ELSEVIER

Biophysical Chemistry 51 (1994) 375–392

Biophysical
Chemistry

Incorporation of reaction field effects into density functional calculations for molecules of arbitrary shape in solution

A.A. Rashin [a], M.A. Bukatin [b], J. Andzelm [c], A.T. Hagler [c]

[a] *BioChemComp, Inc., 543 Sagamore Ave, Teaneck, NJ 07666, USA,*
[b] *Computer Science Department, Brandeis University, Waltham, MA 02254, USA,*
[c] *Biosym Technologies, 9685 Scranton Rd., San Diego, CA 92121, USA*

Received 9 January 1994; accepted 27 January 1994

Abstract

An attempt is made to combine continuum reaction field approaches with DFT ab initio calculations for quantitative evaluation of solvation effects in chemical processes. The formalism of the combined method is delineated along with its possibilities and limitations, and applied to several small model systems. It is found that DFT can provide dipole moments in vacuum and in solution (e.g., for water) with accuracies (0.1 D) that have not been reported with other methods. The results obtained suggest that agreement within ~ 1 kcal/mole can be expected between calculated and experimental hydration enthalpies of polar uncharged solutes. The results for ions are not as consistent as for dipolar molecules, suggesting that accurate multipole representations of the electron density of solutes may be required especially for ionic solutes.

Key words: Reaction field effects; Density functional calculations; solvation effects

1. Introduction

Quantum mechanics is the most fundamental theoretical approach that can be applied to the study of molecular processes and can provide insights into their mechanisms that are not available from experiments and simpler empirical methods. However, quantum mechanics can be applied only to systems with small to moderate number of atoms. This does not allow direct quantum mechanical calculations of solvent or medium effects on chemically or biologically important molecules and reactions. Such effects can be crucial in solution chemistry and biochemistry, and mechanisms of enzymatic reactions. To account for medium effects in quantum mechanical calculations requires that some classical description of the medium be incorporated into the quantum mechanical formalism. There are two types of quantum mechanical formalism that can be used in a combined (hybrid) approach: ab initio [1,2] and semiempirical [3,4], as well as two major classical descriptions of the medium: an atomic [5,6] and a continuum one [7,8]. A part of a classical medium can often be represented in atomic detail while the rest of it can be represented as a continuum (e.g., refs. 7, 9, 10), or its grid analog [11].

Attempts to perform quantum calculations in solution started a few decades ago with quantum

SSDI 0301-4622(94)00059-S

calculations in a spherical or elliptic cavity in a continuum (e.g., [12]). This simplified approach is still alive [13-15], and is used in Gaussian 92 [16]. The advantage of this simplification is a fast calculation of the reaction fields in analytical form allowing for easy evaluation of gradients in geometry optimization. First attempts to do SCF calculations in molecular-shaped cavities in continuum are due to Tomasi's group [17]. This group continues to rectify and apply its approach [18-20]. Grant et al. [21] have reported a study along similar lines but with more efficient and consistent continuum method. Methodologically very similar approach was reported more recently by Colvin and Melius [22]. A recent attempt to perform semiempirical AM1 and MNDO quantum calculations in molecular-shaped cavities in a continuum has been presented by Ford and Wang [23,24]. Same semiempirical quantum methods were used in combination with Langevin dipoles representation of the solvent and with computationally much more expensive molecular simulations of the solvent by Luzhkov and Warshel [25]. Analysis of some methodological aspects of a similar combination of semiempirical quantum mechanics and continuum theory has been reported by Fox et al. [26]. Cramer and Truhlar combined semiempirical quantum mechanics with an approximation of the classical reaction field [27,28]. An interesting recent work combining semiempirical AM1 method with classical molecular simulations has been presented by Gao and Xia [29]. This work also contains references to a number of related attempts. Older attempts to do quantum calculations in proteins were reviewed by Drummond [30]. More recent examples are given in refs. [31-36]. A consistent empirical approach to the study of enzymatic mechanisms is advocated in ref. [37].

