ELSEVIER

Biophysical Chemistry 51 (1994) 359–374

Biophysical Chemistry

Quantitative evaluation of hydration thermodynamics with a continuum model

Alexander A. Rashin [a], Lynn Young [b], Igor A. Topol [b]

[a] *BioChemComp Inc, 543 Sagamore Ave, Teaneck, NJ 07666, USA*

[b] *Structural Biochemistry Program, Frederick Biomedical Supercomputing Center, NCI-Frederick Cancer R&D Center, PRI / DynCorp., Frederick, MD 21702, USA*

Received 23 December 1993; accepted 30 January 1994

Abstract

We attempt to analyze whether experimental entropies, enthalpies and free energies of hydration of small uncharged molecules can be quantitatively rationalized with a continuum model including a classical reaction field formalism. We find that a simple proportionality to accessible surface with five different atom types allows satisfactory (within 1–1.5 kcal/mol) reproduction of hydration entropies ($T\Delta S$) of over 40 solutes. The agreement with experiment can possibly be improved if proximity effects and configurational contributions to transfer entropies are taken into account. In calculations of hydration enthalpies a reasonable agreement with experimental data can be obtained only when solute polarizability is taken into account. Electrostatic contributions to calculated hydration enthalpies exhibit strong dependencies on both the magnitude and the direction of molecular dipole moments. We demonstrate that for 20 molecules with experimentally measured vacuum dipole moments density functional calculations with DZVPD basis set including diffuse functions on d-orbitals allows prediction of experimental dipole moments within 0.1 D. At a fixed direction of the molecular dipole moment, μ, the electrostatic component of hydration enthalpy varies as μ^2. Thus an uncertainty of 0.1 D corresponds to uncertainties of 0.5–0.7 kcal/mol in hydration enthalpies of most small dipolar solutes. A 30° change in the direction of the molecular dipole together with the corresponding change in the quadrupole moment can result in a change of hydration enthalpy of 3 kcal/mol. Changes in the quadrupole moment alone can result in hydration enthalpy changes of over 1 kcal/mol. Representations of multipole expansions by point charges on nuclei fitted to molecular electrostatic potentials cannot accurately reproduce all these factors. Use of such point charges in calculations of hydration enthalpies predictably leads to discrepancies with experiment of ≈ 3 kcal/mol for some solutes. However, errors in hydration enthalpies and hydration entropies are usually compensating leading in most cases to agreement between calculated and experimental free energies of hydration within 1.5 kcal/mol.

Key words: Hydration; Continuum model; Thermodynamics

1. Introduction

Many chemical and most biochemical reactions occur in solution, and thermodynamics of hydration plays a critical part in all such reactions (e.g., see refs. [1–5]). Thus a quantitative understanding of the thermodynamics of hydration is crucial for understanding a multitude of impor-

SSDI 0301-4622(94)00058-R

tant phenomena including protein folding and protein–drug interactions. A variety of empirical schemes, integral equation and computer simulation techniques exists for evaluation of the thermodynamics of hydration (see refs. [1–5] for reviews). Most of the methods can be used for evaluation of only one thermodynamic characteristic of hydration (often the free energy). Empirical schemes are usually designed for evaluations of the free energy alone, however, in the molecular dynamics or Monte Carlo perturbation methods the inability to calculate thermodynamic characteristics other than the free energy comes from large errors (of the order of the result itself) in calculated enthalpies and entropies of hydration [1]. Integral equation techniques, which can be used for the evaluation of free energies, enthalpies and entropies of hydration, have not been tested on a large number of solutes, and in a number of cases yielded results that significantly deviate from experimental values [1]. A very successful empirical scheme for evaluation of a variety of thermodynamic characteristics of hydration based on group additivity has been reported [6,7]. A closer look shows, however, that the very good agreement with experimental data in this scheme comes at a price of physically unreasonable values of hydration entropies and enthalpies of some groups. In particular, enthalpies and entropies of hydration of tetragonal CH and C, and of trigonal C are assigned rather large positive values. It is difficult to imagine that transfer of any of these groups from vacuum to water could be significantly unfavourable enthalpically, or would increase the entropy of water. These non-physical parameters probably reflect an inadequacy of the group-additive approach. However, they also lead to an important question: "Is a physically reasonable and accurate method with a small number of parameters feasible?"

In the last eight years a physically transparent reaction-field approach based on classical continuum electrostatics became increasing popular [1,8–12]. This approach uses numerical solutions of the Poisson (or Poisson–Boltzmann) equation for a system of charges inside a dielectric boundary of an arbitrary shape. Numeric solutions can be obtained with finite difference (FDM) [10,13], boundary element (BEM) [8,14] or finite element (FEM) [15] methods. The reaction potential due to polarization of the continuum medium by charges can be easily extracted from the numeric solutions. Half of the energy of interaction of charges with this reaction potential is usually taken as the electrostatic part of enthalpy [8] or free energy [10,12] of hydration (see ref. [16] in this Issue for a discussion). This is, however, correct only for spherically symmetric systems [1,8]. For solutes that are not spherically symmetric effects of a mutual polarization of the solute and the solvent can make significant contributions (up to 30%) to the thermodynamics of hydration [1,8,12]. Solute polarization can be incorporated into the continuum formalism through point atomic polarizabilities within BEM [8], and through an assignment of a dielectric constant to solutes and an artificial scaling of their charges within FDM [12] (and, possibly, FEM). As the reaction field incorporates only electrostatic components of the solute-medium interactions, empirical terms are required to account for 'non-electrostatic' terms [8,12]. Continuum approaches based on classical electrostatics demonstrated a degree of success in predicting hydration energies (within ≈ 2 kcal/mol of experimental values) similar to that of molecular simulations [8,12]. This leads to two major questions: (1) How far physically clear albeit simplified continuum theories can go in reproducing physically meaningful picture of hydration phenomena? (2) How close can these theories come to accurately reproducing experimental thermodynamics of hydration? The first question is addressed in ref. [16] in this Issue. An attempt to consider the second one is the subject of this report. While we have not arrived at a final answer, we identify a number of factors critical for achieving a high accuracy by continuum methods. Disregard for these factors may disqualify claims for accuracy from some continuum as well as molecular studies of hydration.

2. Methods

2.1. Electrostatic contribution to hydration energetics

Details of BEM used here for the numeric solution of the Poisson equation can be found elsewhere [1,8,17]. Here we present only a brief summary of the method. The effects of mutual polarization of a polar solute and a solvent can be expressed through the density of the surface polarization charges, σ_i, that can be determined from the boundary conditions for the normal components of the electric field and the electric displacement

$$\sigma_i = \frac{1}{4\pi}\frac{D_{in}-D_{out}}{D_{in}} E_{out}\cdot n_i, \qquad (1)$$

and through dipoles, μ_m, induced on atoms, characterized by point polarizabilities, α_m,

$$\mu_m = \alpha_m\left(E_{su}+E_\sigma+E_\mu\right), \qquad (2)$$

where $D_{in}=1$ and $D_{out}=78$ are 'inside' and 'outside' dielectric constants characterizing 'solute' and solvent parts of the system; σ_i characterizes the surface density of polarization charges that is approximated by a constant on the surface of ith element of the dielectric boundary ($1 \leqslant i \leqslant$ ne); n_i is a normal to the i-th boundary element directed to the outside medium at some 'representative' point, r_i, of this element, and E_{out} is the electric field at the 'representative' point, r_i, on the surface of the boundary element [8,17]. E_{out} has contributions from the charge distribution of the solute, E_{su}, and from polarization charges on all boundary elements, E_σ, and from dipoles induced on solute atoms, E_μ. Eqs. (1) and (2) can be transformed into a system of linear algebraic equations for densities of surface polarization charges on all boundary elements and for components of induced dipoles on all atoms of the solute,

$$\sum_j^{l,ne} K_{ij}\sigma_j + \sum_m^{l,na}\alpha_m\sum_a^{1,3} V^a_{im}\mu^a_m = -\sum_a^{1,3} E^a_{su}n^a_i, \qquad (3)$$

$$\alpha_m\sum_i^{1,ne} U^a_{mi}\sigma_i + \alpha_m\sum_{j\neq m}^{1,na}\sum_a^{1,3} W^a_{mj}\mu^a_m = -\alpha_m E^a_{su}, \qquad (4)$$

where ne is the number of surface elements, na is the number of atoms, the superscript a denotes components of a vector, and all summations over repeating indeces are explicit. More detailed expressions for the coefficients of these equations can be found in ref. [8]. After σ_i and μ_i are found from simultaneous solution of Eqs. (3) and (4), the total electrostatic contribution to the free energy of hydration, ΔG_{el}, can be calculated as [1]

$$\Delta G_{el} = \tfrac{1}{2}\left(\sum_m q^{su}_m\sum_j \sigma_j S_j/|r_m-r_j|\right) + \tfrac{1}{2}\left(\sum_n q^{su}_n\sum_{m\neq n}\mu_m S_j(r_m-r_n)/|r_m-r_n|^3\right). \qquad (5)$$

