Theoretic Information Approach to Protein Stabilization by Solvent Engineering

Roberto Di Domenico, Roberto Lavecchia, and Alessandro Ottavi
Dept. of Chemical Engineering, University "La Sapienza," I-00184 Roma, Italy

A novel approach based on molecular thermodynamics and the information theory is proposed to quantify the influence of water-miscible additives on protein stability. According to the two-state mechanism of inactivation, solvent effects are described in terms of perturbation of the equilibrium between the folded and unfolded protein forms. The model provides the dependence of the protein's melting temperature on the additive concentration. Effects of the latter are accounted for by an empirical parameter related to the free energy of transfer of the protein from the pure to the mixed solvent. The model was tested using experimental data relative to the influence of hydroxylic and aminoacidic additives on the thermal unfolding of hen egg lysozyme and erythrocyte carbonic anhydrase. Fitting parameters were correlated in terms of a theoretic information index characterizing the additive's molecule and incorporating an atomic-composition term and a topological contribution. Model calculations agreed very well with experimental data, suggesting that the molecular information content of the additive can be used effectively to correlate solvent-induced perturbations of stability. The procedure was also used to predict melting temperatures in systems containing binary mixtures of additives and to reconstruct thermal unfolding curves in the different media.

Introduction

The three-dimensional (3-D) structure resulting from the *in vivo* folding of the polypeptide chain provides the unique arrangement and orientation of functional groups that give the protein its biological activity. The folded conformation relies on a delicate balance of intramolecular forces, primarily hydrogen and van der Waals bondings, and protein–solvent interactions (Privalov and Gill, 1988; Dill, 1990; Pace et al., 1996). Analysis of the available experimental data shows that the Gibbs energy difference between folded and unfolded states is of the order of 40 kJ mol^{-1}, less than $0.1kT$ per residue, where k is the Boltzmann's constant and T is the temperature (Pace, 1975, 1990; Privalov, 1979). Native proteins are thus only marginally stable, and destined to denature once released from their natural environment. This intrinsic lability dramatically manifests itself in industrial operations involving protein extraction and purification, or during storage (Asenjo and Patrick, 1990). It also affects all existing enzymatic industrial processes where, as pointed out by Klibanov (1983), the enzymes are used not because of their high stability but despite their large instability. The problem therefore arises of how to stabilize a protein in order to preserve a high degree of biological activity in an artificial, nonnative environment.

Different approaches have been proposed to hindering denaturation. They include isolation of more stable protein forms from extremophiles, chemical or genetic modifications of the macromolecule, and use of stabilizing additives (Volkin and Klibanov, 1989; Tomazic, 1991). None of them is free of drawbacks, but in many instances addition of a water-miscible component to the protein medium is the simplest and most effective means for increasing stability (Timasheff and Arakawa, 1989; Cioci and Lavecchia, 1998).

Several substances belonging to the classes of polyols, inorganic salts, amino acids and methylamines have been proved capable of stabilizing proteins (Arakawa and Timasheff, 1983; Fernandez et al., 1991; Lozano et al., 1994; Cioci and Lavecchia, 1994, 1999; Cioci, 1995a,b; Cioci et al., 1996; Matsumoto et al., 1997). Preferential hydration phenomena pro-

Correspondence concerning this article should be addressed to R. Lavecchia.

vide a basis for a unifying interpretation of their influence. In particular, the action of all known stabilizers appears to be a reflection of their ability to be preferentially excluded from the protein domain, that is, to concentrate themselves into the bulk solvent when added to the protein solution (Arakawa and Timasheff, 1982, 1983). The connection between exclusion and stabilization is to be found in the fact that preferentially excluded additives increase the chemical potential of the protein to an extent that is proportional to the protein–solvent contact surface (Wyman and Gill, 1990; Timasheff, 1992). Since denatured proteins have a much greater surface area exposed to solvent than do native proteins, a preferentially excluded component would give rise to a thermodynamically unfavorable situation, causing the displacement of the unfolding equilibrium toward the native protein. The detailed mechanisms of exclusion are not fully established, however, particularly when specific additive–protein interactions are involved (Arakawa et al., 1990; Bhat and Timasheff, 1992). Under these conditions the surface properties of the macromolecule could neutralize the action of common stabilizers or shift it toward destabilization (Volkin and Klibanov, 1989; Cioci et al., 1994; Cioci and Lavecchia, 1997a). On the other hand, the presence of an effective stabilizer could depress the protein's conformational flexibility and activity in such a way as to negate its beneficial influence on stability (Laane et al., 1987; Russell and Vierheller, 1995).

Because of the complexity of the underlying phenomena, attempts to model solvent effects from a quantitative point of view have met only limited success (Schellman, 1990, 1994; Cioci and Lavecchia, 1997b, 1998). As a result, a considerable experimental effort is required to establish which additives should be appropriate to each particular protein.

This article is intended to present a novel approach to the problem in question, based on the concept of information content of a molecular structure. In particular, we explore the possibility of interpreting the solvent-induced changes in stability in terms of some molecular features of the added components. We concentrated on two globular proteins, carbonic anhydrase and lysozyme, differing in size and stability. Literature data were first analyzed concerning the influence of hydroxylic components on the two proteins. To assess the potential of the proposed approach on a broader basis, experiments were also performed using a family of glycine-based amino-acid derivatives as stabilizing components. The molecular structure of each additive was characterized by a theoretical information index incorporating an atomic-composition term and a topological contribution.

Experimental Studies

Materials

Carbonic anhydrase (EC 4.2.1.1) from bovine erythrocytes was purchased from Sigma Chemical Co. as a dialyzed and lyophilized powder. The claimed activity was 2500 Wilbur-Anderson U/mg, where 1 U corresponds to the amount of protein causing the pH of a 20-mM Trizma buffer to drop from 8.3 to 6.3 per minute at 0°C. Prior to use, the protein was dissolved in 0.1-M phosphate buffer ($KH_2PO_4 + Na_2HPO_4$) at pH 7. Glycine, *N*-methylglycine (sarcosine), *N*,*N*-dimethylglycine and *N*,*N*,*N*-trimethylglycine (betaine) were obtained from Fluka, with a purity $>98\%$ for betaine, and $>99\%$ for the others. All other chemicals were of reagent grade and used without further purification.

Methods

Thermal unfolding of carbonic anhydrase was followed by ultraviolet difference spectroscopy. Optical measurements were performed by a double-beam UV-VIS spectrophotometer (Perkin Elmer; Lambda 5) equipped with a water-jacketed cell holder. The cell holder was connected to an external thermostatic bath controlled by a digital temperature programmer (Haake, PG 41). To calibrate the heating system a high-precision electrical resistance microprobe was inserted into the cell filled with bidistilled water.

Protein solution at a concentration of 0.5 mg mL^{-1} were prepared gravimetrically by dissolving the lyophilized powder in buffer (with or without the additives). The solution was passed through a 0.22-μm Millipore filter to remove dust and improve the signal-to-noise ratio. About 1 mL of the same protein solution was then poured into two 1-cm matched quartz cuvettes, which were used as the reference and the sample cell. The former was thermostated at 25 ± 0.05°C, while the second was heated according to a specified thermal program. Most runs were made by continuously heating the solution at a rate of 0.2°C min^{-1}; lower heating rates or different time–temperature profiles did not produce significant differences in the results obtained.

On increasing temperature, carbonic anhydrase was subjected to unfolding, and hence to the transfer of some light-absorbing side chains of hydrophobic residues from the protein interior to the external solvent. This phenomenon gives rise to the blue shift of the UV spectrum of the protein, and to the appearance of a number of distinctive features in the 280–320-nm region of the difference spectrum. Melting curves were determined at constant wavelength (291 or 300 nm) by recording the first derivative of the difference adsorption spectrum against temperature, according to the procedure of Pace et al. (1989). In all experiments, no change was observed in the pH of the protein solution during denaturation.

The reversibility of unfolding was carefully checked by reconditioning, at 25°C, some previously denatured protein solutions. Spectrophotometric analysis of these solutions showed that their absorbance was nearly identical to that of native carbonic anhydrase.

