ELSEVIER

Biophysical Chemistry 59 (1996) 259–275

Biophysical Chemistry

The enthalpy of transfer of unfolded proteins into solutions of urea and guanidinium chloride

John A. Schellman *, Nadine C. Gassner

Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, USA

Received 24 May 1995; accepted 24 July 1995

Abstract

G. Makhatadze and P. Privalov [J. Mol. Biol., 226 (1995) 491] have recently measured the enthalpy of transfer of three proteins into urea and guanidinium chloride solutions as a function of concentration and temperature. The present paper applies the solvent-exchange model [J.A. Schellman, Biopolymers, (1994)] to the data and compares it with the binding model utilized in the original publication. Both calculations assume identical binding sites. It is found that the data may be fit tolerably well using either procedure, but that the parameters describing the binding vary considerably. Consideration of the transfer properties of amino acid moieties and small peptides leads to the conclusion that solvation sites are heterogeneous and that the quantities determined by both methods are statistical averages. The parameters describe an identical-site system that has (approximately) the same properties as the real heterogeneous system. The results have mainly heuristic and mechanistic value. One quantity determined with these simplified isotherms, $\Sigma k_j \Delta h_j$, is a property of the real system and can serve as a measure of a thermal binding capacity for a protein. The appendices contain a resume of the solution theory required for the exchange model of solvation as well as the development of a number of empirical equations for the thermodynamic properties of urea and guanidinium chloride solutions.

Keywords: Protein unfolding; Urea; Guanidinium chloride; Solvation; Denaturation; Cytochrome *c*; Lysozyme; Ribonuclease A

1. Introduction

There are important problems which linger in the forefront of scientific inquiry for decades. One of these is the interaction of solvent with biological macromolecules. When one of the authors (JAS) was working very closely with Bill Harrington about 40 years ago, a continuing topic of discussion was the interactions of denaturing agents, not only with proteins, but with water, the effects of these interactions on "water structure", and ultimately the effect of water structure on the stability of macromolecules. Bill would probably have approved of the fact that water, in the form of its activity and activity coefficient, is an active partner in the discussion of the denaturant reactions in this paper. On the other hand we have not yet reached the level of tangible, physical interpretation that was his constant goal in research.

2. Background

Three years ago G. Makhatadze and P. Privalov published very extensive data on the enthalpy of

* Corresponding author.

SSDI 0301-4622(95)00130-1

interaction of three proteins with urea and guanidinium chloride in concentrated aqueous solution [1]. (In the following this work will be referred to by the initials M&P.) Studies of this kind are clearly of importance but had been long delayed because of the intimidating problems caused by the large heats of dilution of these substances. In the original paper the results were interpreted by a standard model for multi-site binding. In the course of discussions with them about their results one of us (JAS) suggested that the solvent-exchange model [2,3] might be more appropriate for the interpretation of the experiments. P. Privalov very generously sent us the data prior to its publication and suggested that we provide this alternative interpretation either as co-author or in a separate publication. In this paper we are very belatedly following the second course of action.

3. The procedure of M&P

Three proteins were investigated: hen egg-white lysozyme, ribonuclease A and cytochrome *c*. These were studied in their native forms and in their denatured forms. Apocytochrome *c* was used as a model for denatured cytochrome *c*; the denatured forms of lysozyme and ribonuclease were simulated by reducing and blocking the SS bridges. All experimental details can be found in the original publication. Since this paper is mainly a comparison of models of interpretation, it will be restricted to the unfolded forms of these proteins where the data is not complicated by the unfolding reaction, though this "complication" was a main focus of interest in the study.

The major measurement of the paper was the enthalpy of transfer of the proteins from aqueous solution to solutions of guanidinium chloride and urea. This is the enthalpy of transfer or the excess enthalpy, $\overline{H}^{ex}$, a component of the total enthalpy which is given by

$$\overline{H} = \overline{H}^{\circ} + \overline{H}^{ex} \tag{1}$$

where $\overline{H}^{\circ}$ is the molar enthalpy in an ideally dilute solution.

The thermodynamics of both urea and guanidinium chloride solutions have been studied extensively. The data which is required for the interpretation of the experiments in urea solutions is to be found in the review of Stokes [4] and the paper of Egan and Luff [5]. For guanidinium chloride the sources are the studies of Makhatadze et al. [6] and Schrier and Schrier [7].

In the original paper the data was represented by a model with independent binding sites by means of the relation

$$\overline{H}^{ex} = \sum_j \Delta h_j \theta_j = \sum_j \Delta h_j \frac{k_j\{C\}}{1 + k_j\{C\}} \tag{2}$$

where θ_j is the fraction of occupation of site j by the added reagent (urea or guanidinium chloride), given by the standard formula

$$\theta_j = \frac{k_j\{C\}}{1 + k_j\{C\}} \tag{3}$$

and k_j is the intrinsic binding constant at site j. Since this paper will deal with several kinds of activities, we will use the symbol $\{x\}$ to represent an activity where x represents the concentration scale. Thus $\{C\}$ is the activity on the molar scale. Δh_j is the enthalpy of the exchange reaction described by Eq. 1. It is not the standard enthalpy, Δh_j°, which refers to the reaction at infinite dilution. In one approximation it is given as

$$\Delta h_j = \Delta h_j^{\circ} - \overline{L}_3 \tag{4}$$

where $\overline{L}_3$ is the partial relative enthalpy of the cosolvent at the ambient concentration. This correction factor was recognized, but ignored, by M&P. It turns out to be troublesome and will be discussed later.

To proceed further it is assumed that the protein interactions may be represented by an average over the sites by means of the formula

$$\overline{H}^{ex} = \frac{n\Delta hk\{C\}}{1 + k\{C\}} \tag{5}$$

Here k, Δh and n are average values for the binding constant, enthalpy and number of sites, which provide a representation of the data. The significance of Eq. 5, which is often quite accurate, for systems which are actually heterogeneous, has recently been discussed at some length [2]. Since the activities are known as a function of concentration, an experimen-

tal isotherm can be solved for the best values of k and $n\Delta h$ at any given temperature.

The procedure of M&P was to obtain values of k and $n\Delta hk$ at three temperatures and then determine Δh by a Van't Hoff plot of $\ln k$ vs. $1/T$. With Δh and k known, n is then calculated from $n\Delta h$. Their results are discussed below.

4. The solvent-exchange model

This model also assumes a denumerable set of independent sites. The main difference is that the site reaction takes explicit account of the role of the principle solvent

$$S1 + 3 \rightarrow S3 + 1 \tag{6}$$

where S1 represents a site on the protein molecule which is occupied by a "1" molecule (the principal solvent) and S3 represents occupation by a "3" molecule (cosolvent, urea or guanidine). For very weak binding and high concentrations of cosolvent, this model differs qualitatively from the standard binding model for properties like the free energy and preferential interaction. These properties depend not on the extent of occupation by the cosolvent, but on the excess occupation over that in the bulk solution. The enthalpy, on the other hand, is directly related to occupation. Contacts between a protein and contiguous molecules contribute to the energy whether they are in excess or not. Consequently, the formula is rather similar to the simple binding model

$$\begin{aligned}\bar{H}^{ex} &= \sum_j \left\{ \left(\Delta h_j^\circ \theta_j - \bar{L}_3 \theta_j + \bar{L}_1 (1 - \theta_j) \right) \right\} \\ &= \sum_j \left\{ \Delta h_j \theta_j + \bar{L}_1 (1 - \theta_j) \right\} \end{aligned} \tag{7}$$

