Hydration changes upon protein unfolding: cosolvent effect analysis

Hassan O. Hammou, Isabel M. Plaza del Pino and Jose M. Sanchez-Ruiz*

Departamento de Quimica Fisica, Facultad de Ciencias, Instituto de Biotecnologia, Universidad de Granada 18071, Granada, Spain

We have characterized the unfolding energetics of ribonuclease a and hen egg white lysozyme as a function of temperature, pH and concentration of several cosolvents (sucrose, glucose, glycerol, polyethyleneglycol 8000) that are expected to be preferentially excluded from the surface of native proteins, and we have calculated the corresponding unfolding changes in preferential hydration ($\Delta\Gamma_{21}$) at 25 °C. We find no significant cosolvent concentration effect on the unfolding enthalpy and heat capacity values, which suggests that the cosolvents do not interact strongly with the proteins at the comparatively low cosolvent concentrations employed in this work. In spite of this, the $\Delta\Gamma_{21}$ values are significantly smaller than theoretical estimates of the unfolding change in the number of water molecules corresponding to first monolayer coverage (ΔN_1), even when, for the purpose of the ΔN_1 calculation, the solvent accessibility in the unfolded state is modelled on the basis of compact fragments extracted from folded protein structures (T. P. Creamer, R. Srinivasan and G. D. Rose, *Biochemistry*, 1997, **36**, 2832). An analysis in terms of the two-domain (local-bulk) solvent model shows that the low values found for $\Delta\Gamma_{21}$ could be the result of the entrance of rather small amounts of cosolvent in the local domain of the native and/or the unfolded protein. In general, the two-domain model suggests that even a weak protein–cosolvent interaction may significantly distort the membrane-free, osmotic stress estimates of the number of water molecules involved in protein conformational changes.

The preferential hydration of a protein in a water–cosolvent mixture (Γ_{21}) is defined as the amount of water that must be added upon protein addition to the mixture in order to maintain constant the water chemical potential.[1] The unfolding change in preferential hydration (that is, the preferential hydration of the unfolded protein minus that of the native protein: $\Delta\Gamma_{21}$) is an important denaturation parameter,[2,3] since (*i*) it is related through simple thermodynamic equations with the cosolvent effect on protein stability and (*ii*) it can potentially lead to estimates of the number of water molecules that newly interact with the protein upon unfolding. This second point deserves additional comment.

The energetics of protein folding is largely determined by the fact that, upon unfolding, groups that are (totally or partially) buried in the native structure become (totally or partially) exposed to the solvent. This increase in exposure to solvent is usually quantified on the basis of the *accessible surface area* (ASA)[4] and, in fact, studies on the relation between structure and energetics in proteins rely heavily on ASA values calculated for different types of protein surfaces.[5,6] Unfolding changes in the ASA are calculated from the X-ray structure of the native protein, together with one of several plausible models for the exposure to solvent of the amino acid residues in the unfolded state. Such models include tripeptides of sequence Gly—X—Gly or Ala—X—Ala, a single-blocked residue, a fully extended chain and protein fragments excised from the folded structures of proteins.[7,8] It is not clear to what extent these models actually represent the (average) exposure to solvent of the amino acid residues in the unfolding ensemble and, in particular, it appears that some of the models most often used in the past (tripeptides, for instance) may clearly overestimate the exposure. See ref. 7 for a discussion on this matter.

A different approach to the determination of the unfolding increase in exposure to solvent would involve the calculation from experimental data of the number of water molecules that newly interact with the protein upon unfolding. In principle, this number would be given by the unfolding change in preferential hydration ($\Delta\Gamma_{21}$), provided that this change is calculated from the effect on the unfolding energetics caused by a cosolvent that is completely excluded from the hydration shell of both the native and unfolded proteins; this type of calculation is entirely equivalent (see Appendix A) to the osmotic stress method employed in recent years to estimate the number of water molecules involved in ligand binding, channel opening and other biological molecular processes (for a recent review, see ref. 9). Following this approach, we recently calculated $\Delta\Gamma_{21}$ for ribonuclease a (RNase a) from the effect of the stabilizing osmolyte sarcosine on the unfolding energetics as monitored by differential scanning calorimetry (DSC). $\Delta\Gamma_{21}$ turned out to be a surprisingly low number (about 70 molecules of water per molecule of protein) from which no clear conclusion could be reached. Thus, $\Delta\Gamma_{21} \approx 70$ suggests that the unfolded state is compact (with an average exposure to solvent of the amino acid residues only slightly larger than that in the native state) and/or that the osmolyte is not totally excluded from the surface of, at least, the unfolded protein.

In this work, we further explore the possibility of determining the number of water molecules involved in folding-unfolding processes from the effect of cosolvents on the unfolding energetics. To this end, we have characterized the effect of several cosolvents on the unfolding energetics of two model proteins: RNase a and hen egg white (HEW) lysozyme. We have chosen cosolvents that are expected to be preferentially excluded from the surfaces of native proteins: sucrose,[10] glucose,[11] glycerol[12] and polyethyleneglycol 8000 (PEG_{8000}).[13] To determine $\Delta\Gamma_{21}$ we have followed our previous analysis[2] and employed an approach based on the thermodynamic linkage between temperature, pH and cosolvent effects. The study reported in this work involves two proteins and four cosolvents, and required a large number (about 400) of denaturation experiments; in most cases, optical absorption in the UV region was used to monitor unfolding, since this

* Fax: +34 958 272879; e-mail: sanchezr@goliat.ugr.es

allowed us to perform several thermal denaturation experiments within a few hours. The characterization of the unfolding energetics on the basis of optical profiles is acceptable in this case, since (*i*) we are dealing with well-characterized model proteins that are known to conform acceptably to the two-state equilibrium unfolding model in aqueous solution; (*ii*) we study the effect of comparatively low cosolvent concentrations, which are not likely to cause a drastic alteration in the equilibrium unfolding mechanism and (*iii*) as will be shown in this work, an adequate characterization of the unfolding energetics (including an estimate of the unfolding heat capacity change) may be derived from a global analysis of the non-calorimetric unfolding profiles. We must emphasize, nevertheless, that, although this study is mainly based on optical unfolding profiles, we did carry out a significant number of DSC experiments in order to check the correctness of the energetic parameters derived from the global analysis of the optical data.

Results

Profiles of absorbance *vs.* temperature for ribonuclease a and lysozyme unfolding were obtained in water and in the following water–cosolvent mixtures: 0.5, 1, 1.5 and 2 M glucose; 1, 2, 3 and 4 M glycerol; 10, 15, 20 and 25 wt% of PEG_{8000}. In addition, absorbance *vs.* temperature profiles for ribonuclease unfolding were also obtained in water–sucrose mixtures of 0.25, 0.5, 0.75 and 1 M sucrose (similar profiles in water–sucrose, however, could not be obtained with lysozyme, since, at the comparatively high temperatures required to monitor lysozyme unfolding in water–sucrose, a strong UV absorption was observed, which was not due to the protein but, likely, to some substance resulting from sucrose decomposition at high temperature). For each protein and at each solvent composition, seven unfolding profiles were obtained at seven different pH values within the approximate ranges of 2–5 (RNase a) and 1.8–3 (lysozyme) and additional experiments were performed to check the heating rate effect on the unfolding profiles and the reversibility of the unfolding. Thus, the heating rate dependence and the reversibility were checked for both proteins in all the studied solvent compositions and, at least, at two pH values corresponding to the extremes of the studied pH ranges (occasionally, we also carried out experiments on a third pH value in the middle of the studied pH range).