Molecular Thermodynamic Framework

The denaturation of a globular protein can be described, to a first approximation, as a cooperative equilibrium transition between two macroscopical states, the native (N) folded state and the denatured (D) unfolded state (Tanford, 1968, 1970; Pfeil and Privalov, 1979; Lavecchia and Zugaro, 1991): $N \leftrightarrow D$.

The Gibbs energy difference between the N and D states,

$$\Delta G_{\text{unf}} = G_D - G_N, \tag{1}$$

represents the work needed to disrupt the native, biologically active conformation, and can therefore be regarded as a measure of stability (Becktel and Schellman, 1987; Creighton, 1991).

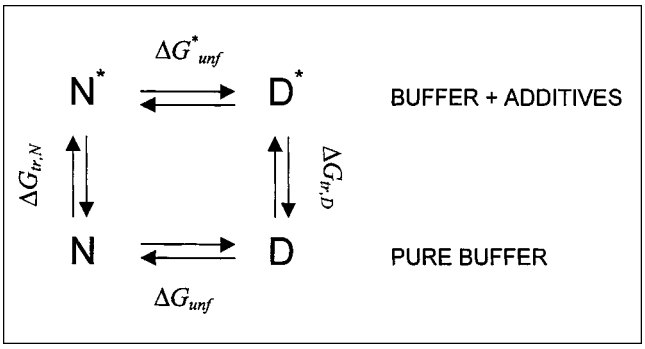


Figure 1. Thermodynamic cycle for calculation of the Gibbs energy of unfolding in the perturbed solvent.

As with any thermodynamic quantity, protein stability depends on the intensive state properties temperature, pressure, and composition. Usually the most pertinent composition variables are pH, ionic strength, and the concentration of any solvent component that preferentially binds the macromolecule in either conformational state. Here we are primarily interested in analyzing the influence of water-miscible additives on the protein's thermal stability. To this aim we consider the hypothetical thermodynamic cycle depicted in Figure 1, where the asterisks mark the perturbed states, that is, the conformational states assumed by the protein in the presence of the additive(s). This cycle allows the Gibbs energy change in the mixed solvent ($\Delta G^*_{\rm unf}$) to be expressed in terms of the same quantity in the pure solvent as

$$\Delta G^*_{\rm unf} = \Delta G_{\rm unf} + \Delta G_{{\rm tr},D} - \Delta G_{{\rm tr},N}, \tag{2}$$

where $\Delta G_{{\rm tr},N}$ and $\Delta G_{{\rm tr},D}$ are the free energies of transfer of the N and D forms from pure solvent to the solution at a specific additive concentration. The last two terms on the righthand side (r.h.s.) of Eq. 2 can be lumped into the following quantity

$$\Delta G_{\rm tr} = \Delta G_{{\rm tr},D} - \Delta G_{{\rm tr},N}, \tag{3}$$

so as to obtain

$$\Delta G^*_{\rm unf} = \Delta G_{\rm unf} + \Delta G_{\rm tr} \tag{4}$$

where $\Delta G_{\rm tr}$ can be assumed to depend only on the functional groups of the protein that are exposed to the solvent in the denatured state, since the contribution of those groups that are exposed to solvent in both states cancel out (Greene and Pace, 1974; Schellman, 1978, 1987). For some additives, such as urea and guanidine hydrochloride, $\Delta G_{\rm tr}$ can be calculated from existing model compound data, provided that the structure of the folded protein is known (Pace, 1975). In all other cases empirical models must be used.

The influence of temperature on the distribution of the protein between the N and D states can be evaluated from the van't Hoff equation

$$\left(\frac{\partial \ln K_{\rm unf}}{\partial T}\right)_P = \frac{\Delta H^0_{\rm unf}(T)}{RT^2}, \tag{5}$$

where $K_{\rm unf}$ is the equilibrium constant on unfolding. For a two-state transition it can be written as

$$K_{\rm unf} = \frac{x_{\rm unf}}{1 - x_{\rm unf}}, \tag{6}$$

where $x_{\rm unf}$ is the fraction of unfolded protein molecules. We note that this quantity can be easily monitored by standard experimental techniques (Schmid, 1989).

Because of the high cooperativity of unfolding, $\Delta H^0_{\rm unf}$ can be assumed to be constant over the temperature range considered, and integration of Eq. 5 between a reference temperature ($T_{\rm ref}$) and T yields

$$\ln \frac{K_{\rm unf}(T)}{K_{\rm unf}(T_{\rm ref})} = \frac{\Delta H^0_{\rm unf}}{R}\left(\frac{1}{T_{\rm ref}} - \frac{1}{T}\right). \tag{7}$$

It may be convenient to take the melting temperature of the protein (T_m) as the reference temperature. Temperature T_m can be directly determined from $x_{\rm unf}$ data, being the temperature at which the concentrations of the N and D species are equal ($x_{\rm unf} = 0.5$). In addition, it is itself an index of stability, in that the higher the T_m, the higher the protein's resistance to unfolding.

At the melting temperature, the equilibrium constant is unitary and Eq. 7 becomes

$$\ln K_{\rm unf}(T) = \frac{\Delta H^0_{\rm unf}}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right). \tag{8}$$

If we consider the equilibrium relation, $\Delta G_{\rm unf} = -RT \ln K_{\rm unf}$, combining Eqs. 4 and 8 leads to the following expression for the equilibrium constant in the perturbed solvent

$$\ln K^*_{\rm unf}(T) = \frac{\Delta H^0_{\rm unf}}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) - \frac{\Delta G_{\rm tr}}{RT}. \tag{9}$$

Again, at T^*_m the equilibrium constant is unitary, and we get

$$T^*_m = T_m\left(1 + \frac{\Delta G_{\rm tr}}{\Delta H^0_{\rm unf}}\right). \tag{10}$$

Equation 10 provides the link between the perturbed and unperturbed melting temperatures. To proceed further, we have to make explicit the dependence of $\Delta G_{\rm tr}$ on solvent composition. The simplest expression satisfying the condition $\Delta G_{\rm tr} \to 0$ as the additive concentration goes to zero is

$$\Delta G_{\rm tr} = \alpha c, \tag{11}$$

where c is the additive concentration and α is a parameter characterizing the additive-protein system. Schellman (1978, 1987) and Dill (1985) have shown that, under proper conditions, such an assumption can reasonably be based on theoretical grounds. Experimental supports have also been reported (Santoro and Bolen, 1988, 1992; Pace and Laurents,

1989; Shortle et al., 1989). Substitution of Eq. 11 into Eq. 10 gives

$$T_m^* = T_m\left(1 + \frac{\alpha c}{\Delta H_{\text{unf}}^0}\right). \tag{12}$$

According to Eq. 12, the dependence of T_m^* on c can be described in terms of three quantities: T_m, ΔH_{unf}^0 and α, the first two of which characterize the unfolding behavior of the protein in the pure solvent. As a result, experiments performed in the absence of additives can be used to estimate T_m and ΔH_{unf}^0, whereas the third parameter can be obtained by fitting Eq. 12 to thermal unfolding data in the perturbed solvent. Since $\Delta H_{\text{unf}}^0 > 0$, positive values of α are indicative of a stabilizing additive, while negative values point out for a destabilizing component.

Characterization of Additives by Theoretical Information Indices

The application of information theory to molecular structures provides a powerful means for expressing the degree of complexity of a molecular structure in a quantitative form (Kier and Hall, 1976; Bonchev, 1983). This approach considers a molecule as a collection of a finite number (n) of elements that are partitioned, according to a specified equivalence criterion, into k subsets of equivalence $\{n_1, n_2, \ldots, n_k\}$. Association of a probability distribution $\{p_1, p_2, \ldots, p_k\}$ with this partitioning allows evaluation of the quantity

$$I = -\sum_{j=1}^{k} p_j \log_2 p_j, \tag{13}$$

where $p_j = n_j/n$ is the probability for a randomly chosen element to belong to the subset j having n_j elements. The term I has the dimensions of bits/element and can be regarded as the average amount of information contained in each element of the structure.