Here we use a simple scheme of non-interacting induced point dipoles [8] which is reasonable for small molecules. In this scheme atomic charges obtained by a quantum mechanical calculation of the entire solute do not interact with induced dipoles on atoms of this solute [8]. Therefore, the second term in Eq. (5) is set to zero [8]. Atomic polarizabilities, α_m, are taken from ref. [18]. It should be noted that the electrostatic part of the enthalpy of hydration equals the electrostatic part of the free energy of hydration within $\approx 1.77\%$ [8]. Thus, ΔG_{el} increased by 1.77% is used here as the electrostatic part of hydration enthalpy, ΔH_{el}. Radii of atomic cavities used here for the positioning of the dielectric boundary are taken from ref. [17] and reproduced in Table 1. It can be noted that only 5–6 experimental van der Waals radii or corresponding water–solute equilibrium distances are used to calculate all geometric parameters of atoms (e.g., cavity radii are derived from them as described in refs. [8,17]). All hydrogens are treated explicitly in electrostatic calculations (note that cavity radii are different for hydrogens of polar and of hydrocarbon groups [8,17]).

2.2. Entropies of hydration

Entropy calculations with a semicontinuum approach [16,19–21] yield $T\Delta S$ per unit of accessible surface area, SA, between -26 cal/(mol Å^2)

Table 1
Atomic radii used in calculations (Å) [a]

N/NH	O_C	O_H	H_C	H_{ON}	C_{tetr}	C_{trig}	CH_n	SH	Computation of
1.4	1.4	1.4			1.7	1.7	1.725	1.84	ΔH_{np}
1.4	1.4	1.4	1.0		1.7	1.7		1.84	$T\Delta S$
1.5	1.5	1.5	1.71	1.16	2.46	2.46		1.97	ΔH_{el}

[a] Atomic radii used for calculations of ΔH_{np} and $T\Delta S$ correspond to the values widely used in the literature. Radius of 'united atom' CH_n is taken from ref. [17]. Cavity radii used for calculations of ΔH_{el} are obtained from those in previous lines as described in ref. [17]. Accessible surface area, SA, in Eqs. (6) and (7) were calculated with the probe radius 1.4 Å as described in ref. [17]. Atom types are explained in legend to Table 2.

for a nonpolar sphere of 2 Å radius and −35 cal/(mol Å^2) for a sphere of 10 Å radius [21]. A polar hydrogen-type sphere with a positive charge 0.3 e yields $T\Delta S$ of −39 cal/(mol Å^2) [16]. Spheres with a negative charge of −0.5 e and radii 1.4–2.0 Å yield $T\Delta S$ of −18 cal/(mol Å^2), and can drop to −13 cal/(mol Å^2) for larger spheres [16]. These results indicate that hydration entropies per unit surface area can be more negative for polar groups with positively charged hydrogens (e.g., OH or NH), or less negative for polar negatively charged groups (e.g., O in CO) than for non-polar groups. A linear fit of alkane $T\Delta S$ of hydration at 298 K [7] yields: $T\Delta S = -1.3884 - 0.0284 \times SA$ (five data points from ethane to hexane are included). A similar fit of seven $T\Delta S$ for alcohols (data points from ethanol to 1octanol [7] are included) yields: $T\Delta S = -1.0833 - 0.0351 \times SA$. A fit of five $T\Delta S$ for amines (data points from ethanamine to 1hexaneamine [7] are included) yields: $T\Delta S = -2.1616 - 0.0310 \times SA$. All three linear fits are shown in Fig. 1. The slopes of these fits qualitatively agree with magnitudes and trends obtained in semicontinuum calculations of hydration entropies of spherical particles [16,21]. As our goal is to construct a physically reasonable model with a small number of parameters, we assume a simple expression for calculations of hydration entropies

$$T\Delta S_{\text{molecule}} = T \sum_{i}^{\text{SA-types}} \Delta s_i \, SA_i, \qquad (6)$$

where i denotes types of accessible surface area for different types of groups, and Δs_i is not allowed to be far beyond the range obtained in semicontinuum calculations of hydration entropies. Atom/group types used in our calculations of hydration entropies of ≈50 molecules are listed in Table 2. Note that there are only four different values of $\Delta s_i - s$. The $(T\Delta s_i)$ value of −25 cal/(mol Å^2) for trigonal carbons, that is less negative than the corresponding value for tetragonal carbons, can be quantitatively justified by polarity (and small negative charge on aromatic carbons). The value for OH and NH groups (−45 cal/(mol Å^2)) is somewhat more negative than the estimate from semicontinuum calculations of hydration entropies (−39 cal/(mol Å^2)) [16,19]. We discuss it further in section 3 below.

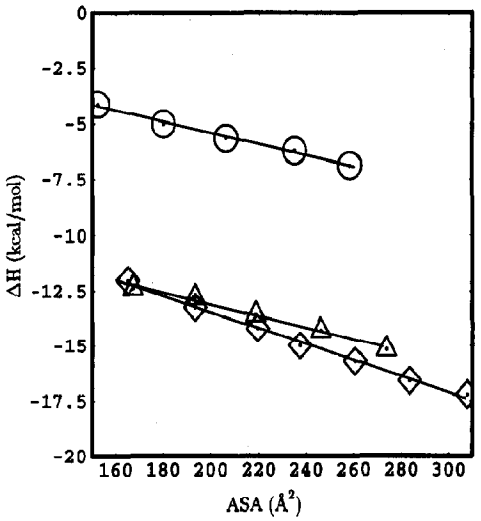

Fig. 1. Linear plots of hydration entropies versus accessible surface areas for alkanes (○), alcohols (□), and amines (△).

Table 2
Contributions of accessible surfaces of individual atoms to $T\Delta S$ of hydration (kcal/(mol Å^2))

N/NH [a]	O_C [b]	O_H [c]	H_C [d]	H_{ON} [e]	C_{tetr} [f]	C_{trig} [g]	CH_n [h]	SH [i]	Computation of
0.045	0.015	0.045	0.035		0.035	0.025		0.035	$T\Delta S$

[a] Nitrogen or NH group.
[b] Oxygen of a carbonyl group or non-hydrogen bearing oxygen of carboxyl group.
[c] Hydrogen-bearing oxygens of hydroxyl and carboxyl groups.
[d] Hydrocarbon hydrogen.
[e] Hydrogen-bonding hydrogens on oxygen or nitrogen.
[f] Tetragonal carbon.
[g] Trigonal carbon.
[h] United atom for methyl ar methylene groups.
[i] Sulphhydryl group.

2.3. Non-electrostatic parts of the enthalpy of hydration

A transfer of molecules from vacuum to water leads not only to electrostatic interactions between solute molecules and the solvent but also to dispersion interactions [17]. Assuming that for non-polar solutes all of the enthalpy of hydration is equal to the energy of solute–solvent dispersion interactions (which is a somewhat simplistic assumption, see section 3), one can estimate the magnitude of non-electrostatic part of the enthalpy of hydration. A linear fitting of hydration enthalpies of alkanes as a function of their accessible surface area yields $\Delta H_{np} = -0.369 - 0.0252$ SA (kcal/mol). A similar fitting of total hydration enthalpies of alcohols and amines yield slopes of -0.0367 kcal/(mol Å^2) and -0.0269 kcal/(mol Å^2), correspondingly. We take the average value of the slope for all these solutes as a measure of the magnitude of non-electrostatic contribution to hydration enthalpy of methyl and methylene groups and assume that

$$\Delta H_{np} = -0.369 - 0.02986\ \mathrm{SA}. \quad (7)$$

In calculations of SA for this contribution hydrogens were not taken into account, and a 'united atom' is used for CH_n groups (see Table 1). The validity of this approximation is further discussed in section 3.