All of the above discussed studies have their drawbacks. While for small molecules it is possible to choose ellipsoidal cavity so that results are similar to those obtained with the molecular shape cavities [23,24], this is unlikely to be true for larger molecules. As the magnitude of the reaction field is very sensitive to the position of the dielectric boundary relative to molecular charges, errors of the order of 10 kcal/mol can be expected for asymmetric single charged ionic species [38]. While many studies report an agreement with experimental data, no consistent attempts have been made to find a sound basis for achieving such an agreement. Problems with conventional ab initio SCF calculations are often indicated [21,22]. In particular, results for water do not seem to converge to a physically reasonable value with increasing size of the basis set [21]. A failure to predict the magnitude of the solution dipole moment of water in AM1 + simulation study [29] is rather illustrative, as are significant discrepancies between the results obtained with AM1 and MNDO methods [23,24]. Inability of traditional quantum mechanical methods to provide accurate molecular dipole moments is, apparently, a persistent problem. While some methods [25,27,28] are made practically useful by fitting numerous parameters to reproduce experimental data, this significantly obscures the meaning of the agreement with experimental results. In some studies [31,34-36] only a small number of water molecules are included in calculation, which would lead to either a neglect or a large distortion of the reaction field and, thus, of the effect of the solvent. In other cases [32,33] when a large number of water molecules was included in calculations, they were submitted to energy minimization with the entire protein-water system, and then left static. This may reasonably reproduce the reaction field on the substrate, but is unlikely to provide a reliable estimate of the energetics of the system that requires either statistical mechanical or reaction field treatments.

It should be noted that semiempirical quantum methods can be applied to systems of a few hundred atoms, while conventional *ab initio* methods to about 20 atoms [4]. Conventional *ab initio* MP2 level calculation for 50 atoms can overload a single processor CRAY for a year [39]. In contrast, *ab initio* density functional theory (DFT) [1] calculations, that came to the forefront of computational quantum mechanics, can handled up to 100 atoms [33]. DFT calculations for 49 atoms (32 of them non-hydrogens) at about 6-31G ** level with correlations took relatively modest 1-2 h of CRAY Y-MP time [32]. Comparable calculation for 74 atoms took 10 h [33].

Thus DFT *ab initio* calculations came a long way in efficiency toward semiempirical methods, but can provide more insights into the physical nature of problems studied with their aid. Thus, a use of DFT *ab initio* methods for quantum calculations in solutions seems quite promising and may have a significant advantages over semiempirical methods. Quoting a recent review: "The most reliable quantum mechanical method is no doubt *ab initio*. There is a systematic way of correcting for the deficiencies of any level applied by progressing to a higher level of theory, and if this systematic progression is not successful, the failure becomes of theoretical interest. With a large enough basis set and sufficient correlation, *ab initio* methods should approach the solution of the time-independent Schrodinger equation" [4].

The last eight years also witnessed a remarkable revival of the continuum approach to the protein and solution energetics based on classical electrostatics, e.g., through numeric solutions of Poisson or Poisson–Boltzmann equations with the finite difference [8], the boundary element [7], and more recently, finite element [40] techniques. Such approaches are now widely accepted as a major proven tool for quantitative studies of the solution energetics [7,8,10,41–49]. It has been found by numerous researchers that the continuum approach yields remarkable agreement with experiment comparable to that of molecular simulations or quantum mechanics (see Refs. [7,8,21–25]).

The major change in the energetics of a chemical reaction in solution relative to that *in vacuo* is determined by the changes in hydration energies of reactants and products. If hydration energies cannot be calculated accurately, then calculated energetics of chemical and biochemical reactions in solution can hardly be reliable. The purpose of this work is: (1) *To combine the best* in modern quantum mechanics and classical methods *for quantum calculations in polarizable media*, and (2) *to analyze* capabilities and *limitations* of such combined approach. The ultimate goal of our effort is a quantum mechanical treatment of the substrate in the field of the protein in solution. In this communication we demonstrate that a combination of continuum reaction field approaches with DFT *ab initio* calculations can provide a reasonable basis for quantitative evaluation of solvation effects in chemical processes. We delineate the formalism to accomplish this, its possibilities and limitations, and apply it to several small model systems. In particular we find that DFT can provide dipole moments in vacuum and solution with accuracies (0.1 D) that have not been reported with other methods. Our results suggest that agreement within ≈ 1 kcal/mol can be expected between calculated and experimental hydration enthalpies of polar uncharged solutes. We, however, find that accurate multipole representations of the electron density of solutes are required especially for ionic solutes. In this communication we do not attempt to include medium effects into the geometry optimization [19]. These effects are expected to be relatively modest for the systems studied here and will be considered elsewhere.

2. Methods

2.1. Hybrid (quantum + reaction field) method

The basic equation of the approach is (see ref. [18])

$$(H_k^0 + V_\sigma)\Psi_k = E_k\Psi_k, \tag{1}$$

where H_k^0, is an unperturbed Hamiltonian of the solute k; V_σ is the classical potential produced by the reaction field of polarization charges, σ, induced by the solute on the dielectric interface; E_k is the quantum mechanically calculated energy of the system, and Ψ_k is the wavefunction of the solute.