Calculating the information content of a molecular structure by the preceding approach therefore requires: (a) identification of the elements of the structure, and (b) selection of the partitioning criterion. The number of indices describing a molecule may be very large, due to the variety of elements and partitioning criteria that can be taken into consideration. We concentrated on two of them: the information index on atomic composition, and a topological index based on the vertex distance matrix.

Information index on atomic composition

The information index based on atomic composition was introduced by Dancoff and Quastler (1953). The subset of equivalence is represented by the kind of the atoms composing the molecule, and the associated probabilities are their relative contents in the molecule. If a_j denotes the number of atoms of the kind j present in the formula of the chemical compound, the corresponding probability distribution is given by

$$p_j = \frac{a_j}{\sum_{j=1}^{k} a_j}. \tag{14}$$

As an example, in the case of glucose ($C_6H_{12}O_6$), the following element partition and probability distribution are obtained

$$n = \{6, 12, 6\} \qquad p = \{0.25, 0.5, 0.25\}.$$

The average information content calculated by Eq. 13 is

$$I = 1.5 \text{ bits/element}.$$

Information index based on the vertex distance matrix

The information index based on the vertex distance matrix belongs to the class of topological indices, that is, quantities reflecting the properties of the molecular graph associated with the chemical structure under consideration. A mathematical graph can be defined as a finite nonempty set of vertices and edges, the latter being unordered pairs of distinct vertices (Wilson, 1972). In the case of a chemical structure, vertices are represented by atoms and edges are represented by valence bonds. When dealing with organic compounds, hydrogen-depleted, or skeletal, graphs are generally considered. The way by which vertices and edges are partitioned identifies the corresponding information index.

Bonchev and Trinajstic (1977, 1978) introduced a number of topological indices based on the vertex distance matrix of a graph. This is a symmetrical $N \times N$ matrix, where N is the number of vertices. The generic element of the matrix is called distance (d_{lm}) and corresponds to the shortest path connecting the l-m pair of vertices. Thus, $d_{lm} = 1, 2, \ldots, d_{\max}$ and, by definition, diagonal elements are all zeros ($d_{ll} = 0$). Partitioning can be made by considering the times a given distance value appears in the matrix. If the distance of a value d appears $2m_j$ times in the distance matrix, the N^2 elements d_{lm} can be partitioned into $d_{\max} + 1$ groups, where $d_{\max}$ is the largest value of d, and 1 denotes the group containing the N zeros of the diagonal elements. The probability distribution associated with this partitioning is

$$p_j = \frac{2m_j}{N^2}. \tag{15}$$

Again, application of Eq. 13 allows calculation of the average information content. Since the distance matrix is symmetric, either of the two triangular submatrices contains all the topological information on the structure in question.

Referring to glucose, whose molecular graph and distance matrix are depicted in Figure 2, the following results are obtained

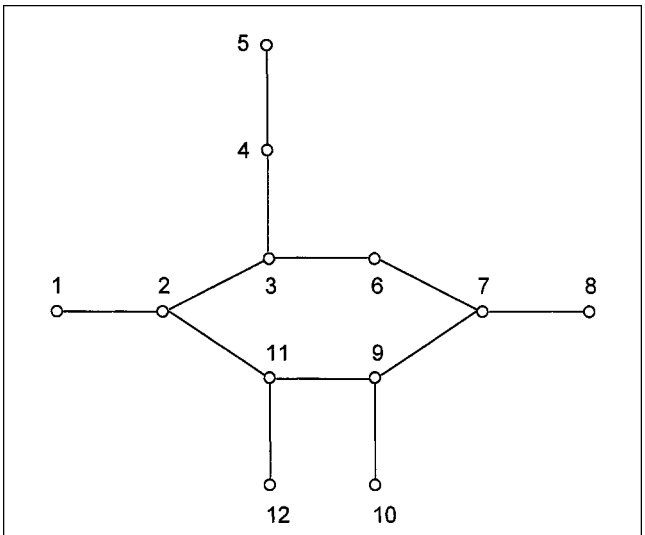
$$d_{\max} = 6 \qquad N = 12$$

$$n = \{12, 24, 34, 38, 24, 10, 2\}$$

$$p = \{0.083, 0.167, 0.236, 0.264, 0.167, 0.069, 0.014\},$$

and Eq. 13 yields

$$I = 2.512 \text{ bits/element}.$$



VERTEX DISTANCE MATRIX

0	1	2	3	4	3	4	5	3	4	2	3
1	0	1	2	3	2	3	4	2	3	1	2
2	1	0	1	2	1	2	3	3	4	2	3
3	2	1	0	1	2	3	4	4	5	3	4
4	3	2	1	0	3	4	5	5	6	4	5
3	2	1	2	3	0	1	2	2	3	3	4
4	3	2	3	4	1	0	1	1	2	2	3
5	4	3	4	5	2	1	0	2	3	3	4
3	2	3	4	5	2	1	2	0	1	1	2
4	3	4	5	6	3	2	3	1	0	2	3
2	1	2	3	4	3	2	3	1	2	0	1
3	2	3	4	5	4	3	4	2	3	1	0

Figure 2. Representation of a molecular hydrogen-depleted graph with the associated vertex distance matrix.

The graph refers to the glucose molecule.

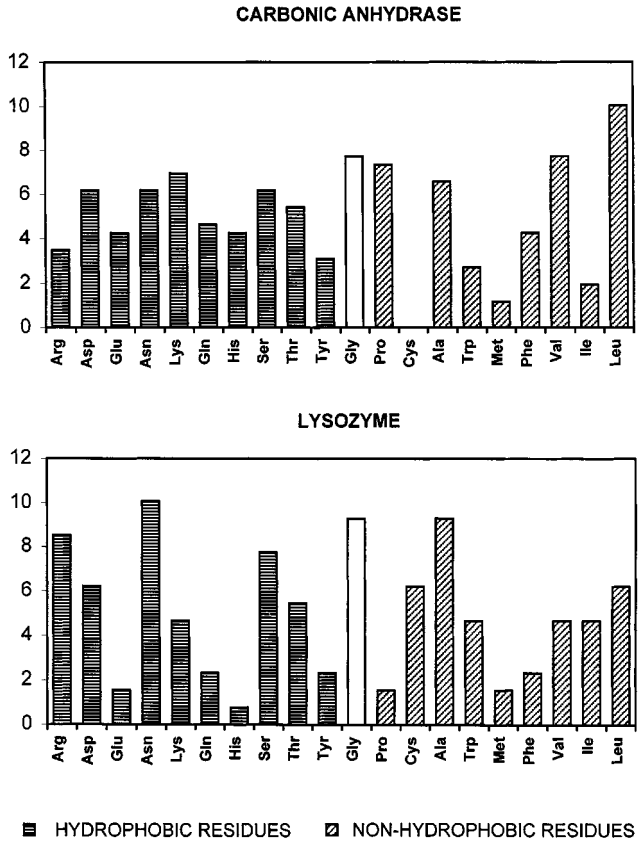


Figure 3. Frequency of occurrence of amino acids in the primary structure of erythrocyte carbonic anhydrase and hen egg lysozyme.

