Thermochimica Acta, 96 (1985) 333–348
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

CALORIMETRY OF MODEL BIOMOLECULES IN AQUEOUS SOLUTION *

ATTILIO CESÀRO

Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, Università di Trieste, Piazzale Europa 1, I-34127 Trieste (Italy)

ABSTRACT

A survey of the thermodynamic properties of aqueous solutions of model compounds is presented. The review is focused on the analysis of the thermal data obtained from dilution and mixing experiments monitoring weak interactions occurring between solvated solute species. The treatment of experimental results in the form of a virial-like expansion and the factorization analysis known as "group-contribution approach" is discussed in order to gain insight into the phenomenological process at molecular level.

Examples are chosen in three dominating classes of biomolecular interactions: purine-like, peptide and carbohydrate molecules.

INTRODUCTION

This paper surveys the solution properties (and solvation effects) of model molecules in aqueous solution. The discussion will be mainly limited to approaches providing information on the thermodynamics of weak, non-covalent interactions (i.e., equilibrium properties) through the use of isothermal microcalorimetry to obtain enthalpy changes for dilution and mixing processes.

All theoretical and experimental approaches to the thermodynamics of solutions must recognize the fundamental problem of the concentration dependence of the properties under examination. Numerical data from a given experiment, indeed, are devoid of interest unless the phenomenological properties are: (i) converted into a proper standard framework, and (ii) related to a formal theory using molecular parameters for the understanding of the macroscopic properties.

As regards the first point, the major source of confusion in developing a formal framework for excess properties of solutions lies in the choice of the standard state, which, though arbitrary, must be clearly specified [1]. In principle, the ideal gas phase is the most sensible convention for the solute reference state, but for our purposes the infinitely dilute solution state, largely adopted for measuring concentration dependence, is the more con-

* Presented at the International Summer School of Calorimetry and Thermal Analysis, 1–5 October 1984, Belgirate, Italy.

0040-6031/85/$03.30

venient reference. This choice, indeed, is almost imperative in view of the paucity of data for the transfer process of solute from the isolated molecule in vacuo to that in solution. The other reference state popular in thermochemistry, i.e., that of pure compounds, is not practical because of the scarcity of heat of combustion data, and because the errors in $\Delta_f H^0$ derived from combustion data are often of the same order of magnitude as the heat effects in solution.

As regards the second point, we may adopt the treatment of solution properties originally presented by McMillan and Mayer [2], and specifically applied to aqueous solutions of electrolytes by Mayer [3] and Friedman [4] and to those of non-electrolytes by Kauzmann and co-workers [5]. In the last ten years, this virial formalism has been extended to many non-electrolyte aqueous systems and has unified the way of presenting numerical parameters.

From the theoretical point of view, rigorous description of these systems is particularly difficult, because of the absence of a long-range structural order as well as a complete statistical disorder. It is shown by X-ray and neutron scattering experiments that short-range correlations do occur. The non-random distribution of molecules may be conveniently described by the radial distribution function $g(r)$, which usually shows damped oscillations and decays more or less rapidly. Knowledge of $g(r)$ for all pairs of molecules, i.e., solvent–solvent, solvent–solute, solute–solute, would allow complete characterization of the solution. The expressions "solvent-structuring" and "solvent-destructuring" have often been used in qualitative description of the macroscopic changes observed in the thermodynamic and non-thermodynamic properties and more or less appropriately attributed to unknown changes in correlation functions.

THERMODYNAMICS

Central to the study of thermal properties of solution with two components is the non-ideality of the solution. Therefore, an activity coefficient, γ_i, for each component ($i = 1, 2$), or alternatively, an osmotic coefficient of the solution, $\phi = -(x_1/x_2) \ln a_1$ can be measured. The general properties of the Gibbs energy functions imply that the expression for the temperature dependence of the activity (or osmotic coefficient) involves an enthalpy function. Reference may be made to general chemical thermodynamics textbooks [6,7] for correct derivation of the following equations. Without forgetting the warnings mentioned earlier, we may simply consider that heat of dilution data therefore provide information about the temperature dependence of the activity and activity coefficient. For the solute species

$$\left(\frac{\partial \ln \gamma_2}{\partial T}\right)_{P,m} = \left(\frac{\partial \ln a_2}{\partial T}\right)_{P,m} = \frac{L_2}{RT^2} \tag{1}$$

The relative partial molar enthalpy, L_2, and the corresponding relative apparent molar enthalpy, L_ϕ, are related to each other by

$$L_2 = \left(\frac{\partial L}{\partial m}\right)_{n_1} = L_\phi + m\left(\frac{\partial L_\phi}{\partial m}\right)_{n_1} \tag{2}$$

and are obtained from the experimentally measured heat effect in a mixing experiment of a solution at initial molality m_i with pure solvent to give a solution at final molality m

$$\Delta H^{\text{dil}}_{(m \leftarrow m_i)} = L_\phi(m) - L_\phi(m_i) \tag{3}$$

where $\Delta H^{\text{dil}}_{(m \leftarrow m_i)}$ is the heat of dilution for a quantity of solution containing one mole of solute.

In the simple case of one solute only, L_ϕ can be expressed in a virial-like form

$$H^{\text{E}} = L_\phi m = h_{xx}m^2 + h_{xxx}m^3 + \ldots \tag{4}$$

where the coefficients h_{xx}, are related to the interactions between pairs, triplets, etc.