L_1 is the partial molar relative enthalpy of the principal solvent, and Δh_j has the same meaning as in Eq. 4. The $\bar{L}_1$ correction is generally less than 5% of the $\bar{L}_3$ correction and will be ignored in the remainder of the paper, mainly because it complicates the non-linear least-square determination of the binding parameters. A detailed discussion of the physical significance of the $\bar{L}_1$ and $\bar{L}_3$ terms has been given [2]. Thus the solvent-exchange model has the same form as the simple binding model, except that the fractional occupation is now given by

$$\theta_j = \frac{K_j \{ \chi_3 \}}{\{ \chi_1 \} + K_j \{ \chi_3 \}} \tag{8}$$

where $\{\chi_1\}$and $\{\chi_3\}$ are activities of the principal and co-solvent on the mole fraction scale, and K_j is the equilibrium constant on the mole fraction scale for the exchange reaction, Eq. 6. It may be shown that the activity on the mole fraction scale is related to the activity on the molarity scale by the relations

$$\{ \chi_3 \} = \{C\} / (m_1 d_1^\circ)$$

where m_1 is the number of moles of principal solvent per kilogram and d_1° is the density of pure solvent. The factor in parentheses has essentially the constant value of 1/55.5 1/mole at all reasonable temperatures. Thus, if K_j is defined as 55.5 k_j, we have $K_j\{\chi_3\} = k_j\{C\}$, and Eq. 8 is identical to Eq. 3 except for the first term in the denominator. This introduces changes in the denominator of up to 20% for the most concentrated guanidinium chloride solutions, but only 7% for the urea solutions. We expect therefore to see results for the exchange model which are closer to the binding model for urea than for guanidinium chloride. At any rate we are not dealing with drastic changes in interpretation. By contrast, calculations of $\bar{G}^{ex}$ and the preferential interaction are strongly affected by the presence of the $\{\chi_1\}$ term. If we represent the interaction by n equivalent sites, we obtain the exchange analog of Eq. 5:

$$\bar{H}^{ex} = \frac{n\Delta hK\{ \chi_3 \}}{\{ \chi_3 \} + K\{ \chi_3 \}} \tag{9}$$

5. Solution thermodynamics

The solvent-exchange model has seen little application by experimentalists. One reason for this is that it is inconvenient relative to simpler models. The theory requires activities on the mole fraction scale, the experimentalist is usually using the molarity scale, and all the relevant thermodynamic data on solutions is published on the molality scale. In order to remove this barrier, at least for urea and guanidinium chloride solutions, we have included appen-

dices in which empirical formulas are developed which permit one to convert molarities directly into the quantities required to utilize Eq. 8 and other solvent-exchange formulas. Appendix A contains a summary of all the necessary relations and definitions; Appendix B demonstrates the procedures actually used and provides empirical formulas for urea; Appendix C discusses changes that are required for ionic cosolvents and provides the constants for guanidinium chloride.

Essentially, the formulas in the Appendices give f_1, f_3, χ_3, and $\bar{L}_3$ as power series in C, the molarity. From these one calculates $\chi_1 = 1 - \chi_3$, $\{\chi_1\} = f_1 \chi_1$, and $\{\chi_3\} = f_3 \chi_3$. These are all the quantities which are required for the theory. Note that for ionic substances f_3 is the mean ionic activity and one must use the relation $\chi_1 = 1 - \nu\chi_3$, where ν is the number of ions produced by dissociation [8] (see Appendix C).

6. Results

The data used for the analyses were obtained entirely from Table 2 of the paper of Makhatadze and Privalov. [1] "Lys", "rib" and "cyt" used in the text and in Tables 1 and 2 refer to the *unfolded* forms of hen egg-white lysozyme, ribonuclease A and cytochrome *c*, respectively. These protein preparations are described in the original paper. Data sets were prepared with $\{\chi_1\}$ and $\{\chi_3\}$ as "independent" variables with Q from the M&P paper as the dependent variable. Values for $\{\chi_1\}$ and $\{\chi_3\}$ (and $\bar{L}_3$ when needed) were computed from the molar concentration of denaturant with the empirical equations given in the appendices. The fact that $\{\chi_1\}$ and $\{\chi_3\}$ are not really independent of one another does not matter for the least-square analysis which treats these quantities as constants. M. Johnson of the University of Virginia supplied the non-linear least-square program as well as advice on parts of the analysis.

6.1. Proteins in urea solutions

The proteins were analyzed in two ways: without and with the explicit introduction of the relative partial molar enthalpy, $\bar{L}_3$ (Eq. 4). In the former case the parameter Δh in the numerator was made equal to Δh°. Since Δh in fact varies because of the $\bar{L}_3$ term, the resulting parameters are averaged over this variation. This is the analysis that closely parallels that of M&P and since $K_j\{\chi_3\} = k_j\{C\}$ for all practical purposes, it differs from their analysis only by the presence of $\{\chi_1\}$ in the denominator. For calculations including the non-ideality correction, $\bar{L}_3$, Eq. 9 was used with Δh given by Eq. 4.

Following the ideas of the original paper of M&P, we used the temperature variation of the isotherms to obtain a value for the enthalpy of interaction. At first we used their procedure. Values of K were obtained

Table 1
Unfolded proteins in urea solution [a]

		k (M^{-1})	Δh° (kJ/mole)	n	$nk\Delta h$ (kJ M^{-1}/mole)	K	$n\Delta h$ (kJ/mole)	w
	lys	0.059	−9	230	−120	3.3	−2030	
M&P	rib	0.052	−8	240	−106	2.9	−2030	
	cyt	0.071	−9	142	−99	3.9	−1390	
	lys	0.065	−13	134	−112	3.6	−1715	0.0206
$\bar{L}_3 = 0$	rib	0.064	−15	106	−101	3.6	−1569	0.00906
	cyt	0.081	−17	67	−93	4.5	−1139	0.00472
	lys	0.042	−12	225	−115	2.3	−2732	0.0173
$\bar{L}_3 \neq 0$	rib	0.044	−14	165	−102	2.4	−2323	0.00827
	cyt	0.061	−16	93	−92	3.4	−1508	0.00489

[a] Note that when $\bar{L}_3$ is assumed zero, $\Delta h^\circ = \Delta h$. Otherwise $\Delta h^\circ = \Delta h + \bar{L}_3$ (Eq. 4).

by least-square analysis of the isotherm (Eq. 9). The Van't Hoff equation was then applied to the values of $\ln K$ at the three temperatures. We found this rather unsatisfactory. The plots were quite non-linear and in the case of lysozyme, $\ln K$ was not even a monotonic function of T.

For this reason a somewhat different strategy was used to analyze the temperature variation. Rather than determine binding constants for each temperature and then analyze the temperature dependence of the k's with Van't Hoff plots, we combined the data for all three temperatures and performed non-linear least-squares on all the data for a given protein. The use of a Van't Hoff plot to determine the enthalpy is equivalent to the assumption that K (or k) may be represented by a function of the form

$$K = w\exp(-\Delta h^\circ/RT) \tag{10}$$

Δh° is used in this equation because K is related to the standard state free energy. There is most likely a variation of Δh° with temperature, but this should be small compared to Δh° for a site interaction and in any case is too refined a quantity to be extracted from the present analysis. $\bar{H}^{ex}$ can then be written in the general form

$$\bar{H}^{ex} = \frac{n\Delta h w\exp(-\Delta h^\circ/RT)\{\chi_3\}}{\{\chi_1\} + w\exp(-\Delta h^\circ/RT)\{\chi_3\}} \tag{11}$$

with parameters Δh°, w and n. In the analyses where $\bar{L}_3$ is ignored, $\Delta h = \Delta h^\circ$ in the numerator. When $\bar{L}_3$ is introduced explicitly, the substitution $\Delta h = \Delta h^\circ - \bar{L}_3$ is made, and $\bar{L}_3$ is calculated for each concentration using the empirical formula given in the Appendix.

The complete urea results are shown in Table 1. The top set of data are the results calculated in the original paper, the lower two sets deal with our results with and without the correction for the partial relative enthalpy. In the absence of data on the temperature variation of $\bar{L}_3$ for urea, the temperature coefficient was ignored and the values for 25°C [5], discussed in Appendix B, were used for all three temperatures. In general the differences amongst the three sets are relatively uniform shifts in the parameters. Notable is the fact that the magnitude of the enthalpy appears to be higher for the newer calculations than for the original ones. This presumably has to do with the variation of $\{\chi_1\}$ with temperature in the current work, which was absent in the original. We note also that the apparent numbers of sites for the $\bar{L}_3 = 0$ set are lower than the M&P results, but this is compensated by higher values of $-\Delta h$ and K. Further discussion will be deferred till after the guanidinium chloride results have been presented.