Systems Studied

Proteins

The proteins studied are hen egg lysozyme and erythrocyte carbonic anhydrase. Lysozyme is a heat-stable protein catalyzing the hydrolysis of sugar polymers containing *N*-acetylglucosamine (Blake et al., 1978). Carbonic anhydrase catalyzes the reversible hydration of carbon dioxide (Pocker and Sarkanen, 1978) and is currently being investigated for the development of enzymatic extracorporeal membrane oxygenators (Salley et al., 1990). The molecular dimensions and other properties for the two proteins are reported in Table 1. As can be seen, lysozyme is smaller and more compact than carbonic anhydrase. The two proteins also differ in aminoacidic composition, as results from the amino-acid profiles shown in Figure 3. The histograms were obtained by evaluating the frequency of occurrence of each amino acid in the protein's primary structure. Amino acids were ordered according to the hydrophobicity of their side chains, with glycine demarcating the two classes of residues: hydrophobic (10 components, from arginine to tyrosine) and nonhydrophobic (9 components, from proline to leucine) (Creigh-

Table 1. Structural and Physicochemical Properties of the Two Proteins

Protein	EC No.	MW	N_R	Dimensions (Å)	D^* (10^{-7} cm$^2\cdot$s^{-1})	ν^{**} (cm$^3\cdot$g^{-1})
Carbonic anhydrase	4.2.1.1	28,800	259	47×41×41	10.7	0.729
Lysozyme	3.1.1.17	14,320	129	45×30×30	11.3	0.703

*Calculated at 20°C in water and extrapolated to zero protein concentration.
**Calculated at 25°C.
Note: MW is the Molecular Weight, N_R is the number of residues, D is the translational diffusion coefficient, and ν is the partial specific volume.
Source: Data were taken from Squire and Himmel (1979) and from Creighton (1993).

ton, 1993). For each class we calculated an average hydrophobicity index as

$$I_h = \frac{\sum_{i=1}^{n_c} f_i \Delta G_{\mathrm{tr},i}}{\sum_{i=1}^{n_c} f_i}, \qquad (16)$$

where n_c is the number of components of the class, f_i is the frequency of occurrence of amino acid i in the primary structure, and $\Delta G_{\mathrm{tr},i}$ is the corresponding free energy of transfer from cyclohexane to water. Free energies of transfer were taken from Creighton (1993) and were obtained using side-chain analogs as model compounds for evaluation of the partition coefficient between cyclohexane and water. Since in a globular protein hydrophobic residues are predominantly in the protein core while nonhydrophobic residues are mainly on the surface of the macromolecule, the I_h-values for the two classes provide a measure of the hydrophobicity of the protein interior and exterior, respectively. We got:

Carbonic anhydrase:

$$I_{h,\mathrm{int}} = 6.77 \text{ kcal}\cdot\text{mol}^{-1}, \qquad I_{h,\mathrm{ext}} = -2.22 \text{ kcal}\cdot\text{mol}^{-1}$$

Lysozyme:

$$I_{h,\mathrm{int}} = 7.83 \text{ kcal}\cdot\text{mol}^{-1}, \qquad I_{h,\mathrm{ext}} = -1.98 \text{ kcal}\cdot\text{mol}^{-1},$$

where $I_{h,\mathrm{int}}$ and $I_{h,\mathrm{ext}}$ represent the hydrophobicity of the internal and external regions. The above results indicate that lysozyme interior is more hydrophobic than that of carbonic anhydrase. This point, along with the more limited dimensions of lysozyme, allow the larger stability of this protein to be justified on a physical ground. Stability data for the two proteins in pure buffer are summarized in Table 2.

Additives

Data taken from the literature were relative to the following classes of additives: alcohols (ethanol, 1-propanol), glycols (ethanediol, 1,2-propanediol, glycerol, erythritol, xylitol) and sugars (glucose, fructose, maltose, sucrose, sorbitol). Most runs were made with individual solvent components, at concentrations ranging from 5% to 50% (w/w). In some cases binary mixtures of additives were used.

The additives investigated by us belong to a family of glycine-based amino-acid derivatives. They comprise glycine, *N*-methylglycine (sarcosine), *N*,*N*-dimethylglycine, and *N*,*N*,*N*-trimethylglycine (betaine). The concentrations explored were between 10 and 30% (w/w).

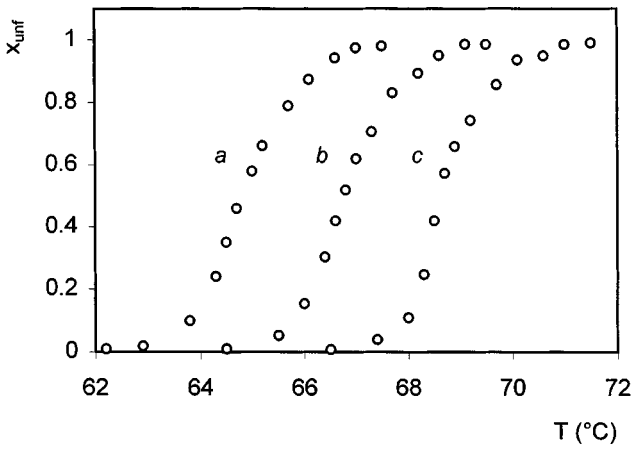


Figure 4. Experimental melting curves of carbonic anhydrase in media containing sarcosine.
(a) 10% w/w; (b) 20%; (c) 30%.

Results and Discussion

Stability of carbonic anhydrase in systems containing aminoacidic components

Some experimental results showing the temperature dependence of the fraction of unfolded protein molecules, that is, the protein's melting curve, in systems containing sarcosine are presented in Figure 4. Data plotted in the figure were obtained from the first derivative of the difference absorption spectrum at constant wavelength ($\delta = dA/d\lambda$). In particular, according to the two-state hypothesis, the experimentally determined δ values were expressed as

$$\delta = x_{\mathrm{fold}}\delta_{\mathrm{fold}} + x_{\mathrm{unf}}\delta_{\mathrm{unf}}, \qquad (17)$$

where x_{fold} and x_{unf} are the fractions of protein molecules in the folded and unfolded conformations, respectively, while δ_{fold} and δ_{unf} represent their contributions to the overall value of δ. Since at any temperature $x_{\mathrm{fold}} + x_{\mathrm{unf}} = 1$, Eq. 17 yields

$$x_{\mathrm{unf}} = \frac{\delta - \delta_{\mathrm{fold}}}{\delta_{\mathrm{unf}} - \delta_{\mathrm{fold}}}. \qquad (18)$$

To evaluate, at a given temperature, x_{unf} from the corresponding δ, it is necessary to know the temperature dependence of δ_{fold} and δ_{unf}. The functions $\delta_{\mathrm{fold}}(T)$ and $\delta_{\mathrm{unf}}(T)$ were calculated from data points before and after the transition, respectively, using a least-squares procedure. The pre-

Table 2. Thermal Unfolding Data for the Two Proteins in Pure Buffer

Protein	Medium	pH	Method*	T_m (°C)	$\Delta H^0_{\mathrm{unf}}$ (kJ·mol^{-1})	Ref.
Carbonic anhydrase	Phosphate 0.1 M	7	UVS	62.9	1062**	†
Lysozyme	Acetate 0.1 M	4	UVS	74.2	597**	††
	Glycine 0.01 M	4	DSC	74.8	473	‡

*Experimental technique used to follow thermal unfolding (UVS: UV difference spectroscopy; DSC: differential scanning calorimetry).
**Estimated from the slope of the thermal unfolding curve.
†Cioci et al. (1996).
††Cioci and Lavecchia (1997a).
‡Gekko (1982).

Table 3. Experimental Melting Temperatures of Carbonic Anhydrase in Media Containing Aminoacidic Additives

Additive	% w/w	T_m (°C)	Additive	% w/w	T_m (°C)
Glycine	10	63.0	Dimethylglycine	10	63.9
	20	63.8		20	65.5
	30	64.5		30	67.1
Sarcosine	10	64.7	Betaine	10	63.7
	20	66.6		20	64.9
	30	68.7		30	66.6

and post-transition base lines so obtained were then extrapolated to the transition region to allow calculation of the fraction of unfolded protein according to Eq. 18. The melting temperature of the protein (T_m) was determined at the midpoint of the transition curve, where $x_{\text{unf}} = 0.5$, leading to the results summarized in Table 3. It can be seen that in all media T_m varies monotonically with the additive concentration. In addition, the aminoacidic components increase the thermal stability of carbonic anhydrase according to the following order: sarcosine > dimethylglycine > betaine > glycine. Thus, in the concentration range explored, the lower the degree of glycine methylation, the greater the ability of the additive to stabilize carbonic anhydrase. These results are in qualitative agreement with reports in the literature concerning other proteins, such as bovine ribonuclease A (Santoro et al., 1992), T7 DNA polymerase (Thakar et al., 1994), and hen egg lysozyme (Santoro et al., 1992).