In a more general case, when two (or more solutes) are involved, the excess enthalpy, H^{E}, of a solution containing solutes x and y can be expressed on the Lewis–Randall molality scale

$$\begin{aligned} H^{\text{E}}_{(m_x,m_y)} &= H_{(m_x,m_y)} - H^0_{\text{w}} - m_x H^0_x - m_y H^0_y \\ &= h_{xx}m_x^2 + 2h_{xy}m_x m_y + h_{yy}m_y^2 + h_{xxx}m_x^3 + 3h_{xxy}m_x^2 m_y \\ &\quad + 3h_{xyy}m_x m_y^2 + h_{yyy}m_y^3 + \ldots \end{aligned} \tag{5}$$

where $H^{\text{E}}_{(m_x,m_y)}$ and $H_{(m_x,m_y)}$ are the excess enthalpy and the enthalpy, respectively, of a solution containing 1 kg of water and m moles of each solute; H^0_{w} is the enthalpy per kg of pure water and H^0_x and H^0_y are the limiting partial molar enthalpies of the solutes; h_{xx}, h_{xxy}, etc., are the enthalpic contributions characterizing pair, triplet, etc., interactions.

To facilitate evaluation of the cross-interaction coefficient from the experimental data, one introduces the function $\Delta H''$ defined as

$$\Delta H'' = \Delta H^{\text{mix}}_{x,y} - \Delta H^{\text{dil}}_x - \Delta H^{\text{dil}}_y \tag{6}$$

where the heat of mixing $\Delta H^{\text{mix}}_{x,y}$ per kg of water of the final solution is given by

$$\Delta H^{\text{mix}}_{(m_x,m_y \leftarrow m_{x_i},m_{y_i})} = H^{\text{E}}_{(m_x,m_y)} - \frac{m_x}{m_{x_i}} H^{\text{E}}_{(m_{x_i})} - \frac{m_y}{m_{y_i}} H^{\text{E}}_{(m_{y_i})} \tag{7}$$

and where

$$\Delta H^{\text{dil}}_{(m_x \leftarrow m_{x_i})} = H^{\text{E}}_{(m_x)} - \frac{m_x}{m_{x_i}} H^{\text{E}}_{(m_{x_i})} \tag{8}$$

is, for instance, the heat of dilution of solute x from the initial molality m_{x_i} to the final molality m_x. Then, from eqns. (4)–(8) it follows that

$$\Delta H'' = H^{E}_{(m_x,m_y)} - H^{E}_{(m_x)} - H^{E}_{(m_y)}$$

$$= 2h_{xy}m_x m_y + 3h_{xxy}m_x^2 m_y + 3h_{xyy}m_x m_y^2 + \ldots \tag{9}$$

The values of the coefficients can be obtained by interpolating $\Delta H''/m_x m_y$ with a least-squares method.

For each of the h coefficients, relationships with the excess free energy and entropy coefficients can be written as for the overall properties

$$h_{xy} = g_{xy} + Ts_{xy} \tag{10a}$$

$$h_{xy} = \left[\partial(g_{xy}/T)/\partial(1/T)\right]_P \tag{10b}$$

and so on.

Before the experimental results are discussed, mention must be made of the virial-type expansion of solution properties [4,5]. The essential result of this theory is a rigorous one-to-one correspondence between the equations of imperfect gas theory and dilute solutions of non-electrolytes: the pressure behavior of the gas corresponds to the osmotic pressure, π, of the solutions which can be fitted by a virial expansion of the form

$$\frac{\pi}{kT} = \rho + B^*\rho^2 + C^*\rho^3 + \ldots \tag{11}$$

where ρ is the number of solute particles per unit volume.

The virial coefficients are formally identical to those of the imperfect gas theory, but instead of the potential U_N, one must use the average potential of mean force of N solute molecules in the pure solvent. The theory can also be extended to the development of a distribution function since, although the solutions are represented by an idealized model, one should keep in mind that its ultimate statistical–mechanical basis lies in the McMillan–Mayer theory. Virial coefficients are therefore described in terms of integrals of the radial distribution function, which is related in turn to the exponential term of the molecular pair interaction potential, $W(r)$

$$B^* = -1/2\int_0^\infty [\langle g(r)\rangle - 1]4\pi r^2 dr \tag{12}$$

$$\langle g(r)\rangle = \exp -\left[\frac{\langle w(r)\rangle}{kT}\right] \tag{13}$$

where the angle brackets ($\langle\rangle$) denote an average overall mutual orientation of the molecules (both solute and solvent). Although possible in principle, evaluation of $\langle g(r)\rangle$ from known molecular parameters is, in practice, a formidable task for non-spherical hydrated solutes. Equation (12) also shows that the orientational averaging process can smooth out any discernible character of $\langle g(r)\rangle$ in experimentally determined B^* values. Any excess

thermodynamic properties, J^E, of a solution can be expressed as a virial expansion of molality m, in the Lewis–Randall molality scale *

$$J^E = J - J_s^0 - m J_i^0 - J_{ideal} = j_{ii} m^2 + j_{iii} m^3 + \ldots \quad (14)$$

where J and J^E are the thermodynamic property and the corresponding excess quantity, respectively, referred to an amount of solution containing m moles of solute i and 1 kg of solvent; J_s^0 and J_i^0 are the standard property for 1 kg of solvent and that of solute in the infinite dilution state, respectively. By considering proper units (when using the molality scale) an equation for excess free energy can be written in a form similar to that of eqn. (1) and experimentally related to the osmotic coefficient, ϕ, of the solution