In each case, the characterization of the heating rate effect involved three unfolding experiments at heating rates of 0.25, 0.5 and 1 K min^{-1}; no significant heating rate effect was detected. The characterization of the reversibility involved the recording of the unfolding profile at a heating rate of 1 K min^{-1} and, immediately, the recording of a refolding profile as the temperature was lowered at a constant cooling rate of 1 K min^{-1}. We always found a good agreement between the transitions observed in the unfolding and refolding profiles, indicating that the denaturation processes studied in this work are reversible. However, for both proteins in water–PEG_{8000} mixtures with 15, 20 and 25 wt% of PEG_{8000} and at the more acidic pH values, the absorbances recorded in the cooling direction were significantly higher than those recorded in the heating direction, the difference being larger at the lower temperatures; we attribute this result to light scattering caused by the presence of small amounts of protein aggregates. Nevertheless, the unfolding (heating) profiles did not appear to be distorted by irreversible (and kinetically controlled) aggregation processes, since no heating rate effect on them could be detected. We concluded that the unfolding profiles under the above conditions reflected an equilibrium unfolding process and, accordingly, they were not excluded from subsequent analyses.

Preliminary analysis of the optical unfolding profiles and comparison with DSC results

Analysis of the optical unfolding profiles (expressed as extinction coefficient *vs.* temperature profiles) was carried out on the basis of the two-state equilibrium model:

$$\varepsilon = \frac{\varepsilon_N + K \cdot \varepsilon_D}{1 + K} \quad (1)$$

where ε is the extinction coefficient at a temperature T, and ε_N and ε_D are the extinction coefficients for the native and unfolded proteins, which were assumed to depend linearly on temperature:

$$\varepsilon_N = \alpha_N + \beta_N \cdot T \quad (2)$$

$$\varepsilon_D = \alpha_D + \beta_D \cdot T \quad (3)$$

Preliminary, individual analyses of the unfolding profiles were carried out assuming that the unfolding equilibrium constant ($K = [D]/[N]$) changes with temperature as given by the following integrated van't Hoff equation:

$$\ln K = -\frac{\Delta H^{vH}}{R} \cdot \left(\frac{1}{T} - \frac{1}{T_m}\right) \quad (4)$$

where ΔH^{vH} is the van't Hoff unfolding enthalpy and T_m is the denaturation temperature (here defined as the temperature at which $K = 1$). Eqn. (4) assumes that the unfolding enthalpy change may be taken as a constant within the comparatively narrow temperature range of the thermal unfolding transition. The nonlinear, least-squares fittings of eqn. (1), together with eqn. (2)–(4), to the experimental unfolding profiles were carried out using the program MLAB (Civilized Software); the fits involved six fitting parameters (α_N, β_N, α_D, β_D, T_m and ΔH^{vH}) and were always excellent (see Fig. 1 and 2 for some examples).

The purpose of our DSC experiments was to check the two-state character of the unfolding processes and to ascertain that, in this case, unfolding enthalpies and denaturation temperatures can be reliably determined from the analysis of the

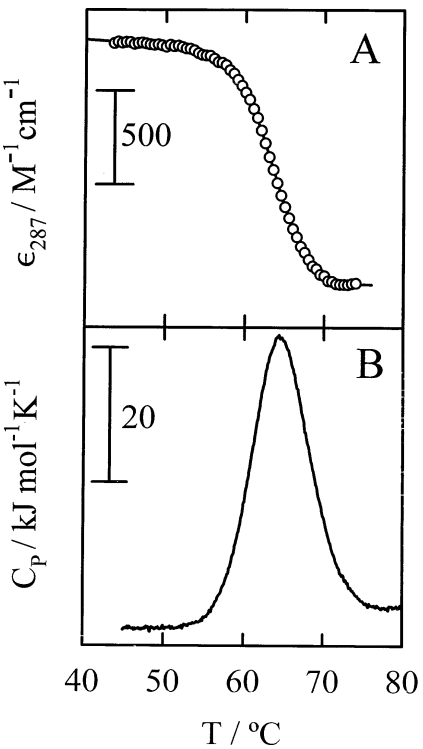

Fig. 1 Thermal unfolding of RNase (1.38 mg mL^{-1}) in 2 M glucose, pH 4.0, as monitored by UV optical absorption and DSC. (*A*) Extinction coefficient at 287 nm *vs.* temperature profile. The line represents the best fit of eqns (1)–(4) to the experimental data (○). (*B*) Experimental heat capacity data corrected for the instrumental baseline

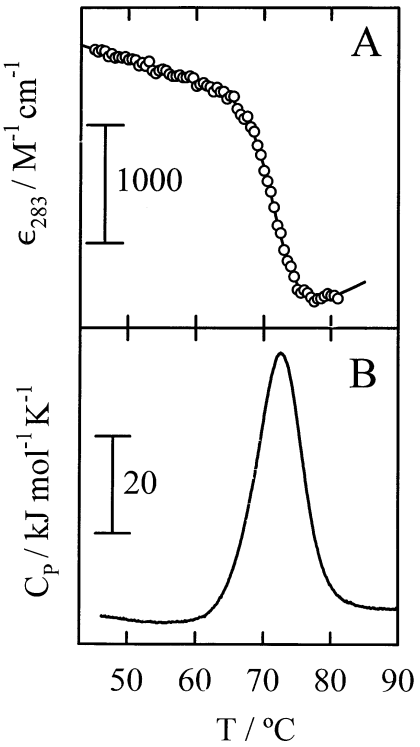

Fig. 2 Thermal unfolding of HEW lysozyme (0.33 mg mL^{-1}) in 2 M glucose, pH 2.3, as monitored by UV optical absorption and DSC. (*A*) Extinction coefficient at 283 nm *vs.* temperature profile. The line represents the best fit of eqns (1)–(4) to the experimental data (○). (*B*) Experimental heat capacity data corrected for the instrumental baseline

optical unfolding profiles. Thus, DSC experiments were carried out with both proteins in water and in water–cosolvent mixtures of the highest cosolvent concentrations employed in this work; in all cases, a pH value in the middle of the pH range of the optical studies was chosen for the DSC test. Simultaneously with any given DSC experiments, an optical unfolding profile was recorded with the same protein solution and at exactly the same heating rate (see Fig. 1 and 2 for some examples). We found an excellent agreement (Table 1) between the energetic unfolding parameters obtained from the optical unfolding profiles and those derived from the analysis of the DSC thermograms. In addition, this latter analysis supports the idea that, in the water–cosolvent mixtures studied in this work, the thermal unfolding of ribonuclease a and HEW lysozyme may be adequately described on the basis of the two-state equilibrium model.