Effect of additives on thermodynamic protein stability

Figure 5 shows typical results on the influence of solvent components on the melting temperature of the two proteins. Data relative to lysozyme refer to media containing an effective stabilizer and a denaturing additive. For carbonic anhydrase, the effect of two stabilizing components is shown. Similar linear trends were observed for the other systems investigated, supporting the ability of Eq. 12 to describe the dependence of melting temperature on additive's concentration. The parameter α appearing in this equation was estimated by a least-square minimization technique after substitution of the pertinent T_m and ΔH^0_{unf} values. Fitting results are presented in Table 4, along with the corresponding experimental concentration ranges. From the estimates of α the following

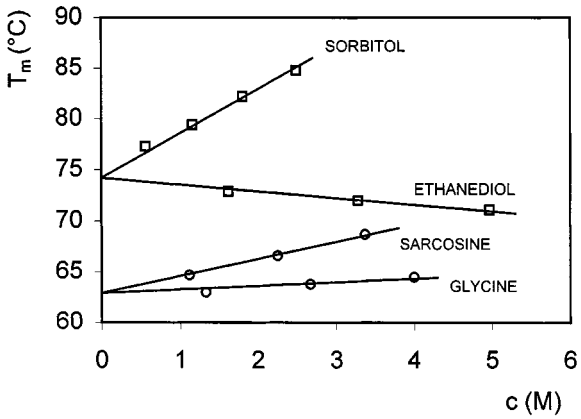


Figure 5. Effect of additives on the melting temperature of hen egg lysozyme (□) and erythrocyte carbonic anhydrase (○).

Experimental data for lysozyme were taken from Cioci and Lavecchia (1997a).

points can be made:

1. According to the sign of α, additives can be partitioned into stabilizers ($\alpha > 0$) and destabilizers ($\alpha < 0$). Glycerol, xylitol, sugars, and glycine-based amino-acid derivatives belong to the first class, whereas ethanol, ethanediol, and 1,2-propanediol are included in the second.

2. For each protein, the absolute value of α provides a measure of the degree of stabilization (or destabilization). Sugars, and particularly disaccharides, are the most effective stabilizers, while ethanol and 1,2-propanediol are the additives inducing the highest destabilization.

Figure 6 illustrates a comparison between experimental and calculated melting temperatures. As is shown, a reasonably good agreement is obtained, with errors on temperature ranging from $-0.97°$ to 1.05°C. In addition, no systematic deviation is observed, as the data points are evenly distributed around the bisection line. The average percent error on T_m^* was 0.5% and the standard deviation was 0.398.

An interesting outcome of the molecular thermodynamic model is the possibility of using the parameter α to reconstruct the protein's thermal unfolding curve, that is, the temperature dependence of the fraction of unfolded protein in a

Table 4. Estimated α Values for the Two Proteins

Protein	Additive	Δc (% w/w)	α (kJ mol^{-2} L)	Protein	Additive	Δc (% w/w)	α (kJ mol^{-2} L)
Carbonic anhydrase	Ethanol	5 – 20	−11.80	Carbonic anhydrase	Sarcosine	10 – 30	5.34
	Ethanediol	5 – 35	−2.54		Dimethylglycine	10 – 30	4.37
	1,2-Propanediol	5 – 25	−7.58		Betaine	10 – 30	4.21
	Glycerol	10 – 35	2.79				
	Xylitol	10 – 30	8.75	Lysozyme	Ethanol	10 – 30	−5.66
	Maltose	10 – 30	16.85		Ethanediol	10 – 30	−1.11
	Glucose	10 – 30	10.48		1,2-Propanediol	10 – 30	−2.19
	Sucrose	10 – 30	16.45		Glycerol	10 – 50	2.01
	Fructose	10 – 30	5.37		Glucose	10 – 30	6.00
	Sorbitol	10 – 30	11.12		Sorbitol	10 – 30	7.52
	Glycine	10 – 30	1.13		Sucrose	10 – 30	11.82

Note: Δc is the additive concentration range.

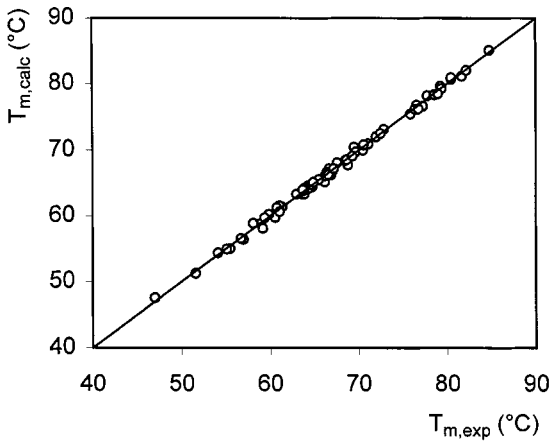


Figure 6. Experimental and calculated melting temperatures for the two proteins; experimental data for lysozyme were taken from Gekko (1982), Cioci et al. (1996), and Cioci and Lavecchia (1997a).

given medium. The analytical form of the function $x_{\text{unf}}(c,T)$ can be derived by combining Eqs. 6, 9 and 11. We get

$$x_{\text{unf}}(c,T)=\left[1+\exp\left(\frac{\Delta H^0_{\text{unf}}}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right)+\frac{\alpha c}{RT}\right)\right]^{-1}. \quad (19)$$

As indicated by Eq. 19 three parameters, α, T_m, and ΔH^0_{unf}, are required to estimate the thermal unfolding curve of the protein. Representative examples of reconstructed melting curves are reported in Figure 7. The excellent agreement between experimental data and model calculations is evident from these diagrams. It should be pointed out that the profiles were not obtained by fitting $x_{\text{unf}}-T$ data; hence they can be considered at least partly predictive.

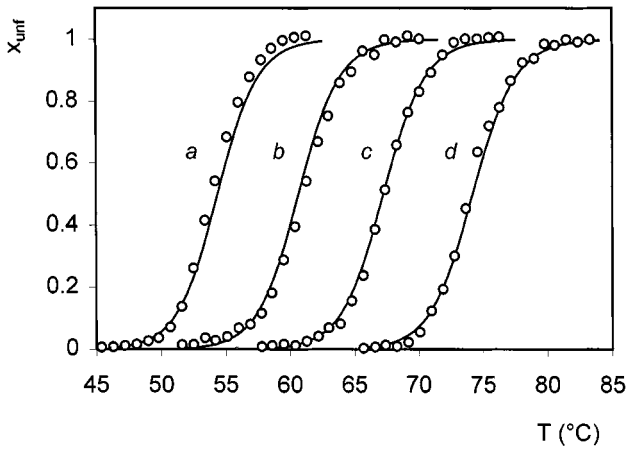


Figure 7. Experimental and calculated melting curves for hen egg lysozyme in media containing ethanol. (a) 30% w/w; (b) 20%; (c) 10%; and (d) in pure buffer.

Experimental data were taken from Cioci and Lavecchia (1997a).

The results obtained strongly validate the molecular thermodynamic approach followed, and in particular the assumed linear dependence of ΔG_{tr} on the additive concentration (Eq. 11). Referring to this assumption, we observe that it appears to be appropriate both in media containing stabilizing additives and in the presence of destabilizing components. This is an important point, since the few experimental supports reported in the literature were obtained for denaturants, mainly urea and guanidine hydrochloride (Santoro and Bolen, 1988, 1992; Shortle et al., 1989). From a physical viewpoint, evidence is therefore provided that the same basic phenomena are involved in stabilization and destabilization.

Correlation of stability data by information indices

We calculated average information indices for each solvent component, obtaining the results shown in Table 5. Looking now at the information index on atomic composition (I_{AC}), we note that additives with the same elemental chemical composition have identical I_{AC} values. This is the case of the two pairs, glucose–fructose and sucrose–maltose. A partial discrimination is achieved when considering the information index based on the vertex distance matrix (I_{VD}). In fact, whereas for glucose and fructose I_{VD} is still the same, the values for sucrose and maltose are different. Such a result is a reflection of the diverse topological features of the molecular graphs associated with the latter components.