$$G^E = \frac{\pi}{m} - RT = RT(\phi - 1) = g_{ii} m + g_{iii} m^2 + \ldots \quad (15)$$

where it should be recalled that the g terms assume non-zero values for an ideal solution due to the very definition of the practical osmotic coefficient in the molality scale (for instance, $g_{ii} = -1/2(M_s RT/1000)$, where M_s is the molar mass of solvent in kg mol^{-1}). Similarly, L_ϕ is by definition the "excess" enthalpy of the solution and expressed in virial form (see eqn. 4)

$$H^E = L_\phi m = h_{ii} m^2 + h_{iii} m^3 + \ldots \quad (16)$$

It is therefore apparent that the experimentally determined coefficents g_{ii}, $g_{iii}, \ldots,$ h_{ii}, h_{iii} are related to the interaction coefficients of the McMillan–Mayer theory. Accurate measurements of osmotic coefficients and dilution enthalpies are the only requirement for evaluating the parameters of eqns. (15) and (16). Most of the recent data have now been analyzed in terms of pairwise interactions, and polynomial coefficients for many solutes have been compared.

GROUP-CONTRIBUTION APPROACH

As to the group-contribution approach [11–20] suffice it to say that the additivity principle is generally applicable in chemical thermodynamics to determine the contribution of groups of atoms to molecular properties, and has also been used recently for excess solution properties. The contributions of each functional group to the pair, triplet, etc., interaction coefficients may be estimated from the osmotic coefficients and heats of dilution and used to predict the polynomial coefficients of other solutes. The agreement between

* Even if the McMillan–Mayer theory is constructed at constant volume and therefore the concentration scale to be used is that of molarities, it is most customary to express the experimental properties in terms of molalities. The correlation between the two different scales has been discussed in the literature [8–10].

predicted and experimental data can be generally considered as fairly good, but it emphasizes that the true cause of the differences in behavior of solutes in water is either of a stereochemical (as in the case of carbohydrates) or of a concerted nature (as for diketopiperazines). It must be noted, in fact, that group contributions do not take into account the stereochemical contributions of the individual components. Accurate determinations of these coefficients will therefore allow construction of a reference framework with which the peculiar interactions of any given system may be compared. The fact that in polynomial coefficients all pairwise interactions in the solution are averaged out, including those between water–water and solute–water, implies that mutually compensatory effects may be occurring, especially if only one thermodynamic property is considered.

ILLUSTRATIVE CASES

Three examples of "weak" interactions in aqueous solution will now be examined. It must be remembered that significant assessment of the interactions between molecules requires not only reliable experimental data, but also that thermodynamic quantities extrapolated to infinite dilution are available, especially in the case of non-electrolyte solutes. The situation is much simpler with solutions of electrolytes, where limiting laws can predict the course of any thermodynamic excess function as $m \rightarrow 0$. Polar non-electrolyte solutes must be regarded with special care since, in the past, they were too often schematically classified as either hydrophobic or solvated-hydrophilic species. An analysis of the more recent data of aqueous solutions suggests an incompatibility with these classical models, in particular for the hydrophobic interaction.

PURINE STACKING

Base-pair interactions take place in nucleic acids by hydrogen bonds between complementary moieties. This picture is more geometrical than energetical, since stabilizing factors predominantly arise from the stacked aggregates of base pairs. These interactions are significant in aqueous solution and have been thought to arise from hydrophobic effects (e.g., water clustering around aggregates of solute molecules). The thermodynamics of base and nucleoside association has provided a major clue to understanding the nature of stacking processes. Caffeine is a good model for this kind of study, since it presents no protolytic products or electrostatic interactions [21–24].

The basic idea for the analysis of solution properties of caffeine (as well as other purine-like molecules) is that all deviations from ideality must arise

from aggregation. In this quasi-chemical approach [25], the total molality, m, is related to the molality of the unassociated monomer, m_1, through the association constant, K, this being assumed to be the same for each step

$$m = m_1 + 2m_2 + 3m_3 + \ldots$$
$$= m_1(1 + 2Km_1 + 3K^2m_1^2 + \ldots)$$
$$= m_1/(1 - Km_1)^2 \tag{17}$$

The practical osmotic coefficient, ϕ, for such a mixture is then given by

$$\phi = \tilde{m}/m \tag{18}$$

where $\tilde{m}$ is the colligative molality (sum over species molality). Assuming that all aggregation steps have the same K value it follows that

$$K = (1 - \phi)/m\phi^2 \tag{19}$$

The activity coefficient of the solute is given by

$$\gamma = m_1/m = \left\{2/\left[1 + (4mK + 1)^{1/2}\right]\right\}^2 \tag{20}$$

A serial expansion of this relation gives

$$\ln \gamma = -2Km + 3K^2m^2 - \tfrac{20}{3}K^3m^3 + \ldots \tag{21}$$

Using the same multiple association hypothesis, analysis of the enthalpy of dilution data (any other thermodynamic property will be similarly defined) gives the following equation

$$L_\phi = K\Delta H^0 m - 2K^2\Delta H^0 m^2 + 5K^3\Delta H^0 m^3 + \ldots \tag{22}$$

where ΔH^0 is the enthalpy of association.