Total enthalpy of hydration, ΔH_{hydr}, is assumed to be [17]

$$\Delta H_{hydr} = \Delta H_{np} + \Delta H_{el}. \quad (8)$$

2.4. Quantum mechanical calculations and fitting of atomic charges

All quantum mechanical calculations in this work have been performed with ab initio GAUSSIAN 92 [22] and density functional (DFT) DGauss [23] [#1] programs. Calculations utilizing GAUSSIAN 92 were performed with 6-31G* basis set, and Merz–Kollman fitting procedure was used to fit charges to electrostatic potentials [24,25]. Calculations utilizing DGauss were performed with DZVP2 [26] basis set for the structure optimization and with DZVPD basis set [27] for calculations of electric multipoles and electrostatic potentials. Molecular structural parameters were optimized at the local spin density (LSD) level using Vosko–Wilk–Nusair (VWN) exchange–correlation potential [28]. Electric multipoles and electrostatic potentials were calculated for LSD optimized structures by self-consistent inclusion of nonlocal and gradient corrections to the exchange–correlation potential (NLSD approximation) according to Becke (exchange) [29] and Perdew (correlation) [30]. In all calculations we used for electron density and exchange–correlation potential auxiliary Gaussian basis sets (8/4/4) for heavy atoms and (4/1) for hydrogens. Our tests have shown that molecular dipoles calculated with DZVPD basis set from geome-

[#1] DGauss and MNDO91 are available as a part of Unichem software from Cray Research, Eagan, MN.

tries optimized with DZVP2 or DZVPD basis sets are practically identical. It has been shown for four dipolar molecules [27] that calculations with DZVPD basis set lead to values of molecular dipole moments that are within 0.1 D of their experimental value. DZVPD basis set contains diffuse functions on heavy atoms as described in ref. [27].

Point charges representing molecular charge distributions are derived by fitting to the DGauss DZVPD electrostatic potential. Given a set of N grid points from DGauss, the electrostatic potential at each point i ($1 \leqslant i \leqslant N$) is

$$V^{DG}(r_i) = \sum_A \frac{Z_A}{|R_A - r_i|} - \int \frac{\rho(r')}{|r_i - r'|} dr'. \quad (9)$$

Σ_A refers to the nuclei and $\rho(r')$ is the electron density. (The probe at r_i is of unit charge.)

Alternatively, the potential at each grid point can be calculated from the partial charges of the molecule:

$$V^{Q}(r_i) = \sum_A \frac{q_A}{|R_A - r_i|}. \quad (10)$$

If the partial charges are unknown, they can be calculated by fitting to $V^{DG}(r_i)$ as follows. Minimization of the function

$$f(q) = \left[(V^{D} - V^{Q}) + \lambda \left(Q_{TOT} - \sum_A q_A \right) \right]^2 \quad (11)$$

(where λ is a Lagrangian multiplier for the charge conservation term) yields a set of n equations for n unknown partial charges.

$$\begin{aligned}
&\sum_i V^{D}(r^i) \left(\frac{1}{|R_1 - r_i|} + \lambda \right) \\
&\quad = \sum_i \sum_A q_A \left(\frac{1}{|R_A - r_i|} + \lambda \right) \left(\frac{1}{|R_1 - r_i|} + \lambda \right) \\
&\quad \vdots \\
&\quad \sum_i V^{DG}(r^i) \left(\frac{1}{|R_n - r_i|} + \lambda \right) \\
&\quad = \sum_i \sum_A q_A \left(\frac{1}{|R_A - r_i|} + \lambda \right) \left(\frac{1}{|R_n - r_i|} + \lambda \right).
\end{aligned} \quad (12)$$

We choose the strongest condition for charge conservation of $\lambda = 1$. All molecules in this study are electrically neutral.

Since DGauss does not allow a user to request specific positions for the grid points, we used grids with cubic cells and density 5–7 points/Å extending 3 Å beyond the molecule. Van der Waal radii were used to remove grid points close to nuclei. For elongated molecules we also used grids with default rectangular cells. The size of the cells could also be changed to higher or lower grid density. The actual choice of the cell shape (cubic or rectangular) and the density of the grid was done as follows. Depending on the shape of the molecule we started with a cubic or rectangular grid. If fitting from a low density grid led to a good reproduction of the DZVPD calculated molecular dipole and its components, then the fitting was considered complete, and the resulting charges were used in BEM calculations. If the fitted charges reproduced the DZVPD dipole reasonably well but not exactly, they were linearly scaled to reproduce the DZVPD dipole. If a DZVPD dipole or its components were reproduced poorly, grids and their densities would be modified until a reasonable agreement was achieved between molecular dipoles from DZVPD calculations and the fitted charges. For pyridine and its derivatives dummy charges were added on lone pairs to achieve such an agreement.

Dipole and quadrupole moments of systems of point charges (obtained with the fitting procedure described above or taken from the literature) were calculated with the usual formulae:

$$\mu_a = \sum_i r_a^i q_i, \quad (13)$$

$$Q_{ab} = \sum_i \left(3 r_a^i r_b^i - \delta_{ab} |r^i|^2 \right) q_i. \quad (14)$$

All quantum calculations were performed on the Cray YMP at the Frederick Biomedical Supercomputing Center.

3. Results and discussion

3.1. Accuracies of multipole distributions and hydration enthalpies

Electrostatic contributions to hydration enthalpies depend on the accuracy of the represen-

Table 3
Effect of scaling of charges on the magnitude of the electrostatic part of hydration enthalpy (kcal/mol) of ethanol

dipole moment, μ [a]	1.14	1.24	1.34	1.44	1.54	1.64	1.74	1.84	1.94	2.04	2.14
ΔH_{el} (BEM) [b]	−2.481	−2.935	−3.430	−3.953	−4.529	−5.143	−5.774	−6.467	−7.178	−7.946	−8.743
$\Delta H_{el} = \Delta H(\mu_0)(\mu/\mu_0)^2$ [c]	−2.485	−2.940	−3.434	−3.965	−4.535	−5.134	−5.790	−6.475	−7.197	−7.958	−8.757

[a] Molecular dipole moments in D: the magnitude of the dipole moment was changed by multiplying all atomic charges by the same constant.
[b] Electrostatic part of hydration enthalpy calculated from the given geometry, atomic charges and cavity radii with BEM.
[c] Electrostatic part of hydration enthalpy calculated from μ_0 with the equation given in this row.

tation of molecular charge distributions by point charges. Thus, the first question in our analysis of the feasibility of a quantitative evaluation of the hydration thermodynamics is concerned with sensitivity of computed results to inaccuracies in charge representations. A dependence of the calculated electrostatic contribution to hydration enthalpy on the magnitude of the dipole moment, μ, when its orientation is fixed is shown in Table 3 for the example of ethanol. DFT calculations with a DZVPD basis set for ethanol yield the value of the dipole moment $\mu_0 = 1.64$ D, which can be reproduced by some set of point charges on nuclei. The first line of Table 3 lists eleven values of the dipole moment, μ (five smaller than μ_0 and five larger than it) obtained by a linear scaling (i.e. by multiplication of all point charges by a constant). The second row of the table shows that ΔH_{el} changes 0.4–0.8 kcal/mol when μ changes by 0.1 D. A comparison of the second and third rows of the table shows that with high accuracy

$$\Delta H_{el}(\mu) = \left(\frac{\mu}{\mu_0}\right)^2 \Delta H_{el}(\mu_0). \qquad (15)$$

Table 3 shows results only for ethanol, but similar results were obtained for other solute molecules. These results show that an error in the gas phase dipole moment of only 0.1 D leads to an error of 0.4–0.7 kcal/mol in calculated hydration enthalpies (or electrostatic parts of hydration free energies). Errors of 0.5 D lead to errors of 2.5–3.5 kcal/mol in calculated hydration enthalpies (free energies). A recent work [31] that varied point charges to achieve good agreement between FDM calculated and experimental free energies of hydration at the expense of up to 0.5 D differences between experimental and calculated vacuum dipole moments conceals actual errors in calculated hydration energies of 3.5 kcal/mol or even larger (see below).

The second question in our feasibility analysis concerns the effect of the orientation of the dipole moment with a fixed magnitude on the calculated enthalpy of hydration. Table 4 provides data for such an analysis using the example of methylamine. Three sets of point charges on atoms of methylamine yield dipole moments of the same magnitude $\mu = 1.31$ D. However, they yield vastly different electrostatic hydration enthalpies: −2.52, −3.36 and −5.50 kcal/mol. The differ-

Table 4
Enthalpy dependence on dipole orientation [a]

Atoms and partial charges of methylamine (e)							Dipole and components (D)				Components of quadrupole				ΔH_{el} (kcal/mol)
C	H	H	H	N	H	H	μ	μ_x	μ_y	μ_z	Q_{xx}	Q_{yy}	Q_{zz}	Q_{xy}	
−0.260	0.122	0.142	0.142	−0.505	0.180	0.180	1.31	1.072	−0.752	−0.002	−3.977	1.906	2.071	3.324	−2.52
0.228	0.000	0.000	0.000	−0.673	0.222	0.222	1.31	1.275	−0.301	−0.002	−5.120	2.992	2.132	3.859	−3.36
0.297	−0.058	−0.009	−0.009	−0.877	0.328	0.329	1.31	1.258	0.365	−0.002	−8.766	6.558	2.207	4.310	−5.50

[a] Molecular geometry in vacuum has been obtained by DFT-'DGauss' optimization with DZVP2 basis set: C(0.005, −0.005, 0.000); H(1.118, −0.020, 0.000); H(−0.336, −0.573, 0.885); H(−0.336, −0.573, −0.885); N(−0.609, 1.310, 0.000); H(−0.330, 1.848, −0.826); H(−0.326, 1.850, 0.824). Atomic charges has been obtained by different procedures and linearly scaled to reproduce experimental dipole moment (1.31 D [32]). Components of the dipole (D) and quadrupole (D Å) moments have been calculated from coordinates and charges as described in the text.