Global analysis of the optical unfolding profiles

The values of the denaturation temperatures, unfolding enthalpy and heat capacity changes we will use in subsequent calculations were derived from global analyses based on eqn. (1)–(3), together with the following van't Hoff expression for the unfolding equilibrium constant:

$$\ln K(T, \mathrm{pH}) = -\frac{\Delta H_R - \Delta C_p \cdot T_R}{R}\left(\frac{1}{T} - \frac{1}{T_m(\mathrm{pH})}\right) + \frac{\Delta C_p}{R}\ln\left(\frac{T}{T_m(\mathrm{pH})}\right) \quad (5)$$

where $K(T, \mathrm{pH})$ is the value of the unfolding equilibrium constant at a given temperature T and pH, T_m(pH) is the denaturation temperature at that pH and ΔH_R is the unfolding enthalpy at a reference temperature, T_R. The T_R values chosen were 50 °C for ribonuclease a and 65 °C for lysozyme (in both cases, values approximately in the middle of the range defined by the experimental T_m values). For each given solvent composition, eqn. (1)–(3) and (5) were used to simultaneously fit (program MLAB) all the measured optical unfolding profiles (seven, in most cases). In this global fitting we impose that the values of ΔH_R and ΔC_p must be the same for all the unfolding profiles, while we allow each profile to have 'its own' value for T_m [eqn. (5)], α_N and β_N [eqn. (2)], α_D and β_D [eqn. (3)]. Note that this global analysis is based on the assumption that ΔC_p and the unfolding enthalpy may be taken as constants (that is, pH-independent values) for a given solvent composition. This should be an excellent approximation in this case, since protein ionization behaviour in the acidic pH range is mainly dominated by carboxylic acid groups, which show a very low ionization enthalpy. Fits were always excellent (see the examples shown in Fig. 3) and the values obtained for ΔH_R and ΔC_p are given in Table 2. We did not detect any significant cosolvent concentration effect on ΔH_R and ΔC_p and the values of ΔC_p derived from the global analyses of optical data were in good agreement with those obtained by DSC (see Table 2). Representative examples, of the T_m *vs.* pH profiles derived from the global analyses are shown in Fig. 4.

The number of protons taken from the solvent upon unfolding

The number of protons taken from the solvent upon unfolding ($\Delta\nu$) is related to the pH effect on the denaturation temperature through the following well-known equation (see, for instance, ref. 2):

$$\Delta\nu = \frac{\Delta H_m}{(\ln 10)\cdot RT_m^2}\left(\frac{\partial T_m}{\partial \mathrm{pH}}\right)_{m_3} \quad (6)$$

where ΔH_m is the denaturation enthalpy at T_m and m_3 stands for cosolvent molality. If we take the unfolding enthalpy to be pH-independent, then linkage thermodynamics indicates that we must also take $\Delta\nu$ to be temperature-independent [see eqn. (9) in ref. 2]; then, it is possible to separate temperature and pH variables in eqn. (6) and integrate at constant m_3 using the

Table 1 Comparison between the energetic unfolding parameters derived from the analysis of DSC thermograms and optical unfolding profiles

			DSC				Optical profiles	
Protein	Cosolvent	pH	T_m/°C	ΔH^{vH}/kJ mol^{-1}	ΔH^{cal}/kJ mol^{-1}	r^a	T_m/°C	ΔH^{vH}/kJ mol^{-1}
RNase	None	4.02	56.4	349	355	1.02	55.8	347
	Sucrose, 1 M	3.60	61.3	394	370	0.94	60.6	387
	Glucose, 2 M	3.54	64.2	419	380	0.91	63.9	398
	Glycerol, 4 M	3.55	59.4	371	399	1.08	59.2	361
	PEG, 25%	3.55	54.9	319	340	1.07	54.4	337
Lysozyme	None	3.01	72.5	500	496	0.99	71.1	469
	Glucose, 2 M	2.20	72.0	520	509	0.98	71.8	518
	Glycerol, 4 M	2.18	65.11	527	452	0.86	64.2	456
	PEG, 25%	2.30	54.8	452	490	1.08	54.9	403

[a] Calorimetric to van't Hoff enthalpy ratio: $r = \Delta H^{cal}/\Delta H^{vH}$.

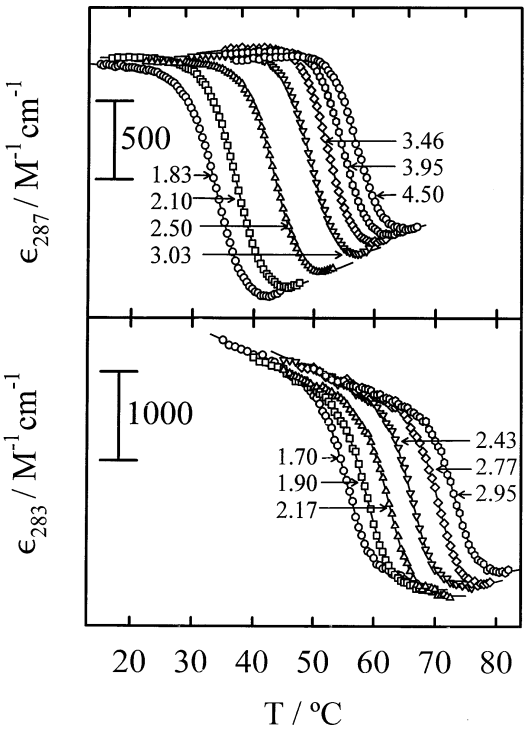

Fig. 3 Representative examples of global analyses of optical unfolding data. Upper panel: extinction coefficient *vs.* temperature profiles for RNase unfolding in 0.75 M sucrose. Lower panel: extinction coefficient *vs.* temperature profiles for HEW lysozyme unfolding in 3 M glycerol. In both cases, the numbers alongside the profiles stand for the pH values and the lines represent the best global fit of eqns (1)–(3) and (5) to the experimental data

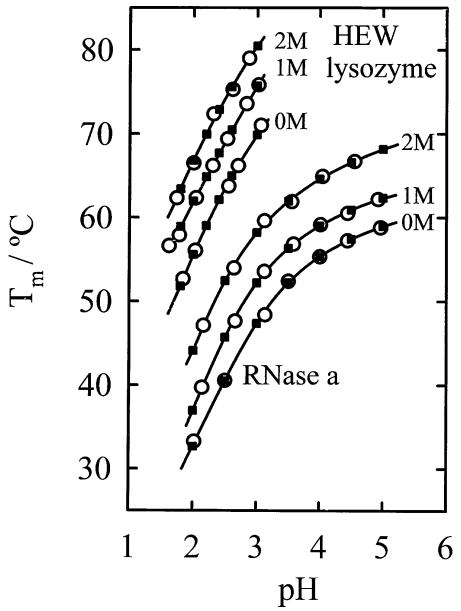

Fig. 4 Representative examples of denaturation temperature (T_m) *vs.* pH profiles for RNase and HEW lysozyme unfolding. The data shown correspond to water–glucose mixtures of the indicated glucose concentrations. (○) T_m values calculated from the global analyses of optical unfolding profiles. (■) Values of T_m interpolated at equally spaced pH values and used in the calculation of the cosolvent concentration dependence of the unfolding Gibbs energies [eqn. (9)]; the interpolation is based on the least-squares fitting of a second-order polynomial in the case of HEW lysozyme, or on eqn. (12) in ref. 2 in the case of RNase. The interpolating functions are shown as continuous lines in the figure

relation $\Delta H_m = \Delta H_R + \Delta C_p \cdot (T_m - T_R)$. The result is:

$$\int_{pH_R}^{pH} \Delta\nu \cdot dpH = -\frac{\Delta H_R - \Delta C_p \cdot T_R}{(\ln 10) \cdot R}\left(\frac{1}{T_m(pH)} - \frac{1}{T_m(pH_R)}\right) + \frac{\Delta C_p}{(\ln 10) \cdot R}\ln\left(\frac{T_m(pH)}{T_m(pH_R)}\right) \quad (7)$$

where $T_m(pH)$ and $T_m(pH_R)$ are the denaturation temperatures corresponding to any given pH and to a pH value taken as reference (pH_R), respectively. A plot of the right-hand side of eqn. (7) *vs.* pH is shown in Fig. 5. It is noteworthy that, for both proteins, there is an excellent agreement between the values calculated for different water–cosolvent mixtures. Since the left-hand side in eqn. (7) is only the pH integral of Δν, the agreement shown in Fig. 5 supports the assumption that Δν may also be taken as independent of solvent composition, at least for the solvent mixtures and composition ranges studied in this work. To conclude, we will consider Δν as a function of