The data of ΔG^0 and ΔH^0 for caffeine in water and in aqueous 1 m solution of urea, guanidinium chloride and potassium chloride, and for other nucleotide systems, are reported in Table 1.

The results provide the following conclusions (at 298.15 K). Stacking is strongly exothermic, $\Delta S^0 < 0$ and $|\Delta H^0| > |T\Delta S^0|$. N-Methyl substitution

TABLE 1

Thermodynamic functions of some purine bases self-association at 298.15 K (kJ mol^{-1})

Compound	ΔG^0	ΔH^0	$T\Delta S^0$
Purine	−1.8	−17.6	−15.8
N-6-Me-purine	−4.6	−25.1	−20.5
Deoxyadenosine	−6.3	−15.5	−9.2
Caffeine (water)	−5.6	−13.4	−7.8
Caffeine (1 *m* urea)	−5.1	−9.6	−4.5
Caffeine (1 *m* Gu·HCl)	−4.6	−8.4	−3.8
Caffeine (1 *m* KCl)	−6.6	−10.9	−4.3

favors aggregation (ΔS^0 less negative). For caffeine, at all events, the entropy of dimerization seems to result from the cratic contribution upon association (in water ca. 34 J K^{-1} mol^{-1}). These statements show that base stacking does not conform to rules characterizing the hydrophobic interactions, though the presence of hidden effects cannot be ruled out.

PEPTIDE–PEPTIDE AND PEPTIDE–UREA INTERACTION

Peptide–peptide and urea–peptide interactions are of great interest in biophysical chemistry because of their role in the conformational stability and transitions of proteins and polypeptides. Studies with model compounds were designed to estimate these energetic contributions and in some cases, but not many, reliable data on the thermodynamic functions characterizing the solute transfer from vapor phase to infinitely dilute aqueous solutions are also available. When series of compounds are studied, there is the attractive possibility of estimating additive contributions to the thermodynamic quantities by certain functional groups [13,15,16,18,20,26–28]. Within the framework of the McMillan–Mayer theory, Wood has obtained an expression for the second coefficient of the excess enthalpy

$$h_{xy} = \sum_{ij} n_i^* n_j^y \{ H_{ij} \} \tag{23}$$

The enthalpic coefficient is therefore given by the sum of all the contributions $\{ H_{ij} \}$ obtained by coupling each group j of the solute molecule y, and i of molecule x (of the same or different species). Although of practical interest, eqn. (23) contains some approximations, whose origins are in part implicit in the cluster-expansion treatment. For example, the additivity of group interactions is invalidated by any cooperative or concerted effect of several groups leading to the specific mutual recognition of two molecules. Equation (23) seems to work well when the coefficients are averaged over a large number of structurally similar solutes, characterized by weak and scarcely specific interactions. Therefore, the approach can provide a useful basis to ascertain whether unexpected effects are present in a given system.

Wood [14] has made some assumptions to reduce to a minimum the number of group interactions to be considered (this number increases according to the function $n(n+1)/2$, n being the number of groups chosen). One can select –CONH–, Urea (U) and $-CH_2$ groups, assuming $CH = 0.5CH_2$ and $CH_3 = 1.5CH_2$ and that all the contributions involving CONH and U have the same value independent of their degree of N-substitution. It has also been found that the $\{ H_{u,CONH} \}$ and $\{ H_{CONH,CONH} \}$ contributions for aqueous solutions of diketopiperazines differ from those obtained for amide solutions. This has been ascribed to the different stereochemistry for the –CONH– group. At least in the solid state, stereo-

TABLE 2

Values of the coefficients of the excess enthalpies used to derive $\{H_{ij}\}$ for amides, peptides and ureas in water at 298.15 K

Solutes		n^x_{CONH} [a]	n^y_{CONH} [a]	$n^x_{CH_2}$ [a]	$n^y_{CH_2}$ [a]	h_{xy} [b] (exp)	h_{xy} [b] (calc)
x	*y*						
NMF	NMF	1	1	2	2	272	126
NMA	NMA	1	1	3	3	236	393
NMP	NMP	1	1	4	4	636	693
NBA	NBA	1	1	6	6	1477	1394
NMF	NMA	1	1	2	3	368	251
NMF	NMP	1	1	2	4	540	377
NMF	NBA	1	1	2	6	883	627
NMA	NMP	1	1	3	4	393	535
NMA	NBA	1	1	3	6	628	819
NMP	NBA	1	1	4	6	1079	1010
FA	FA	1	1	0.5	0.5	−115	−211
AA	AA	1	1	1.5	1.5	12	5
PA	PA	1	1	2.5	2.5	249	256
DMF	DMF	1	1	3.5	3.5	737	539
DMA	DMA	1	1	4.5	4.5	962	856
FA	DMF	1	1	0.5	3.5	155	89
NAGA	NAGA	2	2	2.5	2.5	−220	−208
NAAA	NAAA	2	2	3.5	3.5	268	259
NAVA	NAVA	2	2	5.5	5.5	1259	1294
NALA	NALA	2	2	6.5	6.5	1714	1861
NAGA	NAAA	2	2	2.5	3.5	68	17
NAGA	NAVA	2	2	2.5	5.5	385	468
NAGA	NALA	2	2	2.5	6.5	547	693
NAAA	NAVA	2	2	3.5	5.5	591	743
NAAA	NALA	2	2	3.5	6.5	899	985
NAVA	NALA	2	2	5.5	6.5	1486	1569
NAG_2A	NAG_2A	3	3	3.5	3.5	−646	−635
NAG_3A	NAG_3A	4	4	4.5	4.5	−1499	−1277
NAGA	NAG_2A	2	3	2.5	3.5	−211	−368
NAGA	NAG_3A	2	4	2.5	4.5	−544	−528
NAA_2A	NAA_2A	3	3	5.5	5.5	939	766
NAAA	NAA_2A	2	3	3.5	5.5	641	449
NAAGA	NAAGA	3	3	4.5	4.5	284	49
NAA_3A	NAA_3A	4	4	7.5	7.5	4880	1525
U	U	0	0	0	0	−350	—
U	NMF	0	1	0	2	−109	−47
U	NMA	0	1	0	3	0	28
U	NMP	0	1	0	4	180	102
U	MBA	0	1	0	6	264	251
MMU	MMU	0	0	1.5	1.5	−85	−89
MEU	MEU	0	0	2.5	2.5	160	127
MPU	MPU	0	0	3.5	3.5	292	376
1,3DMU	1,3DMU	0	0	3	3	35	247
1,1DMU	1,1DMU	0	0	3	3	38	247