6.2. Proteins in guanidinium chloride solutions

Equivalent results for the proteins in guanidinium chloride are given in Table 2. Most noticeable is the finding that, whereas the binding model yields binding constants for guanidinium chloride which are ten times those for urea, this factor is reduced to about three for the exchange model. This is compensated for by an increase in the number of sites. We are somewhat hesitant to interpret these changes since the calculations demonstrated that the enthalpy effect associated with the non-ideality of the solutions is

Table 2
Unfolded proteins in guanidinium chloride solution [a]

		k (M^{-1})	Δh° (kJ/mole)	n	$nk\Delta h$ (kJ M^{-1}/mole)	K	$n\Delta h$ (kJ/mole)	w
	lys	0.69	−10	82	−566	38.3	−820	
M&P	rib	0.60	−11	74	−492	33.3	−820	
	cyt	0.51	−12	56	−352	28.3	−690	
Exchange $\bar{L}_3 = 0$	lys	0.186	−10	193	−345	10.3	−1857	0.212
	rib	0.149	−12	169	−303	8.3	−2033	0.0643
	cyt	0.094	−6	370	−213	5.2	−2272	0.437

[a] Note that when $\bar{L}_3$ is assumed zero, $\Delta h^\circ = \Delta h$. Otherwise $\Delta h^\circ = \Delta h + \bar{L}_3$ (Eq. 4).

quite large and has a large temperature dependence. In fact $\bar{L}_3$ is as high as 8 kJ at 10°C and 7 M and varies by a factor of more than three between 10 and 40°C at low concentrations where it is dominated by long-range Coulombic effects [6]. Thus at the highest concentrations, $\bar{L}_3$ has almost the magnitude of $\Delta h°$. The reason that the calculations which include $\bar{L}_3$ are not given in Table 2 is that the non-linear least-square program failed to converge. Evidently the non-ideality is so strong that it masks the concentration dependence of the binding isotherm.

It has been pointed out earlier that the non-ideality correction of Eq. 4 is too large [2]. This can be seen by considering how these terms arise. $\bar{L}_3$ and $\bar{L}_1$ are the ideality corrections for the transfer of the free components 3 and 1 from the standard state to the ambient high concentration of guanidinium chloride. A thermodynamic cycle clearly demonstrates the origin of this correction term [2]. The $\bar{L}_3$ and $\bar{L}_1$ factors arise naturally because they are known standard properties of solutions. On the other hand, the transfer of S3 and S1, i.e. site-bound molecules, is assumed to occur with no change resulting from non-ideality. This is because there is no data and no known way of obtaining it. It certainly must be true that an external guanidinium or water molecule bound to a protein site interacts with the solvent medium since it is in contact with it. One might postulate structures and compare the area of solvent exposure of a bound molecule with that of a free molecule and scale down $\bar{L}_3$ and $\bar{L}_1$ proportionately. We arbitrarily tried scaling $\bar{L}_3$ down by a factor of one half. With this change the program converged leading to a large number of sites: 510, 480 and 640, respectively, for the unfolded forms of lysozyme, ribonuclease and cytochrome *c*. We have concluded that applying the non-ideality correction to the guanidinium chloride data is not practicable at the present time.

7. Discussion

A principal question about the results in Tables 1 and 2 is the extent to which they can be used for a molecular interpretation of the solvation of proteins by denaturing agents. Before going further it will be useful to discuss information which has been gained in the study of small molecules.

Extensive studies of the free energy of transfer in denaturant solutions have been published for amino acids, peptides and molecules related to amino acid residues [9–11]. A few papers have also been directed to the enthalpy of transfer as well. The picture that emerges is that urea and the guanidinium ion interact with every group of every amino acid. Valine differs from leucine, serine from alanine, asparagine from glutamine, etc. For the free energy studies there has been some progress in breaking down the interactions further into basic groups such as $-CH_2$, $-CH_3$, $-OH$, $-C_6H_5$, etc. On the other hand, there has as yet been no successful attempt to predict the transfer properties of amino acid moieties from a more basic set of group transfers. It does seem certain that there are a variety of interactions: polar groups, non-polar groups, aliphatic and aromatic hydrophobic groups, hydrogen-bond donors and acceptors, positive and negative ions and a resulting variety of strengths of interaction.

The conclusion from these facts is that every contact area which can be occupied by a solvent molecule is a potential site for solvation and that these sites are heterogeneous in their binding properties. Consequently, the total number of sites should be at least several times the number of amino acids in the protein.

We now consider the results in Tables 1 and 2.

n: For urea the number of sites obtained in the M&P analysis is about twice the number of amino acids but falls below one per amino acid for guanidinium. The exchange model does the opposite. The sites per amino acid in guanidinium chloride is about 1.5–3 but is less than one in urea.

$\Delta h°$: The enthalpies for both reagents and both methods of analysis appear to be rather high. Experimental estimates of the enthalpy of interaction of these reagents with polar groups have cycled between 0 and -8 kJ over the past 40 years and the issue is still not settled. There is less known about the enthalpy of transfer of hydrophobic groups. The main sources of information are the transfer studies of Bull et al. [12], which refers only to 1 M guanidinium chloride solution, and papers based on the thermodynamics of unfolding [13,14]. These numbers are generally considerably smaller than 8 kJ.

K: The values for the association constant are typical of the weak binding which has been found in studies of the interaction of urea and guanidinium with smaller molecules.

Two other features of the results are (1) the rather strong variability in the parameters which results from a relatively small change in model and (2) consistent deviations of the model from the experimental results. This is illustrated in Fig. 1, which depicts the experimental results for unfolded lysozyme at the three temperatures together with the curves calculated from the parameters listed in line four of Table 1. In most instances the experimental data falls below the theoretical curve at both high and low concentrations and above it at intermediate concentrations like the 10°C curve of the figure. The effect is much more pronounced for the guanidinium fits.

We believe that many of the above factors result from using isotherms derived for identical sites for systems which are heterogeneous. In a previous study [2] it has been shown by numerical simulation that the enthalpy of interaction of a heterogeneous system can be made to fit an isotherm derived for identical sites by a suitable choice of parameters provided that the heterogeneity is not too great. Analysis shows that the binding isotherms that we have been dealing with (Eqs. 4 and 9) are described by the two-parameter equation which for Eq. 9 takes the form

$$\bar{H}^{ex} = \frac{(\Sigma K_j \Delta h_j)\{\chi_3\}}{\{\chi_3\} + K_{eff}\{\chi_3\}} \tag{12}$$

where the first parameter, K_{eff}, is an effective exchange constant and $\Sigma K_j \Delta h_j$, summed over all sites, is the second. It is a global measure of the enthalpy propensity of the interaction. Sites with high enthalpy and low binding, or the converse, contribute little to the enthalpy. It is the combination of both K and Δh which is required for effective contributions. K_{eff} is not a simple average. It is given by

$$K_{eff} = \frac{\Sigma K_j^2 \Delta h_j}{\Sigma K_j \Delta h_j} \tag{13}$$

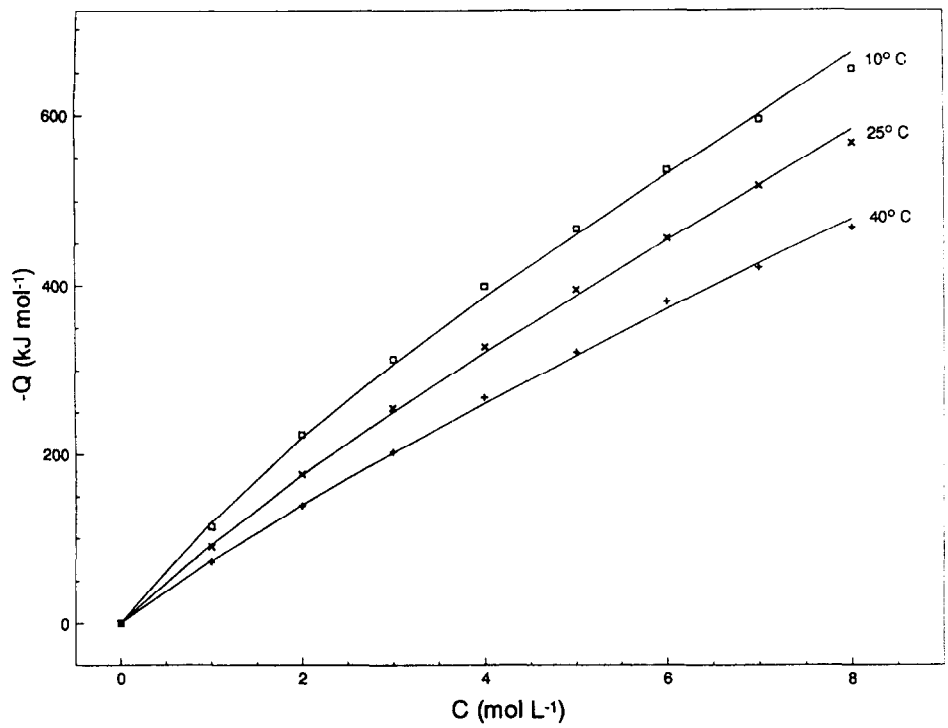

Fig. 1. Experimental and calculated excess enthalpy of unfolded ribonuclease A in urea solution at three temperatures.