The process is iterative: one starts with the unperturbed Hamiltonian, evaluates the electric field at the dielectric boundary corresponding to the charge distribution defined by this hamiltonian, the corresponding induced polarization charges, σ, at the boundary, and the potential, V_σ, they produce; in the second iteration V_σ is substituted in eq. (1), and the partially polarized wave functions Ψ'_k are calculated leading to new V_σ; this is used in the next iteration etc., untill the computed characteristics of the system (in particular its total energy, $E_k^{\rm tot}$) converge. The electrostatic free energy of the system, $G_k^{\rm tot}$ is given by [18]

$$G_k^{\rm tot} = E_k^{\rm tot} - \tfrac{1}{2}\int V_\sigma \Gamma_k'' \, {\rm d}v, \tag{2}$$

where V_σ is the converged potential of surface polarization charges, Γ_k'' is the converged (polarized) charge distribution from the solute's converged wavefunction and nuclear charges, and the integration is over the volume of the system. The last term is the energy expended to polarize the medium. The difference:

$$G_k^{\rm el} = G_k^{\rm tot} - E_k^0 = E_k^{\rm tot} - E_k^0 - \tfrac{1}{2}\int V_0 \Gamma_k'' \, {\rm d}v \tag{3}$$

(where E_k^0 denotes the energy of the unperturbed solute *in vacuo*) corresponds to the electrostatic part of the free energy of solvation. The mechanical part PV is negligible under normal conditions [18]. The entropy of hydration is not fully accounted for by Eq. (3) (see refs. [7,41,42]). It can be evaluated separately [7,48,49].

2.2. Calculation of polarization charges

Two numeric methods are used most often for calculations of the electrostatic reaction field of the solvent around polar polyatomic molecules: boundary element method (BEM) [7,38] and finite difference method (FDM) [8,50]. Numerical versions of BEM [7,38] allow to calculate reaction fields as electric fields of surface polarization charges induced by charge distributions of the solute at the dielectric boundary between the solute and the solvent, modeled as a continuum dielectric. The density of the surface polarization charges, σ_i, can be determined from the boundary conditions for the normal components of the electric field and the electric displacement [7, and references therein] as

$$\sigma_i = \frac{1}{4\pi}\,\frac{D_{\rm in} - D_{\rm out}}{D_{\rm in}}\,\boldsymbol{E}_{\rm out} \cdot \boldsymbol{n}_i, \tag{4}$$

where $D_{\rm in} = {\rm I}$ and $D_{\rm out} = 78$ are 'inside' and 'outside' dielectric constants characterizing 'quantum mechanical' and solvent parts of the system; the dielectric interface is divided into *ne* boundary elements; σ_i characterizes the surface density of polarization charges that is approximated by a constant on the surface of ith element ($l \leqslant i \leqslant {\rm ne}$) of the dielectric boundary ; $\boldsymbol{n}_i$ is a normal to the ith boundary element directed to the outside medium at some 'representative' point, $\boldsymbol{r}_i$, of this element, and $\boldsymbol{E}_{\rm out}$ is the electric field at the 'representative' point, $\boldsymbol{r}_i$, on the surface of ith boundary element [7]. $\boldsymbol{E}_{\rm out}$ has contributions from the charge distribution of the solute (su), $\boldsymbol{E}_{su}$, and from polarization charges on all boundary elements. Thus polarization charge density on each boundary element depends on polarization charge densities on all other elements. Substitution of $\boldsymbol{E}_{su}$ and expressions for electric fields of polarization charges from all boundary elements into Eq. (4) leads to the system of linear algebraic equations for σ_j [7]

$$\sum_j^{1,ne} {\rm K}_{ij}\sigma_j = -\boldsymbol{E}_{\rm su} \cdot \boldsymbol{n}_i. \tag{5}$$

The surface densities of polarization charges, σ_i, for each ith boundary element can be found from this system of equations and used to calculate V_σ in Eqs. (1)–(3). For practical computations we contract continuous surface distribution of polarization charges to a few point charges on each boundary element. These point charges are used as input defining V_σ in quantum mechanical calculations.