Table 5
Enthalpy (kcal/mol) dependence on quadrupole moment (D Å)

Water model charge centres	Charges on water (e)					Dipole μ	Quadrupole			ΔH_{el} (kcal/mol)
	O	H	H	LP	LP		Q_{xx}	Q_{yy}	Q_{zz}	
SPC-type	−0.64	0.32	0.32	0.00	0.00	1.80	0.32	2.60	−2.92	−8.26
Rowlinson		0.32	0.32	−0.32	−0.32	1.80	1.14	3.42	−4.58	−9.63

[a] The geometry, coordinates of charges and their magnitude are taken from ref. [33] (the oxygen of water is at the origin of the coordinate system, and hydrogens are in the XY plane at equal distances from the positive X axis). For SPC-type water model charges similar to that of Rowlinson's model are places on centers of three atoms to make results comparable. Atomic polarizabilities are taken from ref. [18] as in the rest of this work.

ence between the second and third orientations of the dipole is only about 30° which leads to the difference in hydration enthalpies of 2.14 kcal/mol. It should be noted that changes in the orientation of the dipole moment also lead to changes in the quadrupole moment.

Thus, our third question concerns the effect of changes in the quadrupole moment at fixed magnitude and orientation of the dipole moment on the calculated enthalpy of hydration. Table 5 provides data for an estimation of the magnitude of this effect for the example of Rowlinson's four-charge model of water [33] and of its modification with only three charges. This example shows that the hydration enthalpy difference due to the difference in quadrupole moments alone can easily reach 1.3 kcal/mol. The importance of quadrupole moments in accurate calculations of intermolecular interaction energies has been recognized [1,34]. This leads to our next question: 'How well can we predict dipole and quadrupole moments?'

Table 6
Gaussian-calculated, DFT-calculated and experimental dipole moments (D)

No. Molecule	Dipole moment (D) [a]			
	631-G*	DZVP2(A2)/VWN	DZVPD(A2)/BP	exp.
1 ammonia	1.80	1.83	1.48	1.47
2 water	2.35	2.16	1.86	1.85
3 acetamide	4.27	4.04	3.84	3.76
4 *N*-methylacetamide	4.09	3.99	3.76	3.73
5 methanol	1.90	1.85	1.59	1.70
6 ethanol	1.89	1.90	1.64	1.69
7 1-propanol	1.91	1.92	1.68	1.68
8 1 butanol	1.92	1.94	1.69	1.66
9 2-propanol	1.76	1.78	1.52	1.58
10 methylamine	1.42	1.47	1.21	1.31
11 ethylamine	1.38	1.37	1.15	1.22
12 1-propaneamine	1.34	1.31	1.11	1.17
13 2-propanone	3.35	3.20	3.00	2.88
14 2-butanone	3.20	3.01	2.82	2.78
15 acetic acid	2.00	1.75	1.72	1.70
16 propanoic acid	2.11	1.81	1.78	1.75
17 methylbenzene	0.25	0.47	0.43	0.38
18 ethylbenzene	0.20	0.43	0.44	0.59
19 phenol	1.51	1.41	1.28	1.22
20 pyridine	2.35	2.42	2.20	2.22

[a] Calculations with 6-31G* basis set were performed with GAUSSIAN 92 [22]. Calculations with DZVP2 and DZVPD basis sets were performed with DGauss DFT program [23]; results of geometry optimizations with DZVP2 basis set were used in single point DZVPD [27] calculations (see the text).

3.2. Accuracy in predictions of vacuum dipole moments

Experimental and calculated (with different quantum mechanical methods) dipole moments of 20 molecules are listed in Table 6. A comparison of dipole moments in the first two columns shows that dipole moments calculated with the GAUSSIAN 92 6-31G* basis set and with DGauss DZVP2 basis set are similar. However, their comparison with experimental dipole moments in the fourth column shows that dipole moments calculated with these basis sets deviate from experimental values by 0.15–0.35 D. Thus, according to our arguments above, these dipole moments would lead to errors in calculated electrostatic contributions to hydration enthalpies (free energies) of 1–2 kcal/mol. Dipole moments calculated with DGauss DZVPD basis set that includes diffuse functions on heavy atoms agree with experimental values within 0.1 Debye, with only a few exceptions where the deviation is only slightly larger (the deviation of 0.15 Debye for ethylbenzene cannot lead to significant errors in hydration enthalpies because the magnitude of the dipole moment is small). To the best of our knowledge such accuracies in calculated dipole moments have not been reported previously. Accuracies of dipole moments calculated with the DFT DZVPD basis suggest a possibility of accuracies of 0.4–0.7 kcal/mol in calculated hydration enthalpies. It should be noted, however, that we do not know how accurately directions of the dipole moments are predicted, and this, as we have shown above, can also lead to significant errors. In addition, representation of actual charge distributions with point charges on nuclei usually leads to arbitrary values of quadrupole moments introducing additional uncertainties. Note also that 'experimental' dipole moments can vary form one edition of Handbook of Chemistry and Physics [32] to another: The 1991 edition gives for phenol $\mu = 1.45$ while the 1992 edition gives $\mu = 1.22$ D.

3.3. Calculations of hydration enthalpies

The analysis above leads to an expectation that hydration enthalpies cannot be predicted with high accuracy until multipole expansions used in the calculations are accurate enough. Of course, we are not guaranteed that our or any other available method will lead to accurate enthalpy predictions even with accurate multipoles. However, then we at least will know that the problem is not caused by incorrect multipoles. Here we analyze results of computations of hydration enthalpies with representations of molecular charge distributions by atomic charges and with other techniques described above in section 2.

Tables 7 and 8 summarize results of our calculations of enthalpies, entropies and free energies of hydration for ≈ 50 solutes. Table 7 contains results obtained with structures and charges from DFT optimizations, and Table 8 contains corresponding results obtained with Gaussian 92. In both cases atomic charges were fitted to electrostatic potentials as described in section 2. Thus, results of Table 7 are obtained with atomic charges reproducing experimental dipole moments within ≈ 0.1 D, while results of Table 8 are obtained with 6-31G* molecular dipole moments that can significantly deviate from experimental values (e.g., see Table 6). In this subsection we concentrate on columns 2–6 of Tables 7 and 8 concerning hydration enthalpies and leave analysis of results of calculations of entropies and free energies of hydration to subsections below.

A first insight comes from the analysis of nonpolar contributions to hydration enthalpies, ΔH_{np} (third column of Tables 7 and 8). For alkanes the electrostatic contribution to hydration enthalpy, ΔH_{el}, is close to zero (see first five entries in the tables). The difference between calculated, ΔH_{hyd}, and experimental, ΔH_{exp}, hydration enthalpies of alkanes are in the range 0.7–1.2 kcal/mol indicating that the approximation for the non-electrostatic part of hydration enthalpy (Eq. (7)) is not very accurate for alkanes. This is due to our averaging the slopes of ΔH versus accessible areas, SA (Fig. 2) of alkanes, alcohols and amines. As slopes for alkanes and amines are very similar (0.0252 and 0.0269 kcal/(mol Å^2) respectively), the only source of error for alkanes can come from the slope of ΔH/SA plot for alcohols. Attributing the slope of this plot to the contribution of the hydrocarbon part of alcohol

Table 7
Experimental and DZVPD calculated thermodynamic characteristics of hydration (kcal/mol) [a]