In general it is not possible to obtain either an effective value for Δh nor for n without performing additional experiments. M&P were able to extract further information by doing experiments at three temperatures and making use of the Van't Hoff equation. A strict analytical interpretation of the meaning of Δh_{eff}, which can be obtained by applying the Van't Hoff relation, $R\partial \ln K_{\text{eff}}/\partial(1/T) = -\Delta h_{\text{eff}}$, to Eq. 13, gives an expression which is too complicated to be useful.

If we consider that the parameters are statistical averages over heterogeneous sites, a number of the trends and anomalies of Tables 1 and 2 are easier to understand. A possible reason that n is smaller than expected is the fact that in a heterogeneous distribution the isotherm is dominated by the sites with high $K_j\Delta h_j$ and this number is presumably considerably less than the total number of solvation sites. The fact that the experimental data has a more pronounced curvature than can be accommodated by the theory would indicate that different sites are titrated as the concentration increases. At low concentration the sites with the highest $K_j\Delta h_j$ dominate the binding. As the concentration increases other sites, with lower values of $K_j\Delta h_j$, become operational while the low-concentration sites become saturated or partially saturated. In this way the effective value of $\Sigma K_j\Delta h_j$ will gradually diminish at higher concentration leading to a downward curvature. Finally, if the parameters are statistical averages, one can expect them to be sensitive to the changes in weighting caused by changes in model and non-ideality corrections.

We have expressed the view earlier [2] that using an identical-site model to fit the data for the interaction of proteins with solvent components (i.e., weak binding and high concentrations) provides a representation of the data rather than a molecular interpretation. *What is obtained are the parameters for an identical-site system which behaves like the system under consideration.* There is some utility in this process. For example, the free energy of interaction of some proteins with guanidinium chloride [15] and urea derivatives [16] goes through a minimum as the concentration is increased. This minimum is associated with a zero preferential interaction at the same concentration. For other reagents which stabilize proteins, the free energy goes through a maximum [17]. These types of results cannot be reproduced by the standard binding model (Eq. 4) but are a natural result of the exchange model which shows that the effects depend on the non-ideality of the solutions at high cosolvent concentration. For this reason Eq. 9 gives a preferred view of the solvation process and if it applies to a system with identical sites, there is no reason why it should not apply to one with heterogeneous sites. By studying the behavior of a representative system, results are obtained which can be applied to real proteins. But these results are qualitative and refer to mechanism rather than details of the interaction.

The parameters which are obtained from an identical-site analysis have some utility. Analysis of the statistics shows that the parameter in the numerator of both Eqs. 4 and 9 is $\Sigma K_j\Delta h_j$. This is a directly determined global quantity which applies to the real heterogeneous protein. As mentioned above, it is a measure of the binding propensity. This quantity has been listed in Tables 1 and 2, since it is a measure of the thermal affinity of a protein for a denaturing agent, which is independent of concentration. In a relative sense this quantity is quite stable and is not subject to the shifts in order of the other parameters as one goes from one treatment of the data to another. The binding affinity falls in the order lysozyme > ribonuclease > cytochrome c. The ratio of $\Sigma K_j\Delta h_j$ for ribonuclease relative to lysozyme remains almost constant for all five data sets in Tables 1 and 2. Cytochrome c, on the other hand, has uniform ratios relative to the other proteins in both solvents, but the ratio is different in guanidinium chloride. Finally Δh is a measure of a unit binding process. Analyzing the behavior of small molecules via Eq. 9 will be of help in identifying if this process is the binding of a single cosolvent molecule.

Probably the only avenue we have to convert measurements such as those of M&P into molecular information is by way of transfer thermodynamics. This has been attempted with free energies based on the early work and concepts of Kauzmann and Tanford [18,19]. Additivity rules are applied to data for model compounds and the free energy of bringing groups into contact with denaturing solvents is calculated. Though this conception is generally considered valid, the calculations have not yet met with quantitative success [13]. The free energy of unfolding is

usually much smaller than the calculated value. Ahmad and Bigelow [14] have suggested an adjustment of the small molecule data to account for restricted access in large molecules relative to small molecules. It will be difficult to carry this problem further without better data, for example free energy curves for unfolded proteins determined from preferential interaction studies. Most work has been done with the unfolding reaction of proteins where models must be used for the interior of proteins and questions arise on the degree of exposure of the residues to the solvent in the native state. This work has been started by Lapanje, Lilley and their coworkers, but better model systems are available at present for unfolded proteins and it should be possible to carry this work forward.

Enthalpy studies may be more practical than free energy studies. In the paper under study Makhatadze and Privalov have shown that such measurements can be done even in the presence of large heats of dilution, and their methods should be applicable to model compounds as well as proteins. The experiments should be easier and the magnitudes of the enthalpy larger than the free energy. Thermodynamic additivity may also be better for the enthalpy, which is a less subtle property than the free energy. Unfortunately, the data are too few at the present time to provide a proper test.

Finally this study as well as a number of others indicates that urea has a number of advantages over guanidinium chloride as far as thermodynamic interpretation is concerned. The free energy and enthalpy functions of both proteins and model compounds are more linear; the solutions are more ideal; extrapolations to zero concentration are more certain; and least-square analysis of the data is more stable. These remarks apply of course only to studies at moderate temperature where the hydrolysis of urea can be avoided.

8. Glossary

A_k	activity of component k
1, 3	principal solvent, cosolvent
$\{C\}$	activity on molarity scale
d, $d_\circ$	density of solution, principal solvent
Δh	uncorrected binding enthalpy
Δh°	standard state binding enthalpy
Δh_j°	standard state binding enthalpy at site j
Δh_{eff}	effective binding enthalpy
$\overline{H}_2^{\text{ex}}$	excess enthalpy of protein, Q in experimental paper
f_1, f_3, $f_\pm$	activity coefficients on mole fraction scale
k	binding constant on molarity scale
k_j	binding constant on molarity scale at site j
K, K_{eff}	effective binding constant on mole fraction scale
K_j	binding constant on mole fraction scale at site j
$\overline{L}_3$	relative partial molar enthalpy of cosolvent
m_1	number of moles of principal solvent in 1 kg ($1000/M_1$)
M_k	molecular weight of component k
n	effective number of binding sites
W_3	weight fraction of cosolvent
y_1, y_3, $y_\pm$	activity coefficients on molarity scale
γ_1, γ_3, $\gamma_\pm$	activity coefficients on molality scale
θ_j	fractional occupation of site j
μ_3^c, μ_3^m, μ_3^χ	standard state chemical potential on the molar, molal, mole fraction scale
$\{\chi_1\}$, $\{\chi_3\}$	activity of principal solvent, cosolvent on the mole fraction scale

Acknowledgements

This research was supported by NIH Grant GM20159 (JAS) and GM21967 (B. W. Matthews). We also wish to thank C. Schellman for a careful reading of the manuscript.

Appendix A. Empirical formulas for urea and guanidinium chloride

The object of this appendix is to present in one place all the formalism that is required for the thermodynamic analysis of solutes in a mixture of a solvent and cosolvent. It is divided into three sections. The first will be a collection of the standard formulas for evaluating and transforming the thermodynamic variables of solutions. The second will ap-

ply these formulas to urea solutions giving explicit numerical relations for the transformation and interpolation of thermodynamic relations. The third is devoted to guanidinium chloride and will emphasize the additional features required for ionic solutions.