It is possible to calculate surface polarization charges from FDM solution of the Poisson–Boltzmann equation [8,50] as well,

$$\nabla[D(r)\nabla\varphi(r)] - \kappa^2(r)\varphi(r) + 4\pi\rho(r) = 0, \tag{6}$$

where $D(r)$, $\kappa(r)$, $\rho(r)$ and $\varphi(r)$ are the dielectric constant, inverse Debye–Huckel length, charge distribution, and potential at point r. They are defined on rectangular grids for the practical purposes of numerical calculations. If a grid node with subscript 0 is surrounded by eight other neighboring grid points with subscript i, then for a cubic grid the system of the following algebraic equations defines potentials on the grid nodes [8,50,51]:

$$h\sum_i D_i(\varphi_i-\varphi_0) - h^3\kappa^2\varphi_0 + 4\pi q_0 = 0, \qquad (7)$$

or

$$\varphi_0 = \frac{\sum_i D_i\varphi_i + 4\pi q_0/h}{\sum_i D_i + \kappa^2 h^2}, \qquad (8)$$

where h is the length of the side of the cubic grid cell, and q_0 is the charge assigned to zeroth node (all charges are distributed on nodes according to some scheme [50,51]).

The finite difference expression for surface polarization charge, σ_0, on each grid node *assigned* to the boundary (in reality these nodes are not located exactly on the boundary) is [50]

$$\sigma_0 = D_{\text{out}}\Big(6\varphi_0 - \sum_i \varphi_i\Big)/4\pi h. \qquad (9)$$

These charges can be used to compute V_σ in Eqs. (1)–(3). We use both approaches to compare the results obtained. Note that in BEM all polarization charges are placed on the dielectric interface. In FDM they are computed for positions on grid nodes, and then some of them are shifted to the actual dielectric interface and some are left on the nodes [51]. This may lead to small differences between the results obtained with BEM and FDM techniques.

2.3. Representation of the molecular charge distributions

Different ways have been suggested in the literature for calculations of the electric field, E_{su} (Eq. (5)), of the solute or of its potential. These can be calculated either directly from the solute's wavefunctions [17,18] or from some multipole representation of the molecular charge distributions [7,38,49,52]. Calculations directly from the wavefunctions seem to be natural and rather attractive. However, they can lead to a wrong value of the calculated total surface polarization charge as the dielectric interface may be well inside the solute's electron cloud. It can be shown [38] that the total polarization charge at the dielectric interface, $Q^{\text{tot}}_{\text{pol}}$, is

$$Q^{\text{tot}}_{\text{pol}} = Q_{\text{in}}(D_{\text{in}} - D_{\text{out}})/D_{\text{in}}D_{\text{out}}, \qquad (10)$$

where Q_{in} is the solute charge enclosed by the dielectric boundary, and D_{in} and D_{out} are dielectric constants inside and outside of the boundary. The electric field on any surface, $E(s)$, can be obtained directly from the wavefunctions, and the total charge inside this surface can be calculated as a surface integral from the Gauss theorem [38, and references therein], where $\boldsymbol{n}(s)$ is the outside directed normal to the surface:

$$Q_{\text{in}} = \frac{1}{4\pi}\int_s E(s)\cdot \boldsymbol{n}(s)\,\mathrm{d}(s). \qquad (11)$$

Our SCF calculations at near Hartree–Fock limit (unpublished results) with cavity radii from ref. [38,53] (see Table 1) yield the following charges (in electron units) inside of ionic or molecular cavities: 0.99 (Li +, Na +, K +), 0.64 ($\mathrm{F^-}$), 0.78 ($\mathrm{Cl^-}$), 0.18 $\mathrm{H_2O}$). Thus, for anions and molecules with negative partial charges on atoms (which usually have significantly more diffuse electron clouds than cations) the difference between the charge inside the cavity and the total ionic/

Table 1
Cavity radii for different types of atoms [38,53]

	Atom/group											
	CH [a]	NH [b]	HC [c]	$\mathrm{H_{Hb}}$ [d]	$\mathrm{O_{Hb}}$ [e]	$\mathrm{N_{Hb}}$ [f]	C [g]	F	Cl	Li	Na	K
Cavity radius (Å)	2.78	2.17	1.71	1.16	1.5	1.5	2.46	1.423	1.937	1.316	1.68	2.172

[a] United atoms for CH, $\mathrm{CH_2}$, $\mathrm{CH_3}$, $\mathrm{CH_4}$.
[b] N when it cannot form hydrogen bonds of N–HX type (e.g., in derivatives of ammonium ion).
[c] Non-hydrogen-bonding hydrogen (e.g., aliphatic H).
[d] Hydrogen-bonding H (e.g., in water or ammonia).
[e] Hydrogen-bonding oxygen (e.g., in water or OH).
[f] Hydrogen-bonding nitrogen (e.g., in ammonia and its derivatives).
[g] Used here for all carbons.