We used the following empirical expression to correlate the α estimates in terms of the two indices

$$\alpha = a_1 + a_2(I_{AC} + I_{VD}), \quad (20)$$

where a_i are additive-independent parameters specific for each protein. Adoption of such a simple relation can be partly justified, with respect to use of more complex expressions, by the need to limit the number of fitting parameters. Equation 20 implies that an overall information index defined as

$$I_0 = I_{AC} + I_{VD} \quad (21)$$

is used to quantify the information content of the solvent components. As observed by Bonchev (1983), approaches

Table 5. Average Information Indices for the Various Solvent Components

Additive	Formula	N_A	N_V	I_{AC}	I_{VD}
Ethanol	C_2H_6O	9	3	1.224	1.530
Ethanediol	$C_2H_6O_2$	10	4	1.371	1.906
1,2-Propanediol	$C_3H_8O_2$	13	5	1.335	1.939
Glycerol	$C_3H_8O_3$	14	6	1.414	2.171
Xylitol	$C_5H_{12}O_5$	22	10	1.448	2.580
Glucose	$C_6H_{12}O_6$	24	12	1.500	2.512
Fructose	$C_6H_{12}O_6$	24	12	1.500	2.512
Sorbitol	$C_6H_{14}O_6$	26	12	1.457	2.755
Sucrose	$C_{12}H_{22}O_{11}$	45	23	1.510	3.221
Maltose	$C_{12}H_{22}O_{11}$	45	23	1.510	3.247
Glycine	$C_2H_5NO_2$	10	5	1.761	1.939
Sarcosine	$C_3H_7NO_2$	13	6	1.670	2.241
Dimethylglycine	$C_4H_9NO_2$	16	7	1.579	2.268
Betaine	$C_5H_{11}NO_2$	19	8	1.509	2.250

Note: N_A is the number of atoms in the molecule and N_V is the number of vertices for the associated molecular graph.

Table 6. Fitting of α values in Terms of Average Information Indices

Protein	N_P	a_1	a_2	R	σ^2
Carbonic anhydrase	14	-48.52 ± 4.20	13.76 ± 1.07	0.965	4.951
Lysozyme	7	-31.08 ± 1.24	9.13 ± 0.33	0.997	0.294

Note: N_P is the number of data points, R is the correlation coefficient, and σ^2 is the error variance.

based on overall information indices are the most straightforward way to represent molecular complexity. The results obtained by fitting Eq. 20 to the α values are summarized in Table 6. Analysis of correlation coefficients and σ^2 values indicates that the equation describes quite satisfactorily the dependence of α on the overall information index.

A comparison between experimental and calculated results is shown in Figures 8 and 9. Examination of these diagrams reveals that a demarcation between stabilization and destabilization occurs for both proteins at an I_0 value close to 3.5. This value corresponds to $\alpha = 0$, that is, to the unperturbed protein stability. In addition, the stabilizing power of the additives increases with I_0. In other words, highly structured molecules appear to exert a more beneficial influence on stability.

Another important point to be stressed is that data points relative to hydroxylic and amino-acidic components in media containing carbonic anhydrase are disposed around the same line. That is to say that, under the experimental conditions considered, perturbations in stability are not exclusively ascribable to behavioral similarities between the protein and the members of a particular chemical class. Although further experimental support is needed, these results appear to suggest that the overall information index used is capable of capturing at least some of the additive's molecular features that determine its stabilizing or destabilizing power.

Finally we note that, with minor exceptions, the I_0 scale roughly parallels the surface-tension scale. In particular, additives with $I_0 > 3.5$ exhibit a surface tension higher than that of water. This means that the addition addition of these components to the protein solution increases the surface tension of the medium, and hence the interfacial free energy between the protein and the surrounding solvent, hindering in-

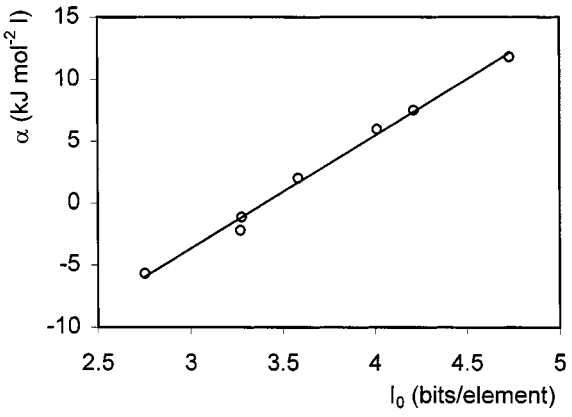


Figure 8. Correlation of α values in terms of the overall information index for hen egg lysozyme.

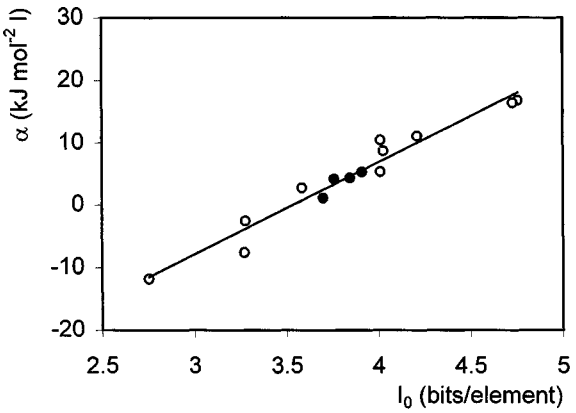


Figure 9. Correlation of α values in terms of the overall information index for erythrocyte carbonic anhydrase in media containing hydroxylic (○) and aminoacidic (●) additives.

activation (Cioci and Lavecchia, 1997a). Interestingly, however, correlation of the α values in terms of surface tension is not as good as that obtained from the additive's information index. We can therefore infer that the surface-tension mechanism is actively involved in the preferential hydration of the protein, but probably in connection with other phenomena. The overall information index defined by Eq. 21 provides a numerical measure of the extent to which the chemical and topological features of the solvent components contribute to such phenomena.

To further validate the theoretic information approach, an attempt was made to predict protein stability in perturbed solvents. The additives examined include 1-propanol, erthyritol, and xylitol. For all of them the experimental data available were relative to a single concentration value, so that they couldn't be used to obtain a reliable estimate of α. This latter was then predicted by Eq. 20. Then the perturbed melting temperature was determined using Eq. 12. The results obtained are presented in Table 7. As can be seen, fairly good predictions result, with an average percent error on temperature of 2.26%.

Protein stability in media containing mixed additives

To estimate or correlate stability data in media containing two or more additives, the information indices for the mixture of interest should first be evaluated. Calculation of mixture properties from those of single components requires the introduction of an appropriate mixing rule. We considered the following relationship

$$I_{\text{mix}} = \frac{\sum_{k=1}^{n} c_k I_k}{\sum_{k=1}^{n} c_k}, \tag{22}$$

where n is the number of additives in the mixture, I_k is the information index for the k-component, and c_k is the corresponding molar concentration. Substitution of I_{mix} into Eq.

Table 7. Experimental and Predicted Melting Temperatures in Media Containing Pure or Mixed Additives

Protein	Additive/es	c (% w/w)	Ref.	I_{AC}	I_{VD}	$T^*_{m,\exp}$ (°C)	T^*_m (°C)
Carbonic anhydrase	1-Propanol	5	*	1.189	1.906	57.7	61.4
	Ethanol + Ethanediol	10 + 10	*	1.286††	1.668††	53.5	53.9
	Ethanol + Ethanediol	10 + 20	*	1.309††	1.748††	51.6	51.8
Lysozyme	Erthyritol	30	**	1.435	2.385	79.8	80.1
	Xylitol	30	**	1.448	2.580	81.3	81.4
	Ethanediol + 1,2-Propanediol	10 + 10	†	1.355††	1.921††	73.8	72.2
	Ethanol + 1,2-Propanediol	10 + 10	†	1.265††	1.682††	65.0	65.8
	Ethanol + 1,2-Propanediol	10 + 20	†	1.283††	1.747††	61.4	64.7
	Ethanol + 1,2-Propanediol	20 + 10	†	1.250††	1.624††	57.1	58.8
	Ethanol + Ethanediol	20 + 10	†	1.262††	1.628††	58.3	58.5

Note: c is the additive concentration.
*Cioci et al. (1996).
**Gekko (1982).
†Cioci and Lavecchia (1997a).
††Calculated by Eq. 22.