TABLE 2 (continued)

Solutes		n^x_{CONH} [a]	n^y_{CONH} [a]	$n^x_{CH_2}$ [a]	$n^y_{CH_2}$ [a]	h_{xy} [b] (exp)	h_{xy} [b] (calc)
x	*y*						
1,3DEU	1,3DEU	0	0	5	5	1011	813
1,1DEU	1,1DEU	0	0	5	5	791	813
U	MMU	0	0	0	1.5	−151	−238
MMU	1,3DMU	0	0	1.5	3	87	60
MMU	MEU	0	0	1.5	2.5	76	11
MMU	1,3DEU	0	0	1.5	5	410	260
1,3DMU	MEU	0	0	3	2.5	131	185
1,3DMU	1,3DEU	0	0	3	5	388	497
MEU	1,3DEU	0	0	2.5	5	508	418
MBU	MBU	0	0	4.5	4.5	1039	659

Abbreviations used in the table: NMF, *N*-methylformamide; NMA, *N*-methylacetamide; NMP, *N*-methylpropionamide; NBA, *N*-butylacetamide; FA, formamide; AA, acetamide; PA, propionamide; DMF, *N,N'*-dimethylformamide; DMA, *N,N'*-dimethylacetamide; NAGA, *N*-acetylglycinamide; NAAA, *N*-acetyl-L-alaninamide; NAVA, *N*-acetyl-L-valinamide; NALA, *N*-acetyl-L-leucinamide; NAG_2A, *N*-acetyl-glycyl-glycinamide; NAG_3A, *N*-acetyl-glycyl-glycyl-glycinamide; NAA_2A, *N*-acetyl-L-alanyl-L-alaninamide; NAAGA, *N*-acetyl-L-alanyl-glycinamide; NAA_3A, *N*-acetyl-L-alanyl-L-alanyl-L-alaninamide; U, urea; MMU, monomethylurea; MPU, mono-*n*-propyl-urea; MBU: mono-*n*-butylurea; 1,3DMU, 1,3(*N,N'*)-sym-dimethylurea; 1,1DMU, 1,1(*N,N*)-asym-dimethylurea; 1,3DEU, 1,3(*N,N'*)-sym-diethylurea; 1,1DEU, 1,1(*N,N*)-asym-diethylurea.

[a] Number of $-CONH-_{trans}$, or $-CON-_{trans}$, or $-CONH_2$ groups and number of $-CH_2-$ groups, or fractions, respectively, on the solute molecule *x* or *y*. Note that the number of urea residues is $n'_U = 1$, for urea and urea derivatives, and $n'_U = 0$, for all other solutes considered.

[b] Units: J mol^{-1} (mol kg^{-1})$^{-1}$.

chemical studies on the –CONH– group show that stable structures are to be expected when the angle of internal rotation about the C(O)–N bond is close to that of the *trans*-conformation. In the cyclic dipeptides, on the other hand, the two –CONH– groups are forced into the domain of the *cis*-conformations to minimize the intramolecular constraints. The most stable angle of internal rotation about the C(O)–N bond depends on the substituents determining the overall conformation of each diketopiperazine ring (quasi-planar, chair or boat). The set of six $\{H_{ij}\}$ contributions arising from the interactions of three selected groups (CONH, CH and U) has been obtained by averaging the available experimental data on the excess enthalpies of aqueous solutions of linear peptides, amides and ureas. The experimental data are listed in Table 2, together with the number of groups present on each of the solute molecules and the values of h_{xy} calculated from eqn. (23) using the mean $\{H_{ij}\}$ contributions given in Table 3 (column 3). The other columns report values obtained over a smaller number of experimental data. The improvement in the confidence limits reported in column 3 is a clear indication that the series of data in Table 2 is homogeneous. Attempts

TABLE 3

Group contributions $\{H_{i,j}\}$ to the second virial coefficients of the excess enthalpies for amides, ureas, linear peptides and diketopiperazines in water at 298.15 K

Functional groups		Amides and linear peptides				Cyclic dipeptides, Ref. 20
i	*j*	Ref. 20	Ref. 27	Ref. 15	Refs. 14,26,28	
CONH	CONH	−307(22) [a]	−311(57)	−252(105)	−251(103)	−470(38) [c]
CONH	U	−196(75) [a]	–	–	–	−525(27) [c]
U	U	−350 [b]	–	–	−280	−350 [b]
CONH	CH_2	92(11) [a]	95(29)	66(37)	42(33)	91(9) [c]
U	CH_2	74(10) [a]	–	–	29	74 [b]
CH_2	CH_2	17(5) [a]	14(13)	26(13)	42(8)	17 [b]

Units: J mol^{-1} $(mol\ kg^{-1})^{-1}$.