No.	Molecule	ΔH_{el}	ΔH_{np}	ΔH_{hyd}	ΔH_{exp}	$\Delta\Delta H$	$T\Delta S$	$T\Delta S_{exp}$	$\Delta T\Delta S$	ΔG	ΔG_{exp}	$\Delta\Delta G$
1	ethane	0.00	−4.91	−4.91	−4.12	*0.79*	−5.79	−5.95	*−0.16*	0.88	1.83	*0.95*
2	propane	−0.02	−5.74	−5.76	−5.01	*0.75*	−6.76	−6.97	*−0.21*	1.00	1.95	*0.95*
3	butane	0.00	−6.52	−6.52	−5.65	*0.88*	−7.79	−7.72	*0.07*	1.27	2.08	*0.81*
4	pentane	−0.03	−7.30	−7.40	−6.21	*1.19*	−8.91	−8.54	*0.37*	1.51	2.33	*0.82*
5	hexane	0.00	−8.07	−8.08	−6.88	*1.19*	−9.88	−9.37	*0.51*	1.80	2.48	*0.68*
6	22-dimethylpropane	0.00	−6.92	−6.92	−5.45	*1.47*	−8.39	−7.95	*0.44*	1.47	2.50	*1.03*
7	methanol	−5.22	−4.48	−9.70	−10.23	*−0.54*	−5.58	−5.12	*0.46*	−4.12	−5.11	*−0.99*
8	ethanol	−5.03	−5.29	−10.12	−12.02	*−1.71*	−6.61	−7.01	*−0.40*	−3.71	−5.01	*−1.30*
9	1-propanol	−5.27	−6.14	−11.41	−13.23	*−1.81*	−7.69	−8.40	*−0.71*	−3.72	−4.82	*−1.10*
10	1-butanol	−5.68	−6.91	−12.59	−14.19	*−1.60*	−8.70	−9.48	*−0.78*	−3.89	−4.71	*−0.82*
11	1-pentanol	−6.34	−7.67	−14.01	−14.99	*−0.98*	−9.79	−10.52	*−0.73*	−4.22	−4.47	*−0.25*
12	1-hexanol	−8.29	−8.40	−16.69	−15.70	*0.99*	−10.91	−11.34	*−0.43*	−5.78	−4.36	*1.42*
13	1-heptanol	−7.31	−9.32	−16.63	−16.58	*0.05*	−11.85	−12.34	*−0.49*	−4.78	−4.24	*0.54*
14	1-octanol	−5.95	−10.20	−16.15	−17.21	*−1.06*	−12.80	−13.12	*−0.32*	−2.70	−4.09	*−1.39*
15	2-propanol	−4.61	−6.14	−10.75	−13.36	*−2.61*	−7.60	−8.60	*−1.00*	−3.15	−4.75	*−1.60*
16	2-butanol	−4.43	−6.82	−11.24	−14.44	*−3.19*	−8.48	−9.86	*−1.38*	−2.76	−4.57	*−1.81*
17	4-heptanol	−5.39	−9.20	−14.59	−17.44	*−2.86*	−11.42	−13.44	*−2.02*	−3.19	−4.00	*−0.84*
18	2-methyl-2-propanol	−5.37	−6.64	−12.01	−14.72	*−2.71*	−8.36	−10.21	*−1.94*	−3.65	−4.51	*−0.86*
19	2-propanone	−5.25	−6.06	−11.30	−9.22	*2.08*	−6.12	−5.37	*0.75*	−5.18	−3.85	*1.33*
20	2-butanone	−4.43	−6.76	−11.18	−10.37	*0.81*	−7.23	−6.74	*0.49*	−3.95	−3.64	*0.31*
21	2-heptanone	−5.66	−9.14	−14.80	−13.05	*1.75*	−10.35	−10.01	*0.34*	−4.45	−3.04	*1.41*
22	3-methyl-2-butanone	−4.76	−7.42	−12.18	−10.48	*1.41*	−7.88	−7.54	*0.34*	−4.40	−3.24	*1.16*
23	4-methyl-2-pentanone	−4.95	−8.25	−13.20	−10.85	*2.36*	−8.86	−7.79	*1.07*	−4.34	−3.06	*1.28*
24	33-dimethyl-2-butanone	−4.31	−7.90	−12.21	−11.37	*0.84*	−8.63			−3.58		
25	acetic acid	−7.32	−5.67	−12.99	−12.07	*0.93*	−5.77	−5.37	*0.40*	−7.22	−6.70	*0.52*
26	propanoic acid	6.99	6.42	−13.40	−12.95	*0.45*	−6.87	−6.48	*0.39*	−6.53	−6.47	*0.06*
27	methylamine	−6.10	−4.58	−10.68	−10.30	*0.38*	−5.58	−5.74	*−0.16*	−5.10	−4.56	*0.54*
28	ethanamine	−5.14	−5.36	−10.50	−12.36	*−1.86*	−6.70	−7.86	*−1.16*	−3.80	−4.50	*−0.70*
29	1-propanamine	4.99	−6.14	−11.13	−12.77	*−1.64*	−7.71	−8.38	*−0.67*	−3.42	−4.39	*−0.97*
30	1-butanamine	−5.90	−6.89	−12.79	−13.56	*−0.77*	−8.69	−9.26	*−0.57*	−4.10	−4.29	*−0.19*
31	1-pentanamine	−6.10	−7.71	−13.80	−14.29	*0.49*	−9.62	−10.20	*−0.58*	−4.18	−4.09	*0.09*
32	1-hexanamine	−5.58	−8.53	−14.38	−15.16	*−0.78*	−10.70	−11.13	*−0.43*	−3.68	−4.03	*−0.35*
33	2-propanamine	−4.15	−6.13	−10.28	−12.60	*−2.32*	−7.59			−2.69		
34	12-ethandiamine	−10.66	−5.74	−16.40	−17.63	*−1.23*	−7.58			−8.85		
35	2-methyl-1-propanamine	−4.36	−6.78	−11.13	−13.10	*−1.97*	−8.44			−2.69		
36	acetamide	−10.57	−5.71	−16.28	−16.32	*−0.04*	−5.83	−6.62	*−0.79*	−10.31	−9.7	*0.75*
37	N-methylacetamide	−8.52	−6.60	−15.12	−17.06	*−1.94*	−7.02			−8.10		
38	N-ethylacetamide	−9.43	−7.52	−16.94	−18.65	*−1.71*	−8.08			−8.86		
39	N-propylacetamide	−10.30	−8.25	−18.55	−19.88	*−1.33*	−9.14			−9.41		
40	N-butylacetamide	−10.14	−8.92	−19.06	−20.88	*−1.82*	−9.98			−9.08		
41	benzene	−1.01	−7.00	−8.01	−7.04	*0.97*	−6.80	−6.18	*0.62*	−1.20	−0.86	*0.34*
42	methylbenzene	−0.93	−7.55	−8.47	−8.12	*0.35*	−7.98	−7.23	*0.75*	−0.49	−0.89	*−0.40*
43	ethylbenzene	−1.33	−8.52	−9.85	−9.07	*0.79*	−8.92	−8.27	*0.65*	−0.93	−0.80	*0.13*
44	phenol	−5.44	−7.18	−12.62	−13.06	*−0.44*	−7.73	−6.44	*1.29*	−5.10	−6.61	*−1.73*
45	2-methylphenol	−4.41	−8.24	−12.65	−14.43	*−1.78*	−8.77	−8.55	*0.22*	−3.88	−6.87	*−1.99*
46	pyridine	−3.75	−6.74	−10.48	−11.36	*−0.88*	−6.71	−6.66	*0.05*	−3.77	−4.70	*−0.93*
47	2-methylpyridine	−3.74	−7.47	−11.21	−12.50	*−1.29*	−7.96	−7.87	*0.09*	−3.25	−4.63	*−1.38*
48	2-ethylpyridine	−1.71	−8.25	−9.96	−12.23	*−2.27*	−8.83	−7.90	*0.93*	−1.13	−4.33	*−3.20*
49	23-dimethylpyridine	−2.13	−8.11	−10.24	−13.24	*−3.00*	−8.85	−8.41	*0.44*	−1.39	−4.82	*−3.43*
50	26-dimethylpyridine	−2.84	−8.46	−11.33	−14.26	*−2.93*	−9.22	−9.66	*−0.44*	−2.11	−4.60	*−2.49*

[a] Geometry optimization for all molecules in Table 7 has been done with DGauss [23] DZVP2 basis set [26], and all partial charges were obtained by fitting to electrostatic potentials from DZVPD [27] calculations as described in the text. The electrostatic contribution to the enthalpy of hydration, ΔH_{el}, is calculated as described by Eq. (5) and the text immediately following it; non-electrostatic contribution to the enthalpy oa hydration, ΔH_{np}, is calculated with Eq. (7); total calculated enthalpy of hydration, ΔH_{hyd}, is the sum of ΔH_{el} and ΔH_{np} (Eq. (8)); $T\Delta S$ is calculated according to Eq. (6) with Δs_i from Table 2; free energy of hydration, ΔG, is calculated as $\Delta G = \Delta H_{hyd} - T\Delta S$; all experimental values, except that for acetamide [35], are from refs. [6,7]; al $\Delta\Delta$s are differences between experimental and calculated values, and they are shown in italic print.