Table 2 Unfolding heat capacity changes and unfolding enthalpies at the reference temperature (50 °C for RNase, 65 °C for HEW lysozyme) derived from global analyses of optical unfolding profiles[a]

Protein	Cosolvent	ΔC_p/kJ K^{-1} mol^{-1}	ΔH_R/kJ mol^{-1}
RNase	None	4.00 ± 0.70 (4.7)	339 ± 3
	Sucrose, 0.25–1 M	4.10 ± 0.45	347 ± 6
	Glucose, 0.5–2 M	4.86 ± 0.57 (3.2)	341 ± 4
	Glycerol, 1–4 M	4.75 ± 0.40 (4.6)	351 ± 6
	PEG_{8000} 10–26%	3.86 ± 0.86 (4.2)	350 ± 3
HEW lysozyme	None	6.38 ± 1.10 (5.6)	439 ± 11
	Glucose, 0.5–2 M	6.01 ± 0.83 (5.1)	463 ± 8
	Glycerol, 1–4 M	5.23 ± 1.16 (5.9)	479 ± 16
	PEG_{8000}, 10–25%	7.05 ± 0.97 (5.3)	478 ± 19

[a] Each value is the average of the results obtained from the global analysis of several (usually four) different data sets of six or seven unfolding profiles each. No significant cosolvent concentration effects were detected. Therefore, for each given cosolvent, results corresponding to different cosolvent concentrations were included in the calculation of the average value. The unfolding heat capacity changes given in parentheses were derived from the analysis of DSC thermograms carried out at the highest cosolvent concentrations employed (1 M sucrose, 2 M glucose, 4 M glycerol, 25% PEG_{8000}).

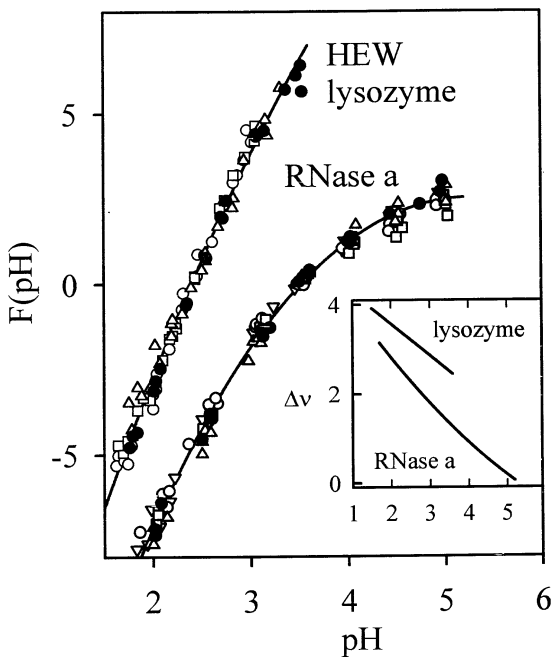

Fig. 5 Plot of the right-hand side of eqn. (7), *F*, *vs.* pH. The value of the reference pH [pH_R in eqn. (7)] was 2.5 for HEW lysozyme and 3.5 for RNase a. Data corresponding to all the water–cosolvent mixtures studied in this work are included in this figure: (●) water (no cosolvent added), (○) water–glucose, (□) water–glycerol, (△) water–PEG_{8000}, (▽) water–sucrose. The continuous lines are the best fits of second-order and third-order polynomials to the *F*/pH data for HEW lysozyme and RNase, respectively. The pH derivatives of these polynomials give the Δν (number of protons taken from the solvent upon unfolding) *vs.* pH dependencies (shown in the inset)

pH exclusively [see the inset in Fig. 5 for $\Delta\nu$ values calculated from the pH dependence of the right-hand term in eqn. (7)]. The consequences of this $\Delta\nu = f(\mathrm{pH})$ approximation have been previously discussed.[2]

The unfolding change in the protein-cosolvent preferential interaction parameter

Following Timasheff,[1] we define the protein-cosolvent preferential interaction parameter (ϕ_{23}) as a measure of the mutual perturbation of the chemical potentials of the protein (component 2) and cosolvent (component 3): $\phi_{23} = (\partial\mu_2/\partial m_3) = (\partial\mu_3/\partial m_2)$. The unfolding change in the preferential interaction parameter [$\Delta\phi_{23} = \phi_{23}(\mathrm{unfolded}) - \phi_{23}(\mathrm{native})$] may be obtained from the cosolvent effect on the Gibbs energy change:

$$\left(\frac{\partial \Delta G}{\partial m_3}\right)_{T,\,\mathrm{pH}} = \Delta\phi_{23} \qquad (8)$$

The application of eqn. (8) requires sets of ΔG values at different cosolvent concentrations, but at the same temperature and pH. These sets were obtained and processed as described below.

Denaturation temperatures for several cosolvent concentrations were calculated at equally spaced pH values (pH_i values) by interpolation in the experimental T_m/pH dependence curves (see Fig. 4). For each pH_i value, the average of several T_m values (corresponding to several cosolvent concentrations) was calculated (T_i) and ΔG values at these $\{\mathrm{pH}_i, T_i\}$ conditions and at several cosolvent concentrations were calculated by using the following integrated Gibbs–Helmholtz equation:

$$\Delta G(T_i, \mathrm{pH}_i, m_3) = \Delta H_m \cdot \left[1 - \frac{T_i}{T_m(\mathrm{pH}_i, m_3)}\right] + \Delta C_p \cdot \left[T_i - T_m(\mathrm{pH}_i, m_3) - T_i \cdot \ln\left(\frac{T_i}{T_m(\mathrm{pH}_i, m_3)}\right)\right] \qquad (9)$$

with the unfolding enthalpy at each denaturation temperature (ΔH_m) calculated as:

$$\Delta H_m = \Delta H_R + \Delta C_p \cdot [T_m(\mathrm{pH}_i, m_3) - T_R] \qquad (10)$$

where $T_m(\mathrm{pH}_i, m_3)$ is the denaturation temperature at a given cosolvent concentration (m_3) and pH value (pH_i). Note that, since T_i has been chosen as the average of the T_m values calculated for a given pH_i, eqn. (9), (10) are applied over a rather small temperature range in all cases (from T_m, at which $\Delta G = 0$, to T_i, at which we wish to calculate the ΔG value), thus minimizing extrapolation errors. Examples of the obtained ΔG *vs.* m_3 dependencies are given in Fig. 6.