[a] Values obtained by the fitting of the data of Table 6, CONH being the –CONH– or –CON– group in *trans* conformation or the $-CONH_2$ group.

[b] Values assumed as fixed parameters.

[c] Value obtained by the fitting of the data of Table 2, CONH being, in this case, the –CONH– or –CON– group in *cis* conformation.

TABLE 4

Values of the coefficients of the excess enthalpies used to derive $\{H_{i,j}\}$ for diketopiperazines in water at 298.15 K [20]

Solutes		n^x_{CONH} [a]	n^y_{CONH} [a]	$n^x_{CH_2}$ [a]	$n^y_{CH_2}$ [a]	$h_{xy(exp)}$ [b]	$h_{xy(calc)}$ [b,c]	$h_{xy(calc)}$ [b,d]
x	*y*							
G_2dkp	G_2dkp	2	2	2	2	−1138	−1085	−429
GAdkp	GAdkp	2	2	3	3	−509	−637	22
A_2dkp	A_2dkp	2	2	4	4	−213	−156	505
GVdkp	GVdkp	2	2	5	5	121	358	1022
Sar_2dkp	Sar_2dkp	2	2	5	5	577	358	1022
G_2dkp	U	2	0	2	0	−917	−900	−242
GAdkp	U	2	0	3	0	−783	−826	−168
A_2dkp	U	2	0	4	0	−674	−751	−93
Sar_2dkp	U	2	0	5	0	−758	−677	−19
Sar_2dkp	MMU	2	0	5	1.5	−339	−278	382
Sar_2dkp	MEU	2	0	5	2.5	−77	−13	648
Sar_2dkp	MPU	2	0	5	3.5	388	253	915
Sar_2dkp	MBU	2	0	5	4.5	486	518	1182

Abbreviations G, A, V and Sar represent glycine, alanine, valine and sarcosine, respectively.

[a] Number of groups on the solute molecule *x* or *y* as in Table 2.

[b] Units: J mol^{-1} $(mol\ kg^{-1})^{-1}$.

[c] Calculated by using the $\{H_{i,j}\}$ values of the last column of Table 3.

[d] Calculated by using the $\{H_{i,j}\}$ values of the third column of Table 3 (amides and linear peptides).

to calculate, with the same coefficients, the h_{xy} coefficients for the diketopiperazine systems produce greatly overestimated values, far beyond the reasonable standard deviations. This means that the set of data for the cyclic dipeptides is not homogeneous with that chosen as a reference. Therefore, the h data reported in Table 4 must be fitted separately from those of Table 2, by using the mean values of $\{H_{ij}\}$ reported in the last column of Table 3. Inspection of columns 3 and 7 of Table 3 shows that the $\{-\text{CONH}-\}_{trans}$ conformer behaves differently from the $\{-\text{CONH}-\}_{cis}$ conformer with respect to U and –CONH– interaction. Whatever the molecular mechanism accounting for the difference, the group contribution approach can discriminate these stereochemical differences. By contrast, the single value of $\{H_{\text{CONH,CH}_2}\}$ proves that the orientational effects (important for interactions between polar groups) play only a marginal role when the $-\text{CH}_2-$ group is involved, regardless of its partners. Accumulation of similarly behaving groups may also invalidate the predictions of the present method, since cooperative or concerted interactions may easily occur.

HYDROPHILIC SOLUTES (CARBOHYDRATES)

After the original, simple proposal that the solution properties of sugars in water could be described in terms of ideal equilibria between a series of hydrated solutes, studies with spectroscopic and dynamic methods have provided some evidence for the "stereochemical" model of hydration [29]. The term "specific hydration" has been mainly used to describe the concerted interaction between water molecules and the polar sites of the solute by hydrogen bonds. Because of the nature of the hydrogen bond and its orientation-dependent potential, it has been inferred that "specific hydration" strongly depends upon the detailed stereochemistry of the interacting groups. Therefore, properties of carbohydrates differing in the steric arrangements could be accounted for, in principle, by conformational factors. Apart from the dynamics of water and solute molecules, the case of carbohydrates is particularly complex owing to the presence of different stereochemical forms. This complexity arises from conformational ring interconversions (not to speak of anomeric interconversion) and the specificity claimed for the water–carbohydrate interaction must, therefore, be viewed with caution. The existence of several conformers suggests that the population of the various conformational states will change in different solvent conditions. As a consequence, it is very hard to establish, at the macroscopic level, an "a priori" dependence of the physicochemical properties of the carbohydrates in solution on their constitution. Fortunately, only a very few of the 26 different pyranoid rings may be important for the equilibrium thermodynamics. The study of conformations and compositions of the equilibrium mixtures formed by sugars in solution (and especially in water) is of great

TABLE 5

Virial coefficients of the excess enthalpies of sugars, polyols, and related compounds at 298.15 K [33]