Table 8
Experimental and 6-31G* calculated thermodynamic characteristics of hydration (kcal/mol)

No.	Molecule	ΔH_{el}	ΔH_{np}	ΔH_{hyd}	ΔH_{exp}	$\Delta\Delta H$	$T\Delta S$	$T\Delta S_{exp}$	$\Delta T\Delta S$	ΔG	ΔG_{exp}	$\Delta\Delta G$
1	ethane	0.00	−4.80	−4.80	−4.12	*0.68*	−5.79	−5.95	*−0.16*	0.99	1.83	*0.84*
2	propane	−0.04	−5.78	−5.82	−5.01	*0.81*	−6.76	−6.97	*−0.21*	0.94	1.95	*1.01*
3	butane	−0.06	−6.52	−6.58	−5.65	*0.94*	−7.79	−7.72	*0.07*	1.21	2.08	*0.87*
4	pentane	−0.05	−7.36	−7.41	−6.21	*1.20*	−8.91	−8.54	*0.37*	1.50	2.33	*0.83*
5	hexane	−0.05	−8.07	−8.13	−6.88	*1.25*	−9.88	−9.37	*0.51*	1.75	2.48	*0.73*
6	22-dimethylpropane	−0.20	−7.03	−7.23	−5.45	*1.78*	−8.39	−7.95	*0.44*	1.16	2.50	*1.34*
7	methanol	−6.75	−4.52	−11.28	−10.23	*1.04*	−5.58	−5.12	*0.46*	−5.70	−5.11	*0.59*
8	ethanol	−6.81	−5.27	−12.08	−12.02	*0.05*	−6.61	−7.01	*−0.40*	−5.47	−5.01	*0.46*
9	1-propanol	−6.69	−6.24	−12.93	−13.23	*−0.30*	−7.69	−8.40	*−0.71*	−5.24	−4.82	*0.42*
10	1-butanol	−6.57	−6.88	−13.46	−14.19	*−0.73*	−8.70	−9.48	*−0.78*	−4.76	−4.71	*0.06*
11	1-pentanol	−6.86	−7.80	−14.65	−14.99	*−0.33*	−9.79	−10.52	*−0.73*	−4.86	−4.47	*0.39*
12	1-hexanol	−6.73	−8.61	−15.34	−15.70	*−0.36*	−10.91	−11.34	*−0.43*	−4.43	−4.36	*0.07*
13	1-heptanol	−6.67	−9.45	−16.12	−16.58	*−0.46*	−11.85	−12.34	*−0.49*	−4.27	−4.24	*0.03*
14	1-octanol	−6.68	−10.21	−16.89	−17.21	*−0.33*	−12.80	−13.12	*−0.32*	−4.09	−4.09	*0.00*
15	2-propanol	−6.62	−6.13	−12.76	−13.36	*−0.60*	−7.60	−8.60	*−1.00*	−5.16	−4.75	*0.39*
16	2-butanol	−6.45	−6.86	−13.31	−14.44	*−1.13*	−8.48	−9.86	*−1.38*	−4.83	−4.57	*0.25*
17	4-heptanol	−6.34	−9.16	−15.51	−17.44	*−1.94*	−11.42	−13.44	*−2.02*	−4.09	−4.00	*0.09*
18	2-methyl-2-propanol	−6.33	−6.75	−13.14	−14.72	*−1.58*	−8.36	−10.21	*−1.94*	−4.78	−4.51	*0.27*
19	2-propanone	−6.65	−5.99	−12.64	−9.22	*3.42*	−6.12	−5.37	*0.75*	−6.52	−3.85	*2.67*
20	2-butanone	−6.06	−6.81	−12.86	−10.37	*2.43*	−7.23	−6.74	*0.49*	−5.64	−3.64	*2.00*
21	2-heptanone	−6.40	−9.11	−15.51	−13.05	*2.46*	−10.35	−10.01	*0.34*	−5.16	−3.04	*2.12*
22	3-methyl-2-butanone	−6.39	−7.33	−13.71	−10.48	*2.94*	−7.88	−7.54	*0.34*	−5.83	−3.24	*2.59*
23	4-methyl-2-butanone	−5.92	−8.12	−14.04	−10.85	*3.19*	−8.86	−7.79	*1.07*	−5.18	−3.06	*2.12*
24	33-dimethyl-2-butanone	−5.97	−7.94	−13.91	−11.37	*2.53*	−8.63			−5.28		
25	acetic acid	−10.17	−5.71	−15.88	−12.07	*3.81*	−5.77	−5.37	*0.40*	−10.11	−6.70	*3.41*
26	propanoic acid	−9.45	−6.37	−15.81	−12.95	*2.86*	−6.87	−6.48	*0.39*	−8.94	−6.47	*2.47*
27	methylamine	−6.97	−4.52	−11.49	−10.30	*1.19*	−5.58	−5.74	*−0.16*	−5.91	−4.56	*1.35*
28	ethanamine	−7.20	−5.29	−12.49	−12.36	*0.13*	−6.70	−7.86	*−1.16*	−5.79	−4.50	*1.29*
29	1-propanamine	−7.53	−6.10	−13.63	−12.77	*0.86*	−7.71	−8.38	*−0.67*	−5.92	−4.39	*1.53*
30	1-butanamine	−7.43	−7.03	−14.46	−13.56	*0.90*	−8.69	−9.26	*−0.57*	−5.77	−4.29	*1.48*
31	1-pentanamine	−7.48	−7.83	−15.30	−14.29	*1.01*	−9.62	−10.20	*−0.58*	−5.68	−4.09	*1.58*
32	1-hexanamine	−7.51	−8.70	−16.21	−15.16	*1.04*	−10.70	−11.13	*−0.43*	−5.51	−4.03	*1.48*
33	2-propanamine	−7.58	−6.19	−13.77	−12.60	*1.17*	−7.59			−6.18		
34	12-ethandiamine	−11.58	−5.97	−17.55	−17.63	*−0.08*	−7.58			−9.97		
35	2-methyl-1-propanamine	−7.32	−6.86	−14.18	−13.10	*1.08*	−8.44			−5.74		
36	acetamide	−13.74	−5.71	−19.45	−16.32	*3.13*	−5.83	−6.62	*−0.79*	−13.62	−9.7	*3.92*
37	N-methylacetamide	−10.92	−6.68	−17.60	−17.06	*0.54*	−7.02			−10.58		
38	N-ethylacetamide	−10.34	−7.48	−17.82	−18.65	*−0.83*	−8.08			−9.74		
39	N-propylacetamide	−10.30	−8.40	−18.70	−19.88	*−1.18*	−9.14			−9.56		
40	N-butylacetamide	−10.28	−8.90	−19.18	−20.88	*−1.70*	−9.98			−9.20		
41	benzene	−1.55	−6.96	−8.51	−7.04	*1.46*	−6.80	−6.18	*0.62*	−1.71	−0.86	*0.85*
42	methylbenzene	−1.29	−7.51	−8.79	−8.12	*0.68*	−7.98	−7.23	*0.75*	−0.81	−0.82	*−0.01*
43	ethylbenzene	−1.37	−8.21	−9.58	−9.07	*0.51*	−8.92	−8.27	*0.65*	−0.66	−0.80	*−0.14*
44	phenol	−7.15	−7.02	−14.17	−13.06	*1.12*	−7.73	−6.44	*1.29*	−6.44	−6.61	*−0.17*
45	2-methylphenol	−6.39	−7.90	−14.29	−14.43	*−0.14*	−8.77	−8.55	*0.22*	−5.52	−5.87	*−0.35*
46	pyridine	−4.07	−6.47	−10.54	−11.36	*−0.82*	−6.71	−6.66	*0.05*	−3.83	−4.70	*−0.87*
47	2-methylpyridine	−4.02	−7.25	−11.27	−12.50	*−1.23*	−7.96	−7.87	*0.09*	−3.31	−4.63	*−1.32*
48	2-ethylpyridine	−3.97	−8.04	−12.01	−12.23	*−0.22*	−8.83	−7.90	*0.93*	−3.18	−4.33	*−1.15*
49	23-diethylpiridine	−4.09	−7.96	−12.06	−13.24	*−1.18*	−8.85	−8.41	*0.44*	−3.21	−4.82	*−1.61*
50	26-dimethylpyridine	−4.04	−8.20	−12.24	−14.26	*−2.01*	−9.22	−9.66	*−0.44*	−3.02	−4.60	*−1.58*

[a] Geometry optimization for all molecules in Table 8 has been done with MNDO91 [22] AM1, and all partial charges were obtained by fitting to electrostatic potentials from 6-31G* calculations [24,25] as described in the text. For the rest of notations, see footnote to Table 7.