A.1. Important relations of solution thermodynamics

This section will consider only two-component, non-ionic solutions. Formulas with an asterisk will require revision for electrolytes. This will be done in the section on guanidinium chloride.

The formulas compiled below are sufficient for most applications, though a few additional ones are necessary for studies which vary pressure as well as concentration and temperature. These relations are to be found in most standard textbooks of thermodynamics, though few textbooks provide a complete discussion of solutions in the variety of concentrations scales required for modern work. There are excellent background sources, however. The author has found the books of Robinson and Stokes [8] and of Eisenberg [20] to be especially useful for general relationships. Discussion of most of the formulas in the first section are to be found in these sources. The monograph of Harned and Owen [21] has a very detailed treatment of electrolyte solutions and their deviations from the Debye theory. The review of Casassa and Eisenberg [22] is the richest source of relationships but, since the formulation is usually for an arbitrary number of components, a bit of work is required to transform the results for applications to a particular case.

A.1.1. Concentration relationships for two-component solutions

Most of these relations arise directly from the definitions of the various concentration units. Conversion from mole ratios or mass ratios to molarity requires a knowledge of the solution density, d, as a function of concentration. Subscripts 1 and 3 refer to the principal solvent and cosolvent respectively. Higher odd numbers designate solutes of low molecular weight. Even numbers are reserved for macromolecular substances. The symbol m_1 is the molarity of the solvent, which by definition is $1000/M_1$ or 55.51 molal for water. See the glossary for other symbols.

The weight fraction and the molality:

$$W_3 = \frac{M_3 m_3}{1000 + M_3 m_3}; \quad m_3 = \frac{1000 W_3}{M_3(1 - W_3)} \tag{A1}$$

The molality and mole fraction:

$$\chi_3 = \frac{m_3}{m_1 + m_3}; \quad \chi_1 = \frac{m_1}{m_1 + m_3}; \quad m_3 = \frac{\chi_3 m_1}{1 - \chi_3} \tag{A2*}$$

The molarity and molality:

$$C_3 = \frac{m_3 d}{1 + M_3 m_3/1000}; \quad m_3 = \frac{C_3}{d - M_3 C_3/1000} \tag{A3}$$

The molarity and mole fraction:

$$\chi_3 = \frac{C_3}{(1 - M_3/M_1)C_3 + d m_1};$$

$$C_3 = \frac{d m_1 \chi_3}{1 - \chi_3 + M_3/M_1 \chi_3} \tag{A4*}$$

Another concentration scale which is frequently used in solution theory is the volume fraction. This is not discussed in the present work.

A.1.2. Activity and chemical potential relationships

One of the tedious barriers to work on the competition between principal solvent and cosolvent is the need to work on several concentration scales. Experimentalists in general usually work with the molarity scale; the solution thermodynamics is usually reported on the molality scale; and the theory of mixed solvents is best understood on the mole fraction or volume fraction scale. This necessitates a number of conversions during data analysis.

In general the chemical potential of a solute is given by

$$\mu = \mu^\circ + RT \ln(\text{conc.}) + RT \ln(\text{act. coeff.})$$

where (conc.) is the concentration on a convenient scale, (act. coeff.) is the activity coefficient on that scale, and μ° is the standard state chemical potential on that scale, i.e., the chemical potential at conc. = 1, extrapolated from the properties of a dilute ideal solution. For the three cases of interest

$$\begin{aligned} &\text{a} \quad \mu_3 = \mu_3^\chi + RT \ln f_3 \chi_3 \\ &\text{b} \quad \mu_3 = \mu_3^m + RT \ln \gamma_3 m_3 \\ &\text{c} \quad \mu_3 = \mu_3^c + RT \ln y_3 C_3 \end{aligned} \tag{A5*}$$

Note the standard symbols for activity coefficients on the three concentration scales. The symbols μ_3^x, μ_3^m, and μ_3^c are not standard and are an attempt by the author to make clear the nature of the reference state. For example, μ_3^m is the reference state at the standard molality (always equal to unity in this discussion) as extrapolated from the dilute ideal solution. The usual superscript, °, for standard or reference states will be used only for general relations, when the concentration scale is unspecified. Eq. A5 is often written in another way

$$\mu = \mu^\circ + RT \ln(\text{conc.}) + \mu^{ex} \qquad (A6)$$

This defines μ^{ex}, which is called the excess chemical potential or often, loosely, the excess free energy. We clearly have

$$\mu^{ex} = RT \ln(\text{act. coeff.})$$

In molecular interpretations and theoretical discussions, μ^{ex} is often a more useful concept than the activity coefficient itself. μ^{ex} is directly related to the free energy of transfer of a substance from one environment to another and may often be expressed as a sum of different types of energy contributions and entropy contributions. In the works of Scatchard, Casassa, and Eisenberg the quantity μ^{ex}/RT is symbolized by β and is fundamental to the analysis of non-ideality.

It is easy to transform from one of these equations to the other. Take for example going from b to a above. To go from molality to mole fraction, substitute for m_3 in A5b using A2 above, separating the $\ln \chi$ term.

$$\mu_3 = \mu_3^m + RT \ln \chi_3 + RT \ln \frac{\gamma_3 m_1}{1 - \chi_3}$$

Since it is constant, combine the $\ln(m_1)$ term with μ_3^m to obtain

$$\mu_3 = \left[\mu_3^m + RT \ln m_1 \right] + RT \ln \chi_3 + RT \ln \frac{\gamma_3}{1 - \chi_3}$$

To conform with A5a, we must have $\mu_3^x = \mu_3^m + RT \ln m_1$ and $f_3 = \gamma_3 / 1 - \chi_3$. These are the desired relations. The latter can be put in the form

$$f_3 = \gamma_3 (1 + m_3/m_1) \qquad (A7*)$$

which is more useful for converting from molarity to mole fraction units. Many of these relations may be found in the book of Robinson and Stokes. The rule is that after concentration units have been changed in the chemical potential expression, the concentration-dependent factors which have been introduced are part of the activity coefficient expression, and the concentration-independent factors are incorporated in the standard chemical potential. Using this procedure with Eq. A3 the molarity and molarity activity coefficients are related by

$$y_3 = \gamma_3 m_3 d_0 / C_3 = \gamma_3 (1 + m_3 M_3 / 1000)(d_0/d) \qquad (A8)$$

For the standard chemical potentials

$$\begin{aligned} \mu_3^x &= \mu_3^m + RT \ln m_1 \\ \mu_3^c &= \mu_3^m - RT \ln d_0 \end{aligned} \qquad (A9*)$$

These relations permit one to go from published activity data on the molality scale to the molar or mole fractions scale.

The principal solvent is treated differently when it is the only volatile component. Thus experimental papers on solution thermodynamics usually report activity coefficients of solutes but osmotic coefficient, ϕ, of the principal solvent. If one is considering the competition between the principal cosolvent and the cosolvent, it is necessary to put components on the same footing by using mole fraction or volume fraction scales. This is easily done. The practical osmotic coefficient, ϕ, is directly related to the activity on the mole fraction scale:

$$\ln A_1 \equiv \ln f_1 \chi_1 = -m_3 \phi / m_1 \qquad (A10*)$$

so that the chemical potential can be written as

$$\mu_1 = \mu_1^x + RT \ln f_1 \chi_1 \qquad (A11)$$

in parallel with Eq. A5a. μ_1^x is the chemical potential of pure solvent.

A.1.3. Enthalpy and heat capacity relationships

The partial molar enthalpy of a component is given by the relationship

$$\overline{H}_k = \left(\frac{\partial \mu_k / T}{\partial (1/T)} \right)_{P,n_j} = \overline{H}_k^\circ + R \left(\frac{\partial \ln A_k}{\partial (1/T)} \right)_{P,n_j} \qquad (A12*)$$

$k = 1$ or 3 for the special case of a mixture of principal solvent and cosolvent.