Compound	$h_{xx(exp)}$ [a]	$h_{xx(calc)}$ [a]	h_{xxx} [b]	g_{xx} [a]
Aldopentoses				
D-Arabinose	177±17	125	10±9	
L-Arabinose	178±9	125	7±5	
D-Ribose	202±8	125	−6±5	
D-Lyxose	243±7	125	−11±4	
L-Xylose	336±8	125	–	
D-Xylose	339±16	125	−19±3	34
Aldohexoses				
D-Galactose	133±8	210	–	
D-Mannose	207±14	210	−14±5	
D-Glucose	343±10	210	−13	70
Ketohexoses				
D-Fructose	264±18	499	−7±4	
L-Sorbose	395±9	499	−16±4	
Deoxysugars				
2-Deoxy-D-galactose	442±22	662	−21±15	
2-Deoxy-D-ribose	468±12	542	−21±8	
2-Deoxy-D-glucose	592±17	662	−36±12	
6-Deoxy-L-mannose (L-rhamnose)	685±32	662	−38±20	140±34
6-Deoxy-galactose (L-fucose)	700±16	662	−21±10	131±36
α-Methylpyranoses				
α-Methyl-D-galactose (α-MGalP)	900±25	928	−42±14	11±7
α-Methyl-D-glucose (α-MGluP)	1097±39	928	–	149±8
α-Methyl-D-xylose (α-MXylP)	1126±5	844	−38±3	92±6
α-Methyl-D-mannose (α-MManP)	1206±14	928	−105±10	122±10
β-Methylpyranoses				
β-M-GalP	1081±28	928	−215±54	
β-M-GluP	1048±32	928	−64±32	
β-M-Xylp	1098±12	844	−78±6	
Oligosaccharides				
Lactose	506±32	491	–	
Sucrose	577±6	491	−34	181±5
Raffinose	811±50	1023	–	333±12
Polyols				
Ethylene glycol	362±4	160	−6	15±1
Glycerol	251	148		37
Mesoerythritol	358±22	140	−11±10	
Pentaerythritol	395±4	641	–	
Adonitol (ribitol)	295±5	76	–	
L-Arabitol	185±3	76	–	
D-Arabitol (lyxitol)	187±3	76	–	
Xylitol	80±11	76	5±4	
Mannitol	66±12	−24	20±12	18±2
Sorbitol	−11±5	−24	24±3	
Dulcitol (galactitol)	−132±50	−24	222±138	
Perseitol (α-mannoheptitol)	−299±20	−161	122±58	

[a] Units: J mol^{-1} $(mol\ kg^{-1})^{-1}$.
[b] Units: J mol^{-1} $(mol\ kg^{-1})^{-2}$.

importance in carbohydrate chemistry. For these reasons, attempts to devise geometrical fits based on the interaction mechanism may, in the case of flexible molecules such as carbohydrates, produce misleading results, since the thermodynamic properties are averaged in the energy space and hence not related to a single structural state, although some states may be more populated than others. The two different approaches, in fact, reflect the very large jump from thermodynamics to interaction mechanisms. The wide gap, geometry versus statistics, originates from the intrinsic difference between the minimum energy approach and that of free-energy space.

A number of new thermal data for carbohydrates at finite concentration have appeared in the very recent literature [30]. The major contribution to the data for enthalpies of dilution is provided by Barone et al. [17,19,31–35], in a series of recent papers where direct calorimetric experiments were carried out over quite a large concentration range. The results are analyzed in terms of the polynomial expansions. In a few cases, they only give significant values of h_{xxx}. The values of the coefficients h_{xx} and the 95% confidence limits are reported in Table 5. The monomeric carbohydrates can be grouped in three different ranges of h_{xx} values

(a) pentoses and hexoses (h_{xx} values ranging from 100 to 400 J mol^{-2} kg);

(b) deoxysugars (400–700 J mol^{-2} kg);

(c) methylglycosides (900–1200 J mol^{-2} kg).

The di- and tri-saccharides studied showed h_{xx} values approximately corresponding to the sum of the h_{xx} values of the constituent monomers. Table 5 also reports the few g_{xx} coefficients available from osmotic coefficient data (isopiestic measurements) or other literature data. The scarcity of Gibbs energy data precludes the drawing of conclusion about similarity of behavior, nonetheless, it can be seen that all g_{xx} and h_{xx} values are positive. The data listed in Table 5 are taken as a selection of representative properties and substances.

TABLE 6

Coefficients $\{H_{ij}\}$ of compounds of Table 5 [a]

$\{H_{ij}\}$	Ref. 33	Refs. 14,38
$\{H_{CH_2\text{-}CH_2}\}$	43±4	40±8
$\{H_{CHOH\text{-}CHOH}\}$	13±4	−4±3
$\{H_{O\text{-}O}\}$	−146±28	−116±102
$\{H_{CHOH\text{-}O}\}$	−49±9	-
$\{H_{OH\text{-}OH}\}$	−8	−26
$\{H_{CHOH\text{-}CH_2}\}$	31±6	33±8
$\{H_{O\text{-}CH}\}$	79±9	71±32

Units: J mol^{-1} (mol kg^{-1})$^{-1}$.

[a] Assumptions made for fitting experimental data are reported in the text for peptide–peptide interactions.

We also report in Table 6 the values of the coefficients calculated with the group contribution method. Although the choice of constituent groups is quite arbitrary and may generate a long debate (of little significance in the present article), it is clear that more data are necessary in order to reveal specific effects, if any, as has been found for diketopiperazines. For a discussion of these results, the reader is referred to the series of articles quoted in ref. 19.