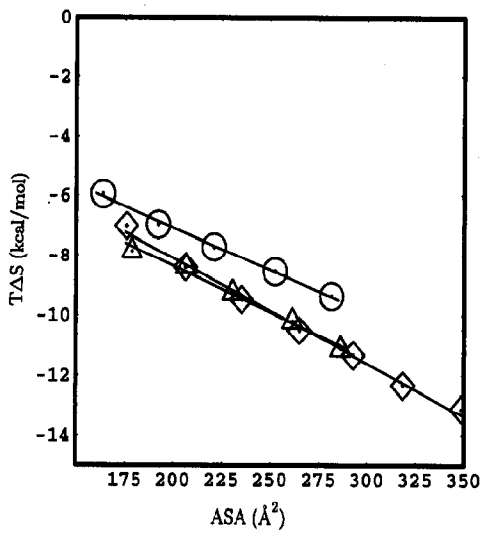

Fig. 2. Linear plots of hydration enthalpies versus accessible surface areas for alkanes, (○), alcohols (□), and amines (△).

chains is justified only if we assume that such contributions are additive for chains longer than methanol (4). This, however, is not necessarily correct. Table 7 shows that the electrostatic contribution, ΔH_{el}, to the enthalpy of hydration increases for the larger alcohols and this may account for larger $\Delta H/SA$ slope for alcohols than for alkanes. However, Table 8 does not show any such increase in ΔH_{el} for larger alcohols. Thus, the nature of the difference in slope for alcohols and alkanes remains unexplaned. One would expect that results of Table 7 are closer to reality because they are based on more accurate dipole moments. This expectation, however, cannot be taken as fully justified because of the uncertainties in the quadrupole moments. Note that the slope under discussion is characteristic only for alcohols and not for amines, and, thus, cannot be explained by a 'nonspecific' difference between nonpolar and polar chains. Calculated hydration enthalpies for alcohols are in all but a few cases smaller in absolute value than the corresponding experimental values. Thus, decreasing the contribution of hydrocarbon groups, that would lead to a better prediction for alkanes, would make predictions for alcohols rather poor. A resolution of this contradiction should await accurate dipole and quadrupole representations of molecular charge distributions. Note that alcohols with hydroxyl group flanked on both sides by methyl groups (secondary alcohols) yield larger discrepancies between the predicted and experimental hydration enthalpies than alcohols with hydroxyl group at the end of the chain (primary alcohols).

It is easy to see from Tables 7 and 8 that in most cases when absolute values of hydration enthalpies of dipolar molecules are larger than the corresponding experimental values, DGauss results (Table 7) are closer to experimental values. When absolute values of hydration enthalpies of dipolar molecules are smaller than the corresponding experimental values, Gaussian results (Table 8) are closer to experimental values. This comes from consistently larger dipole moments from Gaussian calculations leading to larger contributions to electrostatic part of hydration enthalpies in Table 8. Characteristically, discrepancies between experimental and calculated hydration enthalpies are within the range of −3.2–2.4 kcal/mol for DGauss based results and −2.0–3.8 kcal/mol for Gaussian based results, reflecting a larger electrostatic contribution in Gaussian based results. While errors in predicted hydration enthalpies are within the range expected from the analysis in the previous subsection, they are hardly satisfactory. They underscore a necessity for accurate multipole expansions representing molecular charge distributions. Until then we cannot be sure how far the accuracy of our calculations can be improved.

3.4. Calculations of hydration entropies

We can start our analysis of results of entropy calculations with an observation that thirty six out of forty two calculated hydration entropies (Tables 7 and 8) agree with the corresponding experimental values within 'chemical accuracy' (≈ 1 kcal/mol). Two deviations are near 2 kcal/mol, and four are between 1.07 and 1.38 kcal/mol. While this is satisfactory (mean absolute deviation of 0.6 kcal/mol), a further analysis of the results may reveal sources of discrepancies and can lead to further improvement in accuracies of predictions.

Analysis of discrepancies larger than 1 kcal/mol shows that three out of six such discrepancies are for alcohols with hydroxyl group flanked by hydrocarbon groups. All but one calculated hydration entropies of alcohols underestimate experimental loss of entropy upon hydration. One possible source of this underestimation is our use of the same constant per Å^2 of hydrocarbon accessible surface for alkanes and alcohols. Fig. 1 shows, however, that the slope of $T\Delta S/SA$ plot for alcohols is larger than for alkanes (0.0351 versus 0.0284 kcal/(mol Å^2)). The value of this slope for amines lies between the values for alcohols and alkanes. The most plausible explanation we could find for this discrepancy is that conformational entropies of alkanes and alcohols in the gas phase or in solutions are somewhat different, and this difference in conformational entropies is attributed to hydration entropy. For example, it is easy to imagine that conformations of alcohols with hydroxyl group in contact with a hydrocarbon group could be energetically favourable in vacuum, but would be unfavourable in solution because of the loss of hydration by hydroxyl group. This would lead to a narrower conformational space for alcohols in solution than in water, and to a larger total loss of entropy upon transfer from the gas phase to water in agreement with a steeper slope of $T\Delta S/SA$ plot for alcohols. This hypothesis, that cannot be accepted without further detailed verifications, underscores, however, that in both calculations and experiment we should deal with *a set* of conformations with similar enthalpies and entropies (or free energies that are expressed through them). In this work we deal with a single conformation that is assumed to be the same in both vacuum and solution. This is not necessarily correct. A rather cumbersome alternative would require finding quantum mechanically different energy minima in vacuum and in solution.

A part of the difference between the calculated and experimental entropies of hydration for alcohols is compensated by $\Delta s_i = 0.045$ kcal/(mol Å^2) for hydroxyl groups, which is too large. Applied to the water molecule it leads to $T\Delta S_{hyd} = -4.43$ kcal/mol, while the experimental value is -3.65 kcal/mol. We currently see only one physically reasonable way to correct the latter discrepancy. It is to assume that part of the hydrocarbon groups neighboring a hydroxyl group contribute more to the decrease of the water entropy than regular alkanes (proximity effects). This agrees with an observation that the discrepancy between the calculated and experimental hydration entropies increases for alcohols with the hydroxyl group more shielded from water by flanking hydrocarbon groups. We have not implemented such a modification for this study to avoid an introduction of additional arbitrary parameters at this stage. Errors in predicted hydration entropies of ketones ($\Delta s_i = 0.015$ kcal/(mol Å^2) for a carbonyl oxygen) can also be attributed to a neglect of proximity effects. The proximity effect cannot, however, extend beyond one or two methyl groups neighboring a hydroxyl and would not resolve the problem of a difference in slopes of $T\Delta S/SA$ plots (Fig. 1) for alkanes and alcohols.

Thus we have to deal with two major problems in attempts to increase accuracy of calculations of hydration entropies: (a) multiple conformations in vacuum and solution and related differences in conformational entropies, and (b) proximity effects.

3.5. Calculated free energies of hydration

We calculated free energies of hydration as differences between corresponding calculated enthalpies and entropies ($T\Delta S$) of hydration. Examination of Table 7 shows that for 27 out of 42 solutes calculated hydration free energies are within chemical accuracy (≈ 1 kcal/mol) from the corresponding experimental values; 35 out of 42 calculated values are within ≈ 1.5 kcal/mol from the corresponding experimental values. The success rate is only somewhat lower for Table 8 (22 and 30 respectively). The success rate is significantly higher for free energies than for enthalpies where only 21 out of 50 calculated hydration enthalpies are within chemical accuracy from the experimental values (Table 8 has a success rate of 22/50). A higher success rate for predicted free energies than for enthalpies of hydration is due to a compensation of enthalpy and entropy errors. The number of cases where such

compensation occurs is larger than the number of cases where the errors add.

Our success rate for predictions of free energies of hydration is similar to that reported for free energy simulations with 6-31G* derived charges and OPLS Lenard-Jones parameters [35]. Ref. [35] reports 12 out of 13 computed free energies to be within 2 kcal/mol from experimental values and that for acetamide within 3.7 kcal/mol of the experimental value (our 6-31G* acetamide $\Delta\Delta G = 3.9$). Analysis of Table G shows that in electrostatic calculations with correct (within 0.1 D) dipole moments and solute polarizabilities taken into account we often underestimate hydration enthalpies. Simulations of ref. [35] do not account for solute polarizabilities and thus can rely only on overestimation of vacuum dipole moments by 6-31G* calculations to mimic the increase in dipole moment of solutes upon their polarization in solution [25]. Such compensation is not, however, uniform (see Table 6) and is fortuitous. It would lead to errors in situations where polarization is decreased (e.g., for groups partially buried in proteins).