In most cases, plots of ΔG *vs.* m_3 (Fig. 6) show a small curvature and values of $\Delta\phi_{23}$ at cosolvent concentrations in the middle of the studied composition ranges were calculated from the fitting of a second-order polynomial to the $\Delta G/m_3$ data (we note, nevertheless, that fitting of a straight line would yield, essentially the same results). In principle, each $\Delta\phi_{23}$ value thus calculated corresponds to a given cosolvent concentration, pH value (pH_i) and temperature (T_i). However, linkage thermodynamics shows that, if we consider $\Delta\nu$ as independent of cosolvent concentration, then we must consider $\Delta\phi_{23}$ as pH-independent [see eqn. (22) of the Supporting Information in ref. 2]. Accordingly, we will consider $\Delta\phi_{23}$ as a function of cosolvent concentration and temperature, exclusively. Fig. 7 shows the temperature dependence of the calculated $\Delta\phi_{23}$ values at constant m_3. Clearly, $\Delta\phi_{23}$ is only weakly dependent on temperature and the values at 25 °C (given in Table 3 and used in the subsequent calculation of $\Delta\Gamma_{21}$) were obtained by linear extrapolation as shown in Fig. 7.

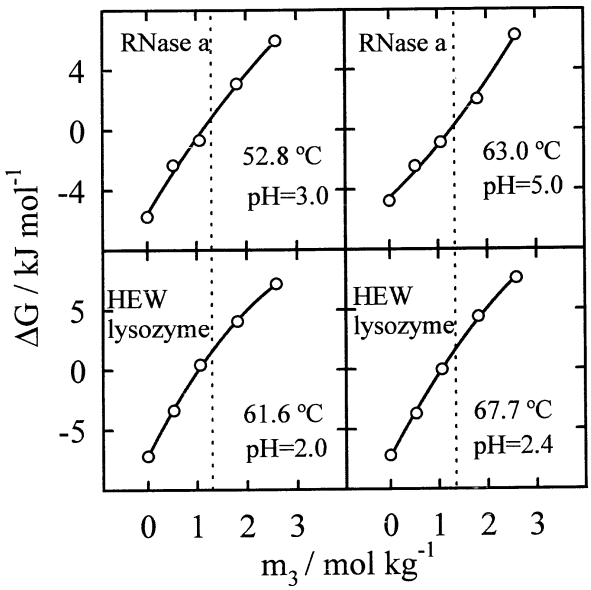

Fig. 6 Representative examples of the calculated cosolvent concentration dependencies of the unfolding Gibbs energy change. The data (○) correspond to RNase and HEW lysozyme in water–glucose mixtures at the indicated temperatures and pH values. The continuous lines represent the best fits of a second-order polynomial to the ΔG *vs.* m_3 data. The dashed lines show the cosolvent concentrations at which $\Delta\phi_{23}$ was calculated as $\partial\Delta G/\partial m_3$ [eqn. (8)]

The unfolding change in preferential hydration

$\Delta\Gamma_{21}$ values at 25 °C were obtained from the corresponding $\Delta\phi_{23}$ values at the same temperature, by using:[2]

$$\Delta\Gamma_{21} = \frac{\Delta\phi_{23}}{v_1 \cdot (\partial\Pi/\partial m_3)_T} \qquad (11)$$

where v_1 is the molar volume of water and Π is the osmotic pressure of the water–cosolvent mixture (with water as the diffusible component and the cosolvent as the non-diffusible one). We show in Appendix A that application of eqn. (11) is entirely equivalent to the osmotic stress calculation of the number of water molecules involved in the unfolding process. The required values of the derivative $(\partial\Pi/\partial m_3)_T$ were calculated from the equations given in ref. 14 in the case of PEG_{8000}, and from the fitting of a second-order polynomial

Table 3 Unfolding changes in the protein-cosolvent preferential interaction parameter $\Delta\phi_{23}$ and the preferential hydration $\Delta\Gamma_{21}$ at 25 °C[a]

	RNase		HEW lysozyme	
Cosolvent	$\Delta\phi_{23}$[b]	$\Delta\Gamma_{21}$[c]	$\Delta\phi_{23}$[b]	$\Delta\Gamma_{21}$[c]
Sucrose, 0.6 mol kg^{-1}	8.04 ± 0.23	157 ± 5	—	—
Glucose, 1.3 mol kg^{-1}	4.41 ± 0.33	92 ± 7	4.06 ± 0.10	85 ± 2
Glycerol, 2.8 mol kg^{-1}	1.43 ± 0.06	30 ± 1	1.51 ± 0.32	32 ± 7
PEG_{8000}, 15%	−0.05 ± 0.09	−14 ± 26	−0.33 ± 0.35	97 ± 104

[a] The uncertainties given are derived from the linear fit employed in the temperature extrapolation of $\Delta\phi_{23}$ to 25 °C (Fig. 7). The actual error associated with the values given in this table may be larger. [b] In kJ kg mol^{-2}. [c] In mol water per mol protein.

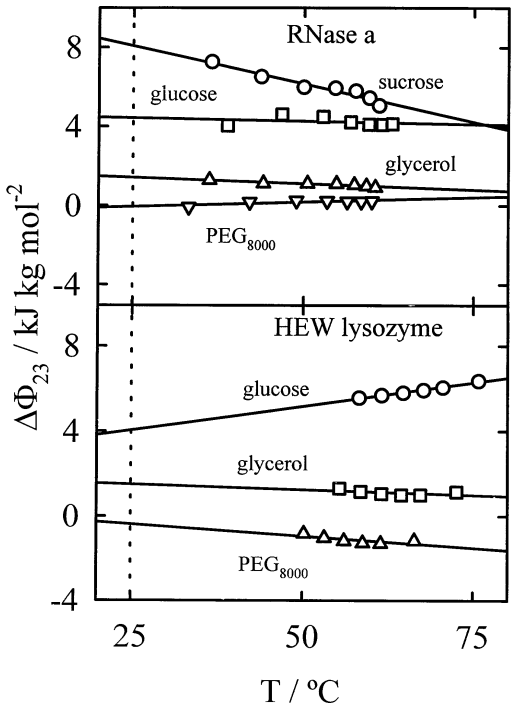

Fig. 7 Temperature dependence of the changes in preferential interaction parameter upon unfolding of HEW lysozyme and RNase in mixtures of water with the indicated cosolvents. The lines represent the least-squares straight lines used to extrapolate the values to 25 °C

to the Π values calculated by using the osmotic coefficients given in ref. 15 in the case of glycerol, glucose and sucrose. The values obtained for $\Delta\Gamma_{21}$ are given in Table 3. These values have been calculated for given cosolvent concentrations (see Table 3); however, the cosolvent concentration dependence of $(\partial\Pi/\partial m_3)_T$ is minor within the ranges studied in this work and, in addition, the curvature seen in plots of ΔG *vs.* m_3 (Fig. 6) is usually very small and does not appear in a consistent manner. It appears likely, therefore, that within the studied cosolvent concentration ranges, $\Delta\phi_{23}$ (the slope of the ΔG *vs.* m_3 plot) and $\Delta\Gamma_{21}$ [eqn. (11)] remain constant (in the values given in Table 3) within experimental uncertainty. A plausible explanation for the constancy of $\Delta\Gamma_{21}$ is offered in the Discussion. We note here, nevertheless, that the $\Delta\Gamma_{21}$ values obtained in mixtures of water with sucrose, glucose and glycerol (Table 3) are positive numbers. This suggests that the stabilizing effect of these cosolvents is related to the fact that the preferential hydration of the denaturated protein is larger than that of the native protein. Similar results have been recently reported by Xie and Timasheff[3,16,17] in connection with the stabilization of RNase by several cosolvents.