The general analysis of the concentration dependence of the available thermodynamic data rules out the simple static hydration model put forward by some authors in the past and brings out some more specific effects which depend upon the stereochemical environment of the solute and are dynamically mediated by water.

CONCLUSIONS

We have shown that heat of dilution data properly presented in the form of a virial expansion can provide useful information. In the case of clear association processes, the virial coefficients are related to both the equilibrium constant and the enthalpy change for the formation of aggregates. In other cases, an approach based on the group contribution factorization can give insight into the interactions occurring in complex systems, and provides a useful scheme for detecting anomalous patterns.

ACKNOWLEDGMENTS

This paper originated from the author's long-lasting collaboration with Prof. G. Barone and Prof. V. Crescenzi. Financial support from the Ministry of Public Education through the University of Trieste is gratefully acknowledged.

REFERENCES

1 F. Franks and D.S. Reid, in F. Franks (Ed.), Water—A Comprehensive Treatise, Vol. 2, Plenum Press, New York, 1973.
2 W.G. McMillan and J.E. Mayer, J. Chem. Phys., 13 (1945) 276.
3 J.E. Mayer, J. Chem. Phys., 18 (1950) 1426.
4 H.L. Friedman, J. Chem. Phys., 32 (1966) 1351.
5 J.J. Kozak, W.S. Knight and W. Kauzmann, J. Chem. Phys., 48 (1968) 675.
6 G.N. Lewis and M. Randall, Thermodynamics, McGraw-Hill, New York, 1961.
7 I.M. Klotz and R.M. Rosenberg, Chemical Thermodynamics, Benjamin, Menlo Park, 1972.
8 H.L. Friedman, J. Solution Chem., 1 (1972) 387.

9 H.L. Friedman, J. Solution Chem., 1 (1972) 413.
10 H.L. Friedman, J. Solution Chem., 1 (1972) 419.
11 H.L. Friedman and C.V. Krishnan, J. Solution Chem., 2 (1973) 119.
12 J.L. Fortier, P.A. Leduc, P. Picker and J.E. Desnoyers, J. Solution Chem., 2 (1973) 467.
13 F. Franks, M. Pedley and D.S. Reid, J. Chem. Soc., Faraday Trans. 1, 72 (1976) 359.
14 J.J. Savage and R.H. Wood, J. Solution Chem., 5 (1976) 733.
15 I.R. Tasker and R.H. Wood, J. Solution Chem., 7 (1982) 295, 469.
16 I.R. Tasker and R.H. Wood, J. Phys. Chem., 86 (1982) 4040.
17 G. Barone, P. Cacace, G. Castronuovo and V. Elia, Carbohydr. Res., 119 (1983) 1.
18 A. Cesáro, E. Russo and G. Barone, Int. J. Peptide Protein Res., 20 (1982) 8.
19 G. Barone, G. Castronuovo, D. Doucas, V. Elia and C.A. Mattia, J. Phys. Chem., 87 (1983) 1931.
20 G. Barone, P. Cacace, V. Elia and A. Cesáro, J. Chem. Soc., Faraday Trans. 1, 80 (1984) 2073.
21 S.J. Gill, M. Downing and G.F. Sheats, Biochemistry, 6 (1967) 272.
22 M.G. Marenchic and J.M. Sturtevant, J. Phys. Chem., 77 (1973) 544.
23 A. Cesáro, E. Russo and V. Crescenzi, J. Phys. Chem., 80 (1976) 335.
24 A. Cesáro, E. Russo and D. Tessarotto, J. Solution Chem., 9 (1980) 221.
25 J.A. Schellman, C.R. Trav. Lab. Carlsberg, Ser. Chim., 29 (1955) 223.
26 G.M. Blackburn, T.H. Lilley and E. Walmsley, J. Chem. Soc., Faraday Trans. 1, 76 (1980) 915.
27 G.M. Blackburn, T.H. Lilley and E. Walmsley, J. Chem. Soc., Faraday Trans. 1, 78 (1982) 1641.
28 R.H. Wood and L.H. Hiltzik, J. Solution Chem., 9 (1980) 45.
29 F. Franks, Philos. Trans. R. Soc. London., Ser. B, 278 (1977) 33.
30 A. Cesáro, in H. Hinz (Ed.), Thermodynamic Data for Biological Sciences, Springer Verlag, Berlin, in press.
31 G. Barone, P. Cacace, G. Castronuovo and E. Elia, Carbohydr. Res., 91 (1981) 101.
32 G. Barone, P. Cacace, G. Castronuovo, V. Elia and F. Iappelli, Carbohydr. Res., 93 (1981) 11.
33 G. Barone, P. Cacace, G. Castronuovo and V. Elia, Gazz. Chim. Ital., 112 (1982) 153.
34 G. Barone, P. Cacace, G. Castronuovo, V. Elia and U. Lepore, Carbohydr. Res., 115 (1983) 15.
35 G. Barone, P. Cacace, G. Castronuovo and V. Elia, J. Solution Chem., 13 (1984) 209, 625.
36 I.R. Tasker and R.H. Wood, J. Solution Chem., 12 (1983) 315.
37 F. Franks and M.D. Pedley, J. Chem. Soc., Faraday Trans. 1, 79 (1983) 2249.
38 B.V. Okamoto, R.H. Wood, J.E. Desnoyers, G. Perrow and L. Delorme, J. Solution Chem., 10 (1981) 139.