For the solute experimentalists usually report the *relative* partial enthalpy, i.e., the partial molar enthalpy relative to the standard state and define a new symbol

$$\bar{L}_3 \equiv \bar{H}_3 - \bar{H}_3^\circ \qquad \text{(A13)}$$

$\bar{L}_3$ may be measured directly from "heats of dilution" experiments. Dividing Eq. A5a or A5b by T and differentiating with respect to $1/T$ gives the relations

$$\bar{L}_3 = R\left(\frac{\partial \ln f_3}{\partial(1/T)}\right)_{P,\chi_3} = R\left(\frac{\partial \ln \gamma_3}{\partial(1/T)}\right)_{P,m_3} \qquad \text{(A14*)}$$

This relation is commonly used to determine the activity coefficient as a function of temperature after it has been measured at one specific temperature. Eq. A12 on the molar scale is more complicated because the molarity changes with temperature as well as the activity coefficient. This is one reason why studies of the precision thermodynamics of solutions usually use the molal scale.

The partial molar entropy is most easily calculated by using the relation $T\bar{S}_k = \bar{H}_k - \mu_k$ and the above equations for $\bar{H}_k$ and μ_k. The partial molar heat capacity is obtained from calorimetric measurements and is obtained by differentiating Eq. A13 with respect to temperature and rearranging the terms

$$Cp_k = Cp_k^\circ + \left(\frac{\partial \bar{L}_k}{\partial T}\right)_{p,n_j} \qquad \text{(A15)}$$

Formulas for other properties such as the partial molar volume and expansibility are easily obtained but will not be used in this analysis. A very useful formula which connects the partial molar properties of one component with those of other components is

$$m_1\left(\frac{\partial \bar{J}_1}{\partial m_3}\right) + m_3\left(\frac{\partial \bar{J}_3}{\partial m_3}\right) = 0 \text{ (constant } T \text{ and } P) \qquad \text{(A16)}$$

where J is any extensive property, V, H, L, Cp, etc. When $\bar{J} = \bar{G} = \mu$, this is the familiar Gibbs–Duhem equation. We shall make use of Eq. A16 to relate the enthalpy of the solvent to that of the solute.

A.2. Urea solutions

The above formalism will now be used to convert the published properties of urea solutions to a form which is useful for the solvent exchange model. In particular we will be interested in activities and activity coefficients on a mole fraction scale. Partial molar enthalpies also enter into the formalism. The raw data are taken from studies of density, activity and enthalpies of solution as a function of concentration. The results will be used to obtain convenient empirical formulas for interpolation and also for the interconversion of units.

Stokes' review [4] presents tables of the activities and osmotic coefficients of urea solutions at several temperatures. Thus the starting point is a table of molalities, molal activity coefficients and osmotic coefficients. Specifically we will be using the 25°C data in Table 10 of Stokes paper. The task is to convert this table to a table which includes activity data for both urea and water on the molar and mole fraction scales as well. All the necessary relations have been given above. The modern spreadsheet is extraordinarily well-suited for this type of calculation. In order to introduce this methodology to those unfamiliar with it, the complete details of the first set of calculations (urea solutions at 25°C) will be presented.

The only missing factor is the density, d, of urea solutions as a function of concentration. This has been determined at 25°C as a function of weight fraction, W_3 (g solute/g solution) by Kawahara and Tanford [23], who have expressed their results with the empirical formula

$$d = d_o\left(1 + d_1 W_3 + d_2 W_3^2\right) = 0.99707\left(1 + 0.2658 W_3 + 0.0330 W_3^2\right) \qquad \text{(A17)}$$

where d_o is the density of pure water at 25°C.

The procedure is now straightforward: W_3 is calculated for each molality using Eq. A1; the density for each entry is calculated from W_3 using Eq. A17; the molarity is calculated using Eq. A3; χ_3 is obtained from Eq. A2 with $m_1 = 55.51$; $\chi_1 = 1 - \chi_3$; Eqs. A7 and A8 provide values for y_3 and f_3 from the tabulated values of γ_3; A_1 (mole fraction scale) is obtained from the osmotic coefficient with Eq. A10 and $f_1 = A_1/\chi_1$; $A_3 = f_3\chi_3$.

A.2.1. Spreadsheet evaluations

This routine set of transformations can be accomplished column by column using a spreadsheet. This is illustrated in Table 3. The first column of Table 3

Table 3
Activities and concentrations of urea solutions

Column	Variable	Equation number	Cell equation
A	m_3	data	
B	ϕ	data	
C	γ_3	data	
D	W_3	A1	$= \mathbf{M}_3^* \underline{A}/(1000 + \mathbf{M}_3^* \underline{A})$
E	d	A17	$= \mathbf{d}_0^* (1 + \mathbf{d}_1^* \underline{D} + \mathbf{d}_2^* \underline{D}^2)$
F	C_3	A3	$= \underline{A}^* \underline{E}/(1 + \mathbf{M}_3^* \underline{A}/1000)$
G	χ_3	A2	$= \underline{A}/(\underline{A} + \mathbf{m}_1)$
H	χ_1	$\chi_1 = 1 - \chi_3$	$= 1 - \underline{G}$
I	A_1	A10	$= \exp(-\underline{A}^* \underline{B}/\mathbf{m}_1)$
J	f_3	A7	$= \underline{C}^* (1 + \underline{A}/\mathbf{m}_1)$
K	f_1	$A_1 = f_1 \chi_1$	$= \underline{I}/\underline{H}$
L	y_3	A8	$= \underline{C}^* (1 + 0.001^* \mathbf{M}_3^* \underline{A})^* \mathbf{d}_0/\underline{E}$
M	$A_3(\chi)$	$A_3 = f_3 \chi_3$	$= \underline{J}^* \underline{G}$
N	$A_3(C)$	$A_3 = y_3 C_3$	$= \underline{L}^* \underline{F}$
O	$A_3(m)$	$A_3 = \gamma_3 m_3$	$= \underline{A}^* \underline{C}$

The entries in the first column of the table are the labels of the columns in the spreadsheet where the variables occur. The second column contains the symbols of the quantities to be stored in the cell of each column. The third column gives references for the source of the data to be stored in each cell. In three cases the data is entered as numbers from the source reference. Symbols like A8, A17 refer to the equations in this appendix which are used to calculate variables in the spreadsheet. In four cases the simple defining formulas are written out. The fourth column contains the cell formula used to calculate the variable. These are direct transcriptions of the equations in column 3. All bold face quantities are to be entered as numerical constants, either by direct entry, or from cells in which they are stored. The underlined symbols are column references found in the first column. Actual cell references have row numbers, A1, B1, etc.

is the label for the column of the spreadsheet in which the variable is stored. We assume that the initial column is the A column. Underlined capital letters indicate the column of the spreadsheet. The second column of Table 3 indicates the variable contained in the spreadsheet column. The third col-

Table 4
Themodynamic functions of aqueous urea solutions at 25°C

A	B	C	D	E	F	G	H	I	J	K	L
m	ϕ (25°C)	γ_3	W_3	d	C	χ_3	χ_1	$A_1(\chi)$	f_3 (25°C)	f_1 (25°C)	y_3
0	1.0000	1.0000	0.0000	0.99707	0.0000	0.00000	0.00000	1.0000	1.0000	1.0000	1.0000
0.5	0.9800	0.9600	0.0292	1.0048	0.4878	0.00893	0.99107	0.9912	0.9686	1.0001	0.9812
1	0.9624	0.9250	0.0567	1.0122	0.9548	0.01770	0.98230	0.9828	0.9417	1.0005	0.9659
1.5	0.9469	0.8950	0.0826	1.0192	1.4025	0.02631	0.97369	0.9747	0.9192	1.0011	0.9544
2	0.9331	0.8670	0.1072	1.0259	1.8317	0.03478	0.96522	0.9669	0.8982	1.0018	0.9439
2.5	0.9208	0.8430	0.1305	1.0322	2.2437	0.04310	0.95690	0.9594	0.8810	1.0026	0.9365
3	0.9096	0.8210	0.1527	1.0383	2.6394	0.05128	0.94872	0.9520	0.8654	1.0035	0.9304
3.5	0.8995	0.8010	0.1737	1.0441	3.0196	0.05931	0.94069	0.9449	0.8515	1.0044	0.9257
4	0.8904	0.7820	0.1937	1.0496	3.3853	0.06722	0.93278	0.9379	0.8384	1.0054	0.9213
5	0.8743	0.7500	0.2309	1.0600	4.0762	0.08263	0.91737	0.9243	0.8176	1.0075	0.9173
6	0.8607	0.7220	0.2649	1.0696	4.7176	0.09755	0.90245	0.9112	0.8000	1.0097	0.9156
7	0.8490	0.6980	0.2960	1.0784	5.3145	0.11199	0.88801	0.8985	0.7860	1.0118	0.9167
8	0.8388	0.6770	0.3245	1.0865	5.8714	0.12597	0.87403	0.8861	0.7746	1.0138	0.9197
9	0.8299	0.6580	0.3509	1.0941	6.3920	0.13952	0.86048	0.8741	0.7647	1.0158	0.9238
10	0.8221	0.6410	0.3752	1.1011	6.8797	0.15265	0.84735	0.8623	0.7565	1.0177	0.9290
11	0.8151	0.6250	0.3978	1.1077	7.3375	0.16539	0.83461	0.8508	0.7489	1.0195	0.9342
12	0.8089	0.6110	0.4188	1.1138	7.7680	0.17776	0.82224	0.8396	0.7431	1.0211	0.9411