20 allows estimation of α and, from Eq. 12, of the perturbed melting temperature.

A comparison between experimental data and model calculations is shown in Table 7. The data considered refer to the melting temperatures of lysozyme and carbonic anhydrase in systems containing binary mixtures of additives. As is apparent, the results are very good. The average percent error on temperature was 1.88%. Some melting curves reconstructed from the I_{mix} values are reported in Figure 10. We note that no experimental information on the mixtures in question was used. Plotted curves can therefore be considered as purely predictive.

To sum up, the results emerging from the analysis of stability data in systems containing pure or mixed additives indicate that molecular thermodynamics and information theory can be effectively used to describe solvent-induced perturbations of stability. Since calculation of the overall information index defined by Eq. 21 simply requires knowledge of the additive's molecular structure, an extensive computational screening can be performed to select the most appropriate additives, once a limited number of experiments on the protein of interest has been carried out. Moreover, a better description of the information content of the added components could probably be achieved by incorporating further structural or topological contributions into the overall index.

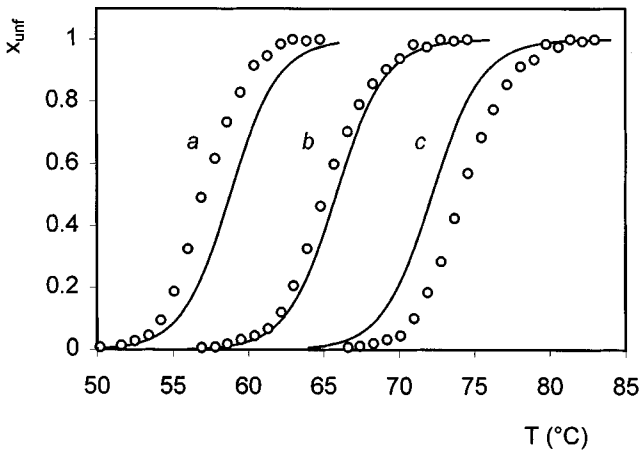


Figure 10. Experimental and predicted melting curves in media containing binary mixtures of additives.

(a) 1,2-propanediol + ethanol (10% + 20% w/w); (b) 1,2-propanediol + ethanol (10% + 10% w/w); (c) 1,2-propanediol + ethanediol (10% + 10% w/w). The solid line represents calculations by Eq. 19. Experimental data were taken from Cioci and Lavecchia (1997a).

Conclusions

In this contribution we have shown that the influence of additives on protein stability can be quantified by an empirical parameter related to the free energy of transfer of the protein from the pure to the mixed solvent. This parameter, along with two easily measurable quantities, the melting temperature of the protein in pure buffer and the corresponding enthalpy change on unfolding, can be used to calculate melting temperatures and reconstruct thermal unfolding curves in media containing pure or mixed additives. Another point to emphasize is the good correlation obtained when expressing the parameter α in terms of theoretic information indices. Moreover, the results pertinent to carbonic anhydrase appear to indicate that the same correlation can be used to describe perturbations in stability induced by additives belonging to different chemical classes.

At the end of this work it seems worth remembering that approaches based on information indices are fully heuristic and justifiable when the complexity of the system under investigation prevents a more fundamental approach. Accordingly, the success of a correlation involving theoretical information indices should not be assessed on the basis of its theoretical significance, but rather of its practical usefulness. The proposed procedure could be helpful for reducing the number of additives to be experimentally tested when stabilizing proteins by controlled changes of the solvent. Solvent engineering is an interesting alternative to protein engineering, which is aimed at increasing the intrinsic stability of a protein by redesigning the macromolecule through changes in its amino-acidic composition. Engineered proteins can exhibit very high stabilities, but undesired irreversible changes often occur in their biological properties. By contrast, the solvent could be modified just when required, for instance, during protein storage or in the course of a heat-treatment step, without inducing any irreversible modification of the protein structure and function.

Acknowledgments

This work was partially supported by grants from the Italian ''Ministero dell'Università e della Ricerca Scientifica e Tecnologica.''

Literature Cited

Arakawa, T., and S. N. Timasheff, ''Stabilization of Protein Structure by Sugars,'' *Biochemistry*, **21**, 6536 (1982).

Arakawa, T., and S. N. Timasheff, ''Preferential Interactions of Proteins with Solvent Components in Aqueous Amino Acid Solutions,'' *Arch. Biochem. Biophys.*, **224**, 169 (1983).

Arakawa, T., R. Bhat, and S. N. Timasheff, ''Why Preferential Hydration does not Always Stabilize the Native Structure of Globular Proteins,'' *Biochemistry*, **29**, 1924 (1990).

Asenjo, J. A., and I. Patrick, ''Large-Scale Protein Purification,'' *Protein Purification Application: A Practical Approach*, E. L. V. Harris and S. Angal, eds., IRL Press, Oxford, p. 1 (1990).

Becktel, W. J., and J. A. Schellman, ''Protein Stability Curves,'' *Biopolymers*, **26**, 1859 (1987).

Bhat, R., and S. N. Timasheff, ''Steric Exclusion Is the Principal Source of the Preferential Hydration of Proteins in the Presence of Polyethylene Glycols,'' *Protein Sci.*, **1**, 1133 (1992).

Blake, C. C. F., D. E. P. Grace, L. N. Johnson, S. J. Perkins, D. C. Philips, R. Cassels, C. M. Dobson, F. M. Poulsen, and R. J. P. Williams, ''Physical and Chemical Properties of Lysozyme,'' *Molecular Interactions and Activity in Proteins*, Excerpta Medica, Amsterdam, p. 137 (1978).

Bonchev, D., *Information Theoretic Indices for Characterization of Chemical Structures*, Wiley, New York (1983).

Bonchev, D., and N. Trinajstic, ''Information Theory, Distance Matrix and Molecular Branching,'' *J. Chem. Phys.*, **67**, 4517 (1977).

Bonchev, D., and N. Trinajstic, ''On Topological Characterization of Molecular Branching,'' *Int. J. Quant. Chem.*, **12**, 293 (1978).

Cioci, F., ''Thermostabilization of Erthyrocyte Carbonic Anhydrase by Polyhydric Addtiives,'' *Enzyme Microb. Technol.*, **17**, 592 (1995a).

Cioci, F., ''Catalytic Activity of *Aspergillus Niger* Glucose Oxidase in Water-Polyol Mixtures,'' *Catal. Lett.*, **35**, 395 (1995b).

Cioci, F., and R. Lavecchia, ''Effect of Polyols and Sugars on Heat-Induced Flavin Dissociation in Glucose Oxidase,'' *Biochem. Mol. Biol. Int.*, **34**, 705 (1994).

Cioci, F., and R. Lavecchia, ''Molecular Thermodynamics of Heat-Induced Protein Unfolding in Aqueous Media,'' *AIChE J.*, **43**, 525 (1997a).

Cioci, F., and R. Lavecchia, ''Interfacial Phenomena and Solvent Effects in Protein Stability,'' *Rec. Res. Dev. Phys. Chem.*, **1**, 369 (1997b).

Cioci, F., and R. Lavecchia, ''Thermostabilization of Proteins by Water-Miscible Additives,'' *Chem. Biochem. Eng. Q.*, **12**, 191 (1998).

Cioci, F., and R. Lavecchia, ''Sorbitol-Mediated Stabilization of Human IgG Against Thermal Inactivation,'' *Biotech. Tech.*, (1999).

Cioci, F., R. Lavecchia, and L. Marrelli, ''Perturbation of Surface Tension of Water by Polyhydric Additives: Effect on Glucose Oxidase Stability,'' *Biocatalysis*, **10**, 137 (1994).

Cioci, F., R. Lavecchia, and L. Marrelli, ''Effect of Surface Tension on the Conformational Stability of Erthyrocyte Carbonic Anhydrase,'' *Fluid Phase Equilibria*, **116**, 118 (1996).