Discussion

The unfolding changes in preferential hydration for RNase and HEW lysozyme are significantly smaller than theoretical estimates of the unfolding change in the number of water molecules in direct contact with the protein surface (first monolayer). We will estimate the unfolding change in the number of water molecules in contact with the protein (ΔN_1) by using:

$$\Delta N_1 = \frac{\Delta\mathrm{ASA}}{\mathrm{ASA_w}} \qquad (12)$$

where ΔASA is the unfolding change in accessible area (the ASA of the unfolded state minus the ASA of the native state) and $\mathrm{ASA_w}$ is the accessible surface area 'covered' (on average) by a water molecule. As we explain below, however, the values calculated by using eqn. (12) depend strongly on (*i*) the model used to calculate the ASA of the unfolded state and (*ii*) the value employed for $\mathrm{ASA_w}$.

Accessible surface areas for unfolded states have often been calculated from ASA values for amino acids in tripeptide models. Rose and colleagues[7,8] have recently argued, however, that this kind of calculation overestimates the unfolded ASA; as an alternative to the tripeptide models, these authors propose the use of two limiting models that are expected to bracket the unfolded state ASA between two reliable limits. Thus, an upper limit is obtained by modelling the protein in an extended conformation, with backbone dihedrals set to $\phi = -120°$ and $\psi = +120°$ and side chain torsions set to 180°,[18] while a lower limit is derived from the accessibility of amino acids in chain segments excised in their native conformations from a data set of 43 proteins. In our calculations (Table 4) we will employ both the lower and upper bounds of the unfolded state ASA and, for illustration, the value derived from tripeptide models.

ASA values are calculated using a spherical probe of radius 1.4 Å (the van der Waals size of a molecule of water); therefore, it would appear reasonable to expect that the protein surface area covered by a water molecule is about 8 Å^2. We believe, however, that this value assumes an unrealistically good packing of water at the protein–water interface; for instance, Gerstein and Chothia[19] have estimated from the analysis of crystal structures that water molecules near protein surfaces occupy a volume of about 24.5 Å^3, which is 18% smaller than the volume of a water molecule in bulk water (30 Å^3), but much larger than the van der Waals volume (11.5 Å^3 for a 1.4 Å radius). We believe that a more realistic value for $\mathrm{ASA_w}$ may be calculated from the estimates of the hydration required to achieve monolayer coverage of native proteins derived from measurements of the non-freezing water and the effect of hydration on heat capacity (for a review on these methods, see ref. 20); these hydration values typically range from about 0.3 to 0.5 g of water per g of protein.[20] Using 0.4 g g^{-1} as a typical value, we compute that about 300 water mol-

Table 4 Theoretical estimates of the unfolding changes in accessible surface area and the number of water molecules in contact with the protein surface[a]

Model for the exposure to solvent in the unfolded state	RNase		HEW lysozyme	
	ΔASA/Å^2	ΔN_1[b]	ΔASA/Å^2	ΔN_1[b]
Tripeptides	14256	634 (1782)	15437	686 (1930)
Extended conformation (upper bound)	10498	467 (1312)	11262	501 (1408)
Fragments from folded structures (lower bound)	6422	285 (803)	7229	321 (904)

[a] ΔASA values are calculated as $\mathrm{ASA^D} - \mathrm{ASA^N}$, where $\mathrm{ASA^N}$ values were taken from ref. 21 and the $\mathrm{ASA^D}$ values for the several models[7,8] were calculated from the sequence by using programs made available by Rose and colleagues (http://cherubino.med.jhmi.edu/~folded/). ΔN_1 values are calculated by using eqn. (12) with $\mathrm{ASA_w}$ equal to 22.5 and 8 Å^2 (values in brackets). [b] In mol water per mol protein.

ecules are required to achieve monolayer coverage of native RNase a and, using 6790 Å^2 as the ASA value for native RNase[21] we obtain 22.5 Å^2 as a rough estimate of ASA_w. We have used $ASA_w = 22.5$ Å^2 in the calculation [eqn. (12)] of the 'reliable' ΔN_1 values given in Table 4; for illustration, we also include in Table 4 the 'less reliable' ΔN_1 values that would be obtained by using $ASA_w = 8$ Å^2 (values within brackets).

It is clear (see Tables 3 and 4) that the unfolding changes in preferential hydration we have obtained for RNase and lysozyme not only depend on the cosolvent used, but are also significantly smaller than all our theoretical estimates of the unfolding change in the number of water molecules in contact with the protein. The discrepancy is unlikely to be due to the possible compact character of the unfolded states, since the $\Delta\Gamma_{21}$ values are even smaller than the ΔN_1 estimates calculated by using the lower bound for the unfolded state ASA (note that the lower bound is based on peptide fragments extracted from folded protein structures and that these fragments are expected to be more compact than unfolded peptides[7,8]). In addition, the discrepancy could hardly be attributed to a strong protein–cosolvent interaction, since we have not detected significant cosolvent concentration effects on the unfolding enthalpy and heat capacity values (Table 2). As we discuss in the following section, however, it appears plausible that the low values found for the unfolding change in preferential hydration are the result of a very weak interaction of the cosolvent with the native and/or the unfolded states.

A view based on the two-domain (local-bulk) solvent model

In this section, we will discuss the values found for $\Delta\Gamma_{21}$ (Table 3) in terms of the two-domain model,[22,23] which assumes that the protein (component 2) affects the distribution of main solvent (water, component 1) and cosolvent (component 3) within its vicinity; as a result, the composition of the 'local' solvent domain that surrounds each protein molecule (*i.e.*, the solvent region affected by the protein) is different than that of the 'bulk' solvent domain that separates the local domains. On the basis of this two-domain model it is straightforward to arrive at the following relationship between preferential hydration and the composition of the local domain (see Appendix B for a simple derivation):

$$\Gamma_{21} = B_1 - \frac{m_1}{m_3} B_3 \tag{13}$$

where B_1 and B_3 are the numbers of moles of components 1 and 3 in the local domain per mole of protein, and m_3 and m_1 stand for the corresponding molalities in the protein solution (note that $m_1 = 55.5$ mol kg^{-1} = constant). According to Schellman's exchange model,[24,25] the entrance of a cosolvent molecule in the local domain would be expected to displace a certain number (S_{13}) of water molecules; accordingly, B_1 and B_3 are related by:

$$B_1 = B_1^0 - S_{13} \cdot B_3 \tag{14}$$

where B_1^0 is the value of B_1 where no molecules of the cosolvent are in the local domain (that is, when the cosolvent is completely excluded from the local domain). Eqn. (13) may then be written as:

$$\Gamma_{21} = B_1^0 - \left(\frac{m_1}{m_3} + S_{13}\right) \cdot B_3 \tag{15}$$

B_1^0 is defined as a constant; on the other hand, B_3 will depend on cosolvent concentration,[23] although the functional form of this dependence is not known. Therefore, we will take a model-free approach and assume that, for cosolvents that do not bind strongly to proteins and at low cosolvent concentration, the $B_3(m_3)$ dependence may be approximated as a Taylor expansion truncated at the linear term:

$$B_3(m_3) \simeq \left(\frac{\partial B_3}{\partial m_3}\right)_{m_3=0} \cdot m_3 = \alpha \cdot m_3 \tag{16}$$

where we have used that $B_3 = 0$ for $m_3 = 0$. We note, nevertheless, that other approaches are possible (invariant particle model,[26,27] for instance).