umn indicates the formula which is used to calculate the variable. The last column of Table 3 indicates the actual equation in the spreadsheet cell, which is used to make the calculation. In this equation the variables are indicated by their column labels and constants are in boldface. Further details are presented in the legend to Table 3.

The end result is Table 4 consisting of 12 columns containing the concentrations and activities in the various units. Columns M to O were omitted because the full table was too large to print. Once the table is set up, any one of the three concentration units can be used as the independent variable and any one of the activity coefficients or activities as the dependent variable. This permits one to do theoretical work in appropriate units such as the mole fraction, but express the results as a function of molarity or molality.

A.2.2. Empirical formulas

Table 4 only contains data for the concentrations listed in the original table, but it can be used to generate a large number of interpolation formulas for other concentrations by means of power series or other relations. For example, using a standard polynomial fitting program, the following formulas can be generated interrelating C and m:

$$C_3 = (9.98188 \cdot 10^{-1})m_3 + (-4.42699 \cdot 10^{-2})m_3^2 + (1.71434 \cdot 10^{-3})m_3^3 + (-3.86441 \cdot 10^{-5})m_3^4 \quad \text{(A18)}$$

and

$$m_3 = (1.00068)C_3 + (4.74096 \cdot 10^{-2})C_3^2 + (9.04338 \cdot 10^{-4})C_3^3 + (2.58333 \cdot 10^{-4})C_3^4 \quad \text{(A19)}$$

Note that theoretically the limiting slopes of both these equations should be unity. For work at low concentrations one would want to fix these coefficients at unity and accept the small error which develops at high concentrations with the modified formula.

One must be careful to use a reasonable number of terms for such series. Too few will produce errors larger than the experimental error of the data; too many will try to fit the error as well as the functionality and can lead to very poor interpolations and divergent series, i.e., each term in the polynomial is larger than the previous one. The series above were terminated at the quadratic term by using a rough version of the χ^2 test. There are 17 entries in Table 4 and a reasonable value for χ^2 (not reduced) would be about 17. In all the empirical formulas given below parameters are added until χ^2 goes below 17 and the analysis is stopped. All "experiments" were assumed to have an error of about a part in one thousand. Actually for urea we are dealing with the smoothed data of Stokes and the smoothed data represented by the formula of Kawahara and Tanford for the density of urea solutions. As a result the error estimate only reflects the number of significant figures in the smoothed data and χ^2 is somewhat arbitrary and does not have its usual interpretation. On the other hand, this procedure should be good enough to estimate the number of meaningful constants required to fit the data.

A large number of empirical relationships can be developed from Table 4, but we need only a small subset for the applications of this paper. Since the experimental data being treated is in terms of molarity, the only required relationships are f_1, f_3, L_1, L_3, and m_3 or χ_3 as functions of C. We already have $m_3(C_3)$

Table 5
Power series coefficients for urea solutions at 25°C

Coeff.	χ_3	f_3	f_1	L_3 (kJ)	L_1 (kJ)
a0	0.0	1.0	1.0	0.0	0.0
a1	$1.806 \cdot 10^{-2}$	$-6.698 \cdot 10^{-2}$	0.0	$-7.038 \cdot 10^{-1}$	0.0
a2	$4.615 \cdot 10^{-4}$	$7.197 \cdot 10^{-3}$	$6.157 \cdot 10^{-4}$	$3.856 \cdot 10^{-2}$	$6.297 \cdot 10^{-3}$
a3	$2.027 \cdot 10^{-5}$	$-4.977 \cdot 10^{-4}$	$-5.349 \cdot 10^{-5}$	$1.084 \cdot 10^{-3}$	$-2.121 \cdot 10^{-4}$
a4	$1.717 \cdot 10^{-5}$	$4.217 \cdot 10^{-6}$	$-2.167 \cdot 10^{-4}$		$-4.582 \cdot 10^{-5}$
a5			$-2.251 \cdot 10^{-7}$		$5.135 \cdot 10^{-6}$
r.m.s. error (%)	0.03	0.03	0.01	0.04	0.08

Empirical formulas relating one thermodynamic variable to another will be accumulating in this appendix and for the most part require little discussion. They are presented in Table 5. Each item in this table represents a power series of the form

$$y(x) = \sum_{k=0} a_k x^k \qquad \text{(A20)}$$

m_3 and χ_3 can all be represented adequately as polynomials in the molarity and the results have been entered in Table 5. The dependent and independent variables are defined at the head of each column. Values for the r.m.s. percentage error between "experimental" and calculated values are given at the bottom of the table. All the series in Tables 5 and 6 were tested for convergence of the higher terms as discussed above. All coefficients are reported to four significant figures, which is sufficient, though often not necessary, to reproduce the data to the r.m.s. accuracy indicated.

A.2.3. Thermal properties

A very extensive study of the heat of solution, heat capacity and density of aqueous urea solutions has been published by Egan and Luff [5]. Data for $\bar{L}_3$ and $\bar{L}_1$ were given for molalities from 0 to 20 m. Curve fitting for such an extended range of concentrations demands more complicated functions and is usually less accurate than that for smaller ranges. Concentrations above 12 m are rarely used in biochemical studies, and we have cut off the data of Luff and Egan at 12 m, which is in line with the other thermodynamic quantities in the tables. Trial showed that the data for $\bar{L}_3$ were well-represented by a 4-term power series representation The coefficients are given in Table 5.

One could attempt to represent $\bar{L}_1$ by a similar power series, but this would be a mistake. The coefficients for $\bar{L}_1$ and $\bar{L}_3$ are interrelated by Eq. A16 (with $J = L$) with the result that the determination of one set automatically fixes the other. Substituting the power series for $\bar{L}_3$ into Eq. A16 and then integrating yields the relation

$$b_k = -\frac{1}{m_1}\frac{k-1}{k}a_{k-1} \qquad \text{(A21)}$$

where b_k is the coefficient of m_k in the series for $\bar{L}_1$. We see that for $\bar{L}_1$ both b_0 and b_1 are zero so that the series starts with the quadratic term and goes to $k = 5$. Had we attempted to represent $\bar{L}_1$ using the linear term, we would have obtained solutions lacking thermodynamic consistency. The $\bar{L}_1$ and $\bar{L}_3$ coefficients in Table 5 were obtained from the reported data and do not exactly satisfy Eq. A21. The reason for this is not clear since Eq. A21 was presumably used to obtain $\bar{L}_1$ in the original publication.

The logs of activities are also related by Eq. A16. We, however, will be dealing with the activity coefficients themselves and no such simple relationship as Eq. A21 applies. It can be shown, however, that if f_3 is represented by a power series starting with a linear term in C, then f_1 will be represented by a power series beginning with a quadratic term.