Creighton, T. E., ''Stability of Folded Conformations,'' *Curr. Opinion Struct. Biol.*, **1**, 5 (1991).

Creighton, T. E., *Proteins: Structures and Molecular Properties*, 2nd ed., Freeman, New York, p. 154 (1993).

Dancoff, S. M., and H. Quastler, *Essays on the Use of Information Theory in Biology*, H. Quastler, ed., Illinois Univ. Press, Urbana (1953).

Dill, K. A., ''Theory for the Folding and Stability of Globular Proteins,'' *Biochemistry*, **24**, 1501 (1985).

Dill, K. A., ''Dominant Forces in Protein Folding,'' *Biochemistry*, **29**, 7133 (1990).

Fernandez, M. M., D. S. Clark, and H. W. Blanch, ''Papain Kinetics in the Presence of a Water-Miscible Organic Solvent,'' *Biotech. Bioeng.*, **37**, 967 (1991).

Gekko, K., ''Calorimetric Study on Thermal Denaturation of Lysozyme in Polyol-Water Mixtures,'' *J. Biochem.*, **91**, 1197 (1982).

Greene, R. F., and C. N. Pace ''Urea and Guanidine Hydrochloride Denaturation of Ribonuclease, Lysozyme, α-Chymotrypsin and β-Lactoglobulin,'' *J. Biol. Chem.*, **249**, 5388 (1974).

Kier, L. B., and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York (1976).

Klibanov, A. M., ''Stabilization of Enzymes Against Thermal Inactivation,'' *Adv. Appl. Microbiol.*, **29**, 1 (1983).

Laane, C., S. Boeren, R. Hilhorst, and C. Veeger, ''Optimization of Biocatalysis in Organic Media,'' *Biocatalysis in Organic Media*, C. Laane, J. Tramper, and M. D. Lilly, eds., Elsevier, Amsterdam, p. 65 (1987).

Lavecchia, R., and M. Zugaro, ''Thermal Denaturation of Erythrocyte Carbonic Anhydrase,'' *FEBS Lett.*, **292**, 162 (1991).

Lozano, P., D. Combes, and J. L. Iborra, ''Effect of Polyols on α-Chymotrypsin Thermostability: A Mechanistic Analysis of the Enzyme Stabilization,'' *J. Biotech.*, **35**, 9 (1994).

Matsumoto, M., K. Kida, and K. Kondo, ''Effects of Polyols and Organic Solvents on Thermostability of Lipase,'' *J. Chem. Technol. Biotechnol.*, **70**, 188 (1997).

Pace, C. N., ''The Stability of Globular Proteins,'' *CRC Crit. Rev. Biochem.*, **3**, 1 (1975).

Pace, C. N., ''Conformational Stability of Globular Proteins,'' *Trends Biochem. Sci.*, **15**, 14 (1990).

Pace, C. N., and D. V. Laurents, ''A New Method for Determining the Heat Capacity Change for Protein Folding,'' *Biochemistry*, **28**, 2520 (1989).

Pace, C. N., B. A. Shirley, and J. A. Thomson, ''Measuring the Conformational Stability of a Protein,'' *Protein Structure: A Practical Approach*, T. E. Creighton, ed., IRL Press, Oxford, p. 311 (1989).

Pace, C. N., B. A. Shirley, M. McNutt, and K. Gajiwala, ''Forces Contributing to the Conformational Stability of Proteins,'' *FASEB J.*, **10**, 75 (1996).

Pfeil, W., and P. L. Privalov, ''Conformational Changes in Proteins,'' *Biochemical Thermodynamics*, M. N. Jones, ed., Elsevier, Amsterdam, p. 75 (1979).

Pocker, Y., and S. Sarkanen, ''Carbonic Anhydrase: Structure, Catalytic Versatility and Inhibition,'' *Adv. Enzymol.*, **47**, 149 (1978).

Privalov, P. L., ''Stability of Proteins,'' *Adv. Protein Chem.*, **33**, 167 (1979).

Privalov, P. L., and S. J. Gill, ''Stability of Protein Structure and Hydrophobic Interaction,'' *Adv. Protein Chem.*, **39**, 191 (1988).

Russell, A. J., and C. Vierheller, ''Protein Engineering,'' *The Biochemical Engineering Handbook*, J. D. Bronzio, ed., CRC Press, New York, p. 1445 (1995).

Salley, S. O., J. Y. Song, G. C. Whittlesey, and M. D. Klein, ''Immobilized Carbonic Anhydrase in a Membrane Lung for Enhanced CO_2 Removal,'' *ASAIO Trans.*, **36**, 486 (1990).

Santoro, M. M., and D. W. Bolen, ''Unfolding Free Energy Changes Determined by the Linear Extrapolation Method,'' *Biochemistry*, **27**, 8063 (1988).

Santoro, M. M., and D. W. Bolen, ''A Test to the Linear Extrapolation of Unfolding Free Energy Changes over an Extended Denaturant Concentration Range,'' *Biochemistry*, **31**, 4901 (1992).

Santoro, M. M., Y. Liu, S. M. A. Khan, L. X. Hou, and D. W. Bolen, ''Increased Thermal Stability of Proteins in the Presence of Naturally Occurring Osmolytes,'' *Biochemistry*, **31**, 5278 (1992).

Schellman, J. A., ''Solvent Denaturation,'' *Biopolymers*, **17**, 1305 (1978).

Schellman, J. A., ''Selective Binding and Solvent Denaturation,'' *Biopolymers*, **26**, 549 (1987).

Schellman, J. A., ''A Simple Model for Solvation in Mixed Solvents. Application to the Stabilization and Destabilization of Macromolecular Structures,'' *Biophys. Chem.*, **37**, 121 (1990).

Schellman, J. A., ''The Thermodynamics of Solvent Exchange,'' *Biopolymers*, **34**, 1015 (1994).

Schmid, F. X., ''Spectral Methods of Characterizing Protein Conformation and Conformational Changes,'' *Protein Structure: A Practical Approach*, T. E. Creighton, ed., IRL Press, Oxford, p. 251 (1989).

Shortle, D., A. K. Meeker, and S. L. Gerring, ''Effects of Denaturans at Low Concentrations on the Reversible Denaturation of Staphylococcal Nuclease,'' *Arch. Biochem. Biophys.*, **272**, 103 (1989).

Squire, P. G., and M. E. Himmel, ''Hydrodynamics and Protein Hydration,'' *Arch. Biochem. Biophys.*, **196**, 165 (1979).

Tanford, C., ''Protein Denaturation. The Transition from Native to Denatured State,'' *Adv. Protein Chem.*, **23**, 121 (1968).

Tanford, C., ''Protein Denaturation. Theoretical Models for the Mechanism of Denaturation'' *Adv. Protein Chem*., **24**, 1 (1970).
Thakar, M., A. Bilenko, and W. J. Becktel, ''Osmolyte Mediation of T7 DNA Polymerase and Plasmid DNA Stability,'' *Biochemistry*, **33**, 12255 (1994).
Timasheff, S. N., ''Water as Ligand: Preferential Binding and Exclusion of Denaturants in Protein Unfolding,'' *Biochemistry*, **31**, 9857 (1992).
Timasheff, S. N., and T. Arakawa, ''Stabilization of Protein Structure by Solvents," *Protein Structure: A Practical Approach*, T. E. Creighton, ed., IRL Press, Oxford, p. 331 (1989).
Tomazic, S. J., ''Protein Stabilization,'' *Biocatalysis for Industry*, J. S. Dordick, ed., Plenum Press, New York, p. 241 (1991).
Volkin, D. B., and A. M. Klibanov, ''Minimizing Protein Inactivation,'' *Protein Function: A Practical Approach*, T. E. Creighton, ed., IRL Press, Oxford, p. 1 (1989).
Wilson, R. J., *Introduction to Graph Theory*, Academic Press, New York (1972).
Wyman, J., and S. J. Gill, *Binding and Linkage: Functional Chemistry of Biological Macromolecules*, Univ. Science Books, Mill Valley, CA (1990).

Manuscript received Mar. 19, 1999, and revision received Jan. 19, 2000.