S_{13} is expected to be of the order of a few units for small cosolvents[23–25] and therefore, $m_1/m_3 + S_{13} \simeq m_1/m_3$ for low cosolvent concentrations. Using this result, together with eqn. (16), eqn. (15) becomes:

$$\Gamma_{21} \simeq B_1^0 - \alpha \cdot m_1 \tag{17}$$

Eqn. (13) and (17) apply to a given protein state (native or unfolded); the corresponding equations giving the unfolding change in preferential hydration are obviously:

$$\Delta\Gamma_{21} = \Delta B_1 - \frac{m_1}{m_3} \Delta B_3 \tag{18}$$

$$\Delta\Gamma_{21} \simeq \Delta B_1^0 - \Delta\alpha \cdot m_1 \tag{19}$$

where ΔB_1 and ΔB_3 are the unfolding changes in number of moles of components 1 and 3 in the local domain, $\Delta\alpha = \alpha^D - \alpha^N$, and ΔB_1^0 is the unfolding change in the number of moles of water in the local domain when the cosolvent is completely excluded. Note that ΔB_1^0 could be considered as a measure of the number of water molecules that newly interact with the protein upon unfolding in the absence of cosolvent; ΔB_1^0 is, therefore, the value we seek (see introduction). However, only $\Delta\Gamma_{21}$ is directly obtained from the experimentally measured cosolvent effects on unfolding energetics and eqn. (19) shows that $\Delta\Gamma_{21}$ may be significantly smaller than ΔB_1^0, even if $\Delta\Gamma_{21}$ remains constant within the low cosolvent concentration range (since both ΔB_1^0 and $\Delta\alpha$ are constants). Note also that eqns (17) and (19) apply rigorously in the $m_3 \to 0$ limit; therefore, the extrapolation of experimental $\Delta\Gamma_{21}$ values to zero cosolvent concentration does not necessarily yield ΔB_1^0.

In addition, a simple calculation shows that very small amounts of cosolvent in the local domain may result in a large difference between ΔB_1^0 and $\Delta\Gamma_{21}$. Thus, we assume that the size of the local domain corresponds to a monolayer and, therefore, that $\Delta B_1^0 \approx \Delta N_1$ (the unfolding change in the number of water molecules in contact with the protein surface). We also assume that ΔN_1 is about 400 mol water per mol protein for RNase (see Table 4). The value of $\Delta\Gamma_{21}$ derived from experiment is about 160 mol water per mol protein for RNase a in water–sucrose (Table 3) and eqn. (19) with $\Delta B_1^0 = 400$ yields a $\Delta\alpha$ of about 4. Now, density measurements[10] show that the preferential hydration of native RNase in water–sucrose in the range 0–1 M sucrose is about 0.46 g of water per g of protein, a value which agrees with that expected for monolayer coverage (about 0.4 g g^{-1}) and suggests that sucrose is completely excluded from the local domain in the native protein (at least if the local domain is assumed to be of the size of a monolayer). Accordingly, we take $\alpha^N \approx 0$ and $\Delta\alpha \approx \alpha^D \approx 4$. Using $\alpha^D = 4$, eqn. (16) indicates that about four molecules of sucrose enter the local domain in the unfolded state when sucrose concentration is 1 *m* (*m* = molal), and only two when $m_3 = 0.5$ *m*. This small amount of cosolvent 'bound' to the unfolded state would be responsible for the large discrepancy between $\Delta\Gamma_{21}$ (160) and ΔB_1^0 (400).

The above calculation is meant for illustration only, since we do not know the actual value of ΔB_1^0. The calculation strongly suggests, nevertheless, that the small values we have found for $\Delta\Gamma_{21}$ in several water–cosolvent mixtures (Table 3) could be the result of the presence of rather small amounts of cosolvent in the local domain of the native and/or the unfolded protein. This interpretation is consistent with the fact that no significant cosolvent concentration effects on

unfolding enthalpy and heat capacity are observed at the comparatively low cosolvent concentrations employed in this work (Table 2).

Experimental

Materials

Ribonuclease a, HEW lysozyme, sucrose, glucose, glycerol and PEG_{8000} were purchased from Sigma Chemical Co. and used without further purification. Urea was purchased from Sigma Chemical Co. and purified by passage of its aqueous solutions through an AG-501-X8(D) ion exchange resin from Biorad prior to adding the buffer components. Aqueous stock solutions of ribonuclease a and HEW lysozyme at about 20 mg mL^{-1} were prepared by exhaustive dialysis against sodium acetate–ClH buffer 50 mM, with pH values of 3.5 (ribonuclease) and 2.5 (lysozyme). Stock solutions of the cosolvents in the same buffer, but at different pH values usually spanning the ranges 2–5 (ribonuclease) and 1.8–3 (lysozyme) were prepared by weight. Protein solutions in the water–cosolvent mixtures (used in the optical and DSC denaturation experiments) were prepared by mixing adequate volumes of the stock solutions of protein cosolvent, in such a way that final protein concentrations of about 1 mg mL^{-1} (ribonuclease) and 0.4 mg mL^{-1} (lysozyme) were achieved. The pH values of these solutions were measured after mixing. Protein concentrations were determined spectrophotometrically using known values for the extinction coefficients.[28,29]

Optical denaturation profiles

Profiles of absorbance at 287 nm (ribonuclease) or 283 nm (lysozyme) *vs.* temperature were obtained with a Cary-210 spectrophotometer. Briefly, a protein solution in a given water–cosolvent mixture and at a given pH was placed in the spectrophotometer cell; the cell was thermostated by circulating water using a Julabo M5 programmable temperature bath. In most cases, the unfolding experiments involved an upwards temperature scan at a constant heating rate of 1 K min^{-1}, although, under some selected conditions (see Results), experiments were carried out at different scanning rates. The temperature was measured by using a platinum probe directly introduced in the protein solution.

Differential scanning calorimetry

DSC experiments were carried out in a DASM-4 calorimeter described by Privalov[30] with 0.47 mL cells, at a heating rate of 1 K min^{-1} and under an overpressure of 2.5 atm to avoid any degassing during the heating. Analysis of the DSC transitions was carried out as described prevously.[2]

Acknowledgements

This research was supported by Grant PB96-1439 from the DGES (Spanish Ministry of Education and Culture). H.O.H. is a recipient of a predoctoral fellowship from the 'Instituto para la Cooperacion con el Mundo Arabe' (Spanish Ministry of Foreign Affairs).

Appendix A

We will show here that the unfolding change in preferential hydration calculated by using eqn. (11) (derived in ref. 2) is equal to the number of water molecules involved in the unfolding process as defined in the membrane-free osmotic stress approach (Note that this latter number is not directly related to a structural model, but must be considered as a thermodynamic value).

