974) *J. Mol. Biol.* 86, 665-684.
- Puett, D. (1973) *J. Biol. Chem.* 248, 4623-4634.
- Ramdas, L., Sherman, F., & Nall, B. T. (1986) *Biochemistry* 25, 6952-6958.
- Saito, Y., & Wada, A. (1983) *Biopolymers* 22, 2123.
- Schellman, J. A. (1975) *Biopolymers* 14, 999-1018.
- Schellman, J. A. (1978) *Biopolymers* 17, 1305-1322.
- Schellman, J. A. (1987) *Biopolymers* 26, 549-559.
- Schnarr, M., & Maurizot, J. C. (1981) *Biochemistry* 20, 6164-6169.
- Shortle, D., & Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet.* 1, 81-89.
- Shortle, D., Meeker, A. K., & Freire, E. (1988) *Biochemistry* 27, 4761-4768.
- Shortle, D., Meeker, A. K., & Gerring, S. L. (1989) *Arch. Biochem. Biophys.* 272, 103-113.
- Spink, C. H., & Wadso, I. (1975) *J. Chem. Thermodyn.* 7, 561-572.
- Stackhouse, T., Onuffer, J. J., Mathews, C. R., Ahmed, S. A., & Miles, E. W. (1988) *Biochemistry* 27, 824-832.
- Stigter, D., & Dill, K. A. (1989) *J. Phys. Chem.* 93, 6737-6743.
- Stigter, D., & Dill, K. A. (1990) *Biochemistry* 29, 1262-1271.
- Stigter, D., Alonso, D. O. V., & Dill, K. A. (1991) *Proc. Natl. Acad. Sci. U.S.A.* (in press).
- Tanford, C. (1968) *Adv. Protein Chem.* 23, 122-217.
- Tanford, C. (1970) *Adv. Protein Chem.* 24, 1-95.
- Tanford, C. (1980) *The Hydrophobic Effect*, John Wiley & Sons, New York.
- Thomson, J. A., & Bigelow, C. C. (1986) *Biochem. Cell Biol.* 64, 993-998.
- Tsunegawa, M., Goto, Y., Kawata, Y., & Hamaguchi, K. (1987) *Biochemistry* 26, 6044-6051.
- Vita, C., Fontana, A., & Chaiken, I. M. (1985) *Eur. J. Biochem.* 151, 191-196.
- Wetlaufer, D. B., Malik, S. K., Stoller, L., & Coffin, R. L. (1964) *J. Am. Chem. Soc.* 86, 508-514.