A degree of success in predicted free energies of hydration demonstrated in this study could be considered rather high by most accepted criteria (the average absolute error is 1.0 kcal/mol). However, the task of predicting protein–drug binding constants imposes higher demands on accuracies of computed free energies than those achieved here. Furthermore, we see that the accuracy of predicted free energies of hydration is often due to cancellation of errors in calculated enthalpies and entropies of hydration. Thus a full thermodynamic characterization of binding would be less accurate than that including only free energies. Lumping errors into some non-physical compensating parameters would lead to large errors for each changing problem. We have not resolved all of the difficulties yet, and this paper is rather a progress report than a solution of the problem acceptable for quantitative drug design. In principle, it is not guaranteed that they can be resolved within any existing method. However, a clarification of difficulties in developing quantitative methods for predictions of the thermodynamics of hydration constitute, in our opinion, an important step towards quantitative computerized drug design. We are currently working on accurate multipole representations of molecular charge distributions and expect to report results of their incorporation in our computational scheme in a near future.

Acknowledgment

We thank S. Burt, R. Cachau and P.L. Privalov for useful discussions and staff and administration of Frederick Biomedical Supercomputing Center for their support of this project. The content of this publication does not necessarily reflect the views of policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

References

[1] A.A. Rashin, Progr. Biophys. Mol. Biol. 60 (1993) 73.
[2] C.L. Brooks, M. Karplus and B.M. Pettitt, Advan. Chem. Phys. 71 (1988).
[3] P.L. Privalov and G.J. Gill, Advan. Prot. Chem. 39 (1988) 191.
[4] K.P. Murphy and E. Freire, Advan. Prot. Chem. 43 (1992) 313.
[5] C.J. Cramer and D.G. Truhlar, Rev. Comput. Chem. 6 (1994), in press.
[6] G.I. Makhatadze and P.L. Privalov, J. Mol. Biol. 232 (1993) 639.
[7] P.L. Privalov and G.I. Makhatadze, J. Mol. Biol. 232 (1993) 660.
[8] A.A. Rashin, J. Phys. Chem. 94 (1990) 1725.
[9] M.E. Davis and J.A. McCammon, Chem. Rev. 90 (1990) 509.
[10] K.A. Sharp and B. Honig, A. Rev. Biophys. Biophys. Chem. 19 (1990) 301.
[11] D. Bashford, Curr. Opin. Struct. Biol. 1 (1991) 95.
[12] B. Honig, K. Sharp and A-S. Yang, J. Phys. Chem. 97 (1993) 1101.
[13] H. Oberoi and N.M. Allewell, Biophys. J. 65 (1993) 48.
[14] B.I. Yoon and A.M. Lenhoff, J. Comput. Chem. 11 (1990) 1080.
[15] T.J. You and S.C. Harvey, J. Comp. Chem. 14 (1993) 484.
[16] A.A. Rashin and M.A. Bukatin, Biophys. Chem. 51 (1994).
[17] A.A. Rashin and K. Namboodiri, J. Phys. Chem. 91 (1987) 6003.
[18] Y.-K. Kang and M.-S. Jhon, Theoret. Chim. Acta 61 (1982) 41.

[19] A.A. Rashin and M.A. Bukatin, J. Phys. Chem. 95 (1991) 2942.
[20] A.A. Rashin and M.A. Bukatin, J. Phys. Chem. 97 (1993) 1974.
[21] A.A. Rashin and M.A. Bukatin, J. Phys. Chem. 98 (1994) 386.
[22] A.M. Frisch, G.W. Trucks, M. Heads-Gordon, P.M.W. Gill, M.W. Wong, J.B. Foresman, B.G. Johnson, H.B. Schlegel, M.A. Robb, E.S. Replogle, R. Comperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzales, R.L. Martin, D.J. Fox, D.J. DeFrees, J. Baker, I.J.B. Stewart and J.A. Pople, GAUSSIAN 92 (Gaussian Inc., Pittsburgh, PA, 1992).
[23] J. Andzelm and E.J. Wimmer, J. Chem. Phys. 96 (1992) 1250.
[24] U.C. Singh and P.A. Kollman, J. Comput. Chem. 5 (1984) 129.
[25] K.M. Merz Jr., J. Comput. Chem. 13 (1992) 749.
[26] N. Godbout, D.R. Salahub, J. Andzelm and E. Wimmer, Can. J. Chem. 70 (1992) 560.
[27] A.A. Rashin, M.A. Bukatin, J. Andzelm and A.T. Hagler, Biophys. Chem. 51 (1994).
[28] S.J. Vosko, L. Wilk and M. Nusair, Can. J. Chem. 58 (1980) 1200.
[29] A.D. Becke, J. Chem. Phys. 84 (1986) 4524.
[30] J.P. Perdew, Phys. Rev. B 33 (1986) 8822; 38 (1986) 7406 (E).
[31] D. Sitkoff, K.A. Sharp and B. Honig, J. Phys. Chem. 98 (1994) 1978.
[32] CRC Handbook of Chemistry and Physics, 73rd Ed., ed. D.R. Lide CRC Press, Boca Raton, 1992/1993).
[33] J.O. Hirschfelder, C.F. Curtiss and R.B. Bird, Molecular theory of gases and liquids (Wiley, New York, 1967).
[34] M.A. Spackman, J. Chem. Phys. 85 (1986) 6587.
[35] H.A. Carlson, T.B. Nguyen, M. Orosco and W.L. Jorgensen, J. Comput. Chem. 14 (1993) 1240.
[36] R.A. Pierotti, J. Phys. Chem. 69 (1965) 281.

Discussion to the paper by A. Rashin, L. Young and I. Topol

Comments

Y. Marcus

(1) The accessible surface area (ASA) is more difficult to calculate than the accessible volume. The latter can be obtained from the Leahy intrinsic volume, adequately expressed by the additive atom-and-bond scheme of McGowan and Abraham (cf. ref. [1]). The accessible volume is then calculated by analogy with spherical molecules of similar volume, taking into account the exclusion volume (i.e. the radius of the solvent). The main point, however, is that the ASA is linearly correlated with the accessible volume to a good approximation for molecules with relative molecular masses 60–600. (For spheres $V_s = (R/3)A_s$, $V_s = (V_s^{1/3}/4\pi)A_s$; if $200 < V_s < 400$ then the coefficient of A_s is between 5.8 and 7.4, i.e. 6.6 ± 0.6, i.e. within 10% constant; cf. ref. [2] for data showing this for practical solute molecules). In view of the uncertainty of the coefficient of SA in Eq. (7) (± 0.006 in 0.030, i.e. 20%) such a correlation is more than adequate.

(2) The hydration enthalpy calculated for methylamine from the dipole moment disregards any enthalpy contribution from hydrogen bonding of the methylamine and water. Both have donor and acceptor H-bonding functions and I expect that H-bonding should be not less important than dipole–dipole (or other multipole and induced dipole) interactions in such cases.

(3) It should be useful to have the sources of the experimental data for the thermodynamic functions of hydration of the molecules shown in Table 7 & 8 specified.

[1] Y. Marcus, J. Phys. Chem. 95 (1991) 8886.
[2] R.S. Pearlman, in: Partition coefficient determination and estimation, eds. W.J. Dunn, J.H. Block and R.S. Pearlman (Pergamon Press, Oxford, 1985).

Responses by A. Rashin to comments

To Y. Marcus

(1) Computer algorithms for the numeric evaluation of ASA are rather simple and are known for about 20 years. We programmed one around 1980, and another with analytical evaluation of the surface areas a few years ago. Therefore this is not a problem. Your argument on the linear relationship between ASA and volume is probably well applicable for small molecules we considered but it does not apply to the sphere with 10 Å radius that was included in the derivation of range of entropies/Å^2 that you refer to. For this sphere $V \approx 4000$, well beyond your range.

(2) The current wisdom (starting with S. Lifson at Weizmann Institute) on the energetics of hydrogen bonds adopted in most empirical poten-

tials is that it is well described by the combination of the electrostatics and Lennard-Jones repulsion. Quantum calculations confirm that numerically total energy of H-bond in vacuum is nearly equal to the energy of electrostatic interaction between partial charges on the donor and acceptor groups (see ref. [1] for a review). We reproduce quite well experimental hydration enthalpies for water and ammonia using only quantum mechanically calculated partial charges, polarizabilities, and the classical reaction field. If explicit account of hydrogen bonding is not necessary for these molecules, why to use it for others?

(3) The data are taken from Privalov's compilations referenced in the legend to Tables 7, 8.

[1] A.A. Rashin, Progr. Biophys. Mol. Biol. 60 (1993) 73–200.