Assume that, as a result of the unfolding process, a certain number of water molecules (ΔN_{OS}) enter a cosolvent-inaccessible region. Then, the unfolding Gibbs energy change will contain a contribution equal to $-\Delta N_{OS} \cdot RT \cdot \ln a_1$, where a_1 is the activity of water and ΔN_{OS} is expressed in moles of water per mole of protein. Water activity and osmotic pressure are related by $\Pi = -(RT/v_1) \cdot \ln a_1$, where v_1 is the molar volume of *pure water*. Therefore, the $-\Delta N_{OS} \cdot RT \cdot \ln a_1$ contribution is equivalent to an osmotic work $\Pi \cdot \Delta V_W$ (where $\Delta V_W = v_1 \cdot \Delta N_{OS}$ is the volume of *pure water* corresponding to ΔN_{OS}) and the dependence of ΔG with osmotic pressure is given by:

$$\left(\frac{\partial \Delta G}{\partial \Pi}\right)_T = \Delta V_W = v_1 \cdot \Delta N_{OS} \tag{A1}$$

which is equivalent to eqn. (A17) in ref. 9. The derivative in the left-hand side of eqn. (A1) may be expressed as:

$$\left(\frac{\partial \Delta G}{\partial \Pi}\right)_T = \frac{(\partial \Delta G/\partial m_3)_T}{(\partial \Pi/\partial m_3)_T} = \frac{\Delta \phi_{23}}{(\partial \Pi/\partial m_3)_T} \tag{A2}$$

and combining eqn. (A1) and (A2) we obtain:

$$\Delta N_{OS} = \frac{\Delta \phi_{23}}{v_1 \cdot (\partial \Pi/\partial m_3)_T} \tag{A3}$$

Finally, comparison of eqn. (A3) with eqn. (11) shows that:

$$\Delta N_{OS} = \Delta \Gamma_{21} \tag{A4}$$

Of course, the definition we have used for ΔN_{OS} is a very simple one. A more realistic interpretation of ΔN_{OS} (and of $\Delta \Gamma_{21}$) is provided by the two-domain model[22,23] as embodied in eqn. (18); see also ref. 9.

Appendix B

We will provide here a simple derivation for the two-domain expression [eqn. (13)] of the preferential hydration.

The preferential hydration of the protein is the amount of water that must be added to the system, upon protein addition, in order to maintain constant the water chemical potential. The chemical potential of water is an intensive property and may be considered as a function of temperature, pressure and *solvent composition in the bulk solvent domain.*[23] We will describe bulk solvent composition in terms of the water–cosolvent mol ratio (χ). Thus for a mixture of N_1 molecules of water and N_3 molecules of cosolvent this ratio is:

$$\chi \text{ (before protein addition)} = \frac{N_1}{N_3} \tag{B1}$$

Assume now that a protein molecule is added to the system. As a result, B_1 molecules of water and B_3 molecules of cosolvent are removed from the bulk solvent in order to make up the local domain. Therefore, after protein addition the composition of the bulk solvent is:

$$\chi \text{ (after protein addition)} = \frac{N_1 - B_1}{N_3 - B_3} \tag{B2}$$

However, Γ_{21} molecules of water must also be added along with the protein if the water chemical potential is to be kept constant. After addition of protein and water, the composition of the bulk domain is:

$$\chi \text{ (after addition of protein plus water)} = \frac{N_1 - B_1 + \Gamma_{21}}{N_3 - B_3} \tag{B3}$$

The condition of constant water chemical potential implies constant bulk solvent composition. That is, the compositions given by eqn. (B1) and (B3) must be equal:

$$\frac{N_1}{N_3} = \frac{N_1 - B_1 + \Gamma_{21}}{N_3 - B_3} \tag{B4}$$

and solving for Γ_{21} we obtain:

$$\Gamma_{21} = B_1 - \frac{N_1}{N_3} B_3 \qquad \text{(B5)}$$

which is equivalent to eqn. (13), since $N_1/N_3 = m_1/m_3$.

References

1 S. N. Timasheff, *Annu. Rev. Biophys. Biomol. Struct.*, 1993, **22**, 67.
2 I. M. Plaza del Pino and J. M. Sanchez-Ruiz, *Biochemistry*, 1995, **34**, 8621.
3 G. Xie and S. N. Timasheff, *Protein Sci.*, 1997, **6**, 211.
4 B. Lee and F. M. Richards, *J. Mol. Biol.*, 1971, **55**, 379.
5 E. Freire, *Annu. Rev. Biophys. Biomol. Struct.*, 1995, **24**, 141.
6 G. I. Makhatadze and P. L. Privalov, *Adv. Protein Chem.*, 1995, **47**, 307.
7 T. P. Creamer, R. Srinivasan and G. D. Rose, *Biochemistry*, 1995, **34**, 16245.
8 T. P. Creamer, R. Srinivasan and G. D. Rose, *Biochemistry*, 1997, **36**, 2832.
9 V. A. Parsegian, R. P. Rand and D. C. Rau, *Methods Enzymol.*, 1995, **259**, 43.
10 J. C. Lee and S. N. Timasheff, *J. Biol. Chem.*, 1981, **256**, 7193.
11 T. Arakawa and S. N. Timasheff, *Biochemistry*, 1982, **21**, 6536.
12 K. Gekko and S. N. Timasheff, *Biochemistry*, 1981, **20**, 4667.
13 R. Bhat and S. N. Timasheff, *Protein Sci.*, 1992, **1**, 1133.
14 V. A. Parsegian, R. P. Rand, N. L. Fuller and D. C. Rau, *Methods Enzymol.*, 1986, **127**, 400.
15 R. H. Stokes and R. A. Robinson, *J. Phys. Chem.*, 1966, **70**, 2126.
16 G. Xie and S. N. Timasheff, *Protein Sci.*, 1997, **6**, 222.
17 G. Xie and S. N. Timasheff, *Biophys. Chem.*, 1997, **64**, 25.
18 R. S. Spolar, J. R. Livingstone and M. T. Record, *Biochemistry*, 1992, **31**, 3947.
19 M. Gerstein and C. Chothia, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 10167.
20 R. B. Gregory, in *Protein–Solvent Interactions*, ed. R. B. Gregory, Marcel Dekker, New York, 1995, pp. 191–264.
21 S. Miller, J. Janin, A. M. Lesk and C. Chothia, *J. Mol. Biol.*, 1987, **196**, 641.
22 H. Inouye and S. N. Timasheff, *Biopolymers*, 1972, **11**, 737.
23 M. T. Record and C. F. Anderson, *Biophys. J.*, 1995, **68**, 786.
24 J. A. Schellman, *Biophys. Chem.*, 1990, **37**, 121.
25 J. A. Schellman, *Biopolymers*, 1994, **34**, 1015.
26 H. Eisenberg, *Biophys. Chem.*, 1994, **53**, 57.
27 A. Tardieu, P. Vachette, A. Gulik and M. Le Maire, *Biochemistry*, 1981, **20**, 4399.
28 R. E. Canfield, *J. Biol. Chem.*, 1963, **238**, 2691.
29 J. F. Brandts and L. N. Lin, *Biochemistry*, 1990, **29**, 6927.
30 P. L. Privalov, *Pure Appl. Chem.*, 1980, **52**, 479.

Received in Montpellier, France, 10*th March* 1998; *revised m/s received* 3*rd July* 1998; *Paper* 8/01992D