A.3. Guanidinium chloride

For ionic substances a part of the preceding formulation must be altered. Those formulas which deal with the neutral substance as the thermodynamic component remain unchanged, but others which describe individual ions, such as activities and activity

Table 6
Power series coefficients for guanidinium chloride at 25°C

Coeff.	χ_3	f_3	f_1	$\bar{L}_3$ (kJ)
a0	0.0	1.0	1.0	0.0
a1/2	0.0	$-6.739 \cdot 10^{-1}$	0.0	1.665
a1	$1.809 \cdot 10^{-2}$	$2.205 \cdot 10^{-1}$	0.0	−3.494
a3/2	0.0	0.0	$8.416 \cdot 10^{-3}$	0.0
a2	$5.706 \cdot 10^{-4}$	$-1.501 \cdot 10^{-2}$	$-1.498 \cdot 10^{-3}$	$3.933 \cdot 10^{-1}$
a3	$4.014 \cdot 10^{-5}$	$8.663 \cdot 10^{-4}$	$2.377 \cdot 10^{-4}$	$-2.145 \cdot 10^{-2}$
a4	0.0	0.0	$-3.56 \cdot 10^{-5}$	0.0
r.m.s. error (%)	0.03	0.20	0.01	0.20

coefficients, take a different form. Because our interest is in guanidinium chloride, we need only consider the case of univalent ions, which simplifies the discussion considerably. More general formulas are given in the book by Robinson and Stokes.

In general the subscript 3 will be used to designate a neutral substance, while ± will be used as a subscript for ionic properties. Because the free energy of a substance is regarded as the sum of the free energies of its component ions the activities, A_3, on the molal, molar and mole fraction scale are given respectively by

$$A_3(m) = A_\pm^2 = \gamma_\pm^2 m_3^2 \quad (\text{Column } \underline{O})$$
$$A_3(C) = A_\pm^2 = y_\pm^2 C_3^2 \quad (\text{Column } \underline{N})$$
$$A_3(\chi) = A_\pm^2 = f_\pm^2 \chi_\pm^2 \quad (\text{Column } \underline{M})$$

For univalent salts the molarity and molality of the ions is the same as the concentration of the substance. For the mole fraction it is customary to treat the individual ions as components

$$\chi_+ = \chi_- = \chi_\pm = \frac{m_3}{2m_3 + m_1} \quad (\text{Column } \underline{G}) \qquad \text{(A2a)}$$

so that $\chi_1 + \chi_+ + \chi_- = 1$. This convention is discussed in the book of Robinson and Stokes. Equation numbers in this section match the equivalent numbers in Section A.1. For example, Eq. A2a replaces A2 and column G in Table 3.Other replacements are as follows:

$$\chi_1 = 1 - 2\chi_\pm \quad (\text{Column } \underline{H})$$
$$A_1 = \exp[-2m_3\phi/m_1] \quad (\text{Column } \underline{I}) \qquad \text{(A10a)}$$
$$f_3 = \gamma_3(1 + 2m_3/m_1) = \gamma_3/\chi_1 \quad (\text{Column } \underline{J}) \qquad \text{(A7a)}$$

In dealing with the thermal properties of electrolyte solutions, it is customary to use molar rather than ionic quantities, so that formulas like A13 and A15 remain intact. On the other hand, Eq. A14, which relates molar enthalpies to ionic activity coefficients, becomes

$$\bar{L}_3 = 2R\left(\frac{\partial \ln f_\pm}{\partial(1/T)}\right)_{P,\chi_3} = 2R\left(\frac{\partial \ln \gamma_\pm}{\partial(1/T)}\right)_{P,m_3} \qquad \text{(A14a)}$$

These are all the relations which we will need to construct the equivalents of Tables 3 and 4 for univalent salts, but the reader is warned that all the relations in Section A.1 which are marked by an asterisk require alteration if they are to be used for solutions of ionic substances. The other entries in Table 3 remain unchanged.

The recent paper of Makhatadze et al. [6] on the density, heat capacity, and heat of dilution together with the earlier study of Schrier and Schrier [7] on the activity and osmotic coefficient provide us with an unusually complete description of the thermodynamics of guanidinium chloride solutions. The density studies of Makhatadze et al., which cover both temperature and concentration variation, are in complete agreement with the earlier study of Kawahara and Tanford at 25°C [23]. The activity results of Schrier and Schrier differ slightly from the earlier work by Tanford's group [24,25]. We have accepted the later and more complete study. The results of Schrier and Schrier are incorporated in the paper of Makhatadze and Privalov with the result that all the information required to construct a table similar to Table 3 may be found there.

The ionic nature of the guanidinium chloride causes changes in the form of the empirical equations used to represent the data. In dilute solution both the log of the activity coefficient and $\bar{L}_3$ depend on the square root of the concentration. This is also true of the activity coefficients themselves, if they are expressed as power series. Consequently, quantities like $f_\pm$ or $\bar{L}_3$ will be expressed by series of the form

$$y = a_{1/2}x^{1/2} + \sum_{k=1} a_k x^k$$

and by arguments similar to those used for urea solutions, this means that quantities like f_1 and $\bar{L}_1$ will be represented by series beginning with the 3/2 power of the concentration,

$$y = a_{3/2}x^{3/2} + \sum_{k=2} a_k x^k$$

It was found that using the Debye–Hückel values for the leading terms in $f_\pm$ and $\bar{L}_3$ did not give successful representations. The Debye–Hückel terms become much too large at high concentrations so that the main function of the rest of the terms turns out to be to negate their contributions. On the other hand, it is not possible to get a good representation without

the 1/2-power or 3/2-power dependency. $a_{1/2}$ and $a_{3/2}$ were treated as parameters in the least square fitting and turned out to have values about 50% of the Debye–Hückel value. The resulting empirical formulas cannot be used for concentrations below the lowest experimental concentration. The results are summarized in Table 6.

References

[1] G.I. Makhatadze and P.L. Privalov, J. Mol. Biol., 226 (1992) 491–505.
[2] J.A. Schellman, Biopolymers, 34 (1994) 1015–1026.
[3] J.A. Schellman, Biophys. Chem., 37 (1990) 121–140.
[4] R.H. Stokes, Aust. J. Chem., 20 (1967) 2087–2100.
[5] E.P. Egan and B.B. Luff, J. Chem. Eng. Data, 11 (1966) 192–194.
[6] G.I. Makhatadze, J. Fernandez, E. Freire, T.H. Lilley and P.L. Privalov, J. Chem. Eng. Data, 38 (1993) 83–87.
[7] M.Y. Schrier and E.E. Schrier, J. Chem. Eng. Data, 22 (1977) 73–74.
[8] R.A. Robinson and R.H. Stokes, Electrolyte Solutions, Butterworth, London, 1955.
[9] Y. Nozaki and C. Tanford, J. Biol. Chem., 238 (1963) 4074–4081.
[10] Y. Nozaki and C. Tanford, J. Biol. Chem., 245 (1970) 1648–1652.
[11] D.R. Robinson and W.P. Jencks, J. Biol. Chem., 238 (1963) PC1558–1560.
[12] H.B. Bull, K. Breese and C.A. Swenson, Biopolymers, 17 (1978) 1091–1100.
[13] S. Lapanje, Biopolymers, 23 (1984) 1943–1950.
[14] F. Ahmad and C.C. Bigelow, Biopolymers, 29 (1990) 1593–1598.
[15] J.A. Schellman, Biophys. Chem., 45 (1993) 273–279.
[16] N. Poklar and S. Lapanje, Biophys. Chem., 42 (1992) 283–290.
[17] T. Arakawa and S.N. Timasheff, Biochemistry, 21 (1982) 6536–6544.
[18] W. Kauzmann, Adv. Protein Chem., 14 (1959) 1–63.
[19] C. Tanford, Adv. Protein Chem., 24 (1970) 1–95.
[20] H. Eisenberg, J. Chem. Phys., 36 (1961) 1837–1843.
[21] H.S. Harned and B.B. Owen, The Physical Chemistry of Electrolyte Solutions, Reinhold, New York, 1950.
[22] E.F. Casassa and H. Eisenberg, Adv. Protein Chem., 19 (1964) 287–395.
[23] K. Kawahara and C. Tanford, J. Biol. Chem., 241 (1966) 3228–3232.
[24] E.P.K. Hade and C. Tanford, J. Am. Chem. Soc., 89 (1967) 5034–5040.
[25] K. Aune and C. Tanford, Biochemistry, 8 (1969) 4586–4590.