molecular charge is rather large. According to Eq. (10) the magnitude of the total surface polarization charge for polar molecules in water is almost equal (within 2%) to the charge inside the cavity. Born-type calculations of hydration enthalpies of ions [53] show that good agreement with experiment can be obtained only if the entire ionic charge is placed inside the cavities with the same radii. Any increase of the radii of anionic cavities (that would enclose more of the electron density of anions) would lead to a decrease of hydration enthalpies (which are proportional to the inverse cavity radius) in disagreement with experiment [38,53]. If the total polarization charge, producing V_σ in Eqs. (1)–(3), corresponds to less than full ionic charge, the resultant total energies calculated with hybrid-quantum methods (Eq. (2)) may be incorrect, with electrostatic contributions (eq. (3)) up to 40% in error (e.g., for F^-). This problem has not been adequately addressed so far.

A practical alternative (see ref. [47] for a discussion) is to use a multipole expansion of the solute charge distribution placed inside the cavity as suggested by the success of the Born type approaches [53]. In this work we approximate multipole expansions by readily available Mulliken [54] charges, corrected to reproduce calculated dipole moments of the molecules under the study. The shortcomings of the Mulliken scheme are well known, and we do not discuss them here. In many of our calculations Mulliken charges reproduce calculated dipole moments rather poorly. The worst disagreements (30–40%) come from calculations with basis sets that include diffuse functions on d orbitals. (These functions are, however, required to obtain accurate vacuum dipole moments).

Therefore, for polar uncharged molecules where the dipole–solvent interactions provide a leading contribution to the hydration energy, we used a simple proportional normalization of the Mulliken charges such that they exactly reproduce calculated dipole moments. For water and ammonia such a procedure is completely adequate. For these molecules the charge on only one atom is unknown, and the charges on other atoms are known multiples/fractions of this single charge due to the electroneutrality. For larger polar uncharged molecules (e.g., CH_3OH and CH_3NH_2) the procedure becomes non-uniquely defined (the number of unknown charges is larger then the number of constraints). Similar procedures have, nevertheless, been widely used [55]. For charged molecules the situation is more complex because the dipole moments (and higher multipoles) depend on the choice of the coordinate system. Adding two charges with opposite signs and a magnitude of the ionic charge to the set of Mulliken charges allows to correctly represent both monopole and dipole. However, this procedure did not significantly change magnitudes of calculated hydration energies suggesting that more sophisticated schemes may be needed.

A number of schemes for accurate and assumption-free calculations of distributed multipole moments have been reported in the literature [56–59]. One of these schemes [57] allows to represent distributed multipoles up to the quadrupole by sets of point charges. This could be, but was not, implemented with other schemes. Representations of multipole distributions by point charges are currently more convenient than those with point multipoles, because most codes for the evaluation of surface polarization charges expect point charges as input. We will report effects of these representations of molecular charge distributions on the results of hybrid quantum calculations elsewhere.

2.4. Classical evaluation of the effects of solute polarization by the solvent

Energetic effects of the solute polarization by the solvent can be also obtained through a number of purely classical schemes [7,60–65]. It is of interest to find out whether purely classical approaches yield results similar to those of a more fundamental 'hybrid' approach in which solute polarization is treated quantum mechanically. In one of the classical schemes polarizability of the solute is accounted for by assigning point polarizabilities to the atoms of the solute, and using classical electrostatics to calculate the effective dipole moment of the solute in the solvent and the energy of solvation [7,61,63–64]. We use this approach here to evaluate its consistency with the hybrid quantum calculations.

Dipoles, $\boldsymbol{\mu}_m$, induced on atoms, characterized by point polarizabilities, α_m, are proportional to the electric field which is a sum of contributions from solute 'vacuum' partial charges, $\boldsymbol{E}_{\text{su}}$, from polarization charges, $\boldsymbol{E}_\sigma$, and from induced dipoles, $\boldsymbol{E}_\mu$

$$\boldsymbol{\mu}_m = \alpha_m(\boldsymbol{E}_{\text{su}} + \boldsymbol{E}_\sigma + \boldsymbol{E}_\mu). \tag{12}$$

$\boldsymbol{E}_\mu$ will also contribute to the right-hand side of Eq. (4). Substituting expressions for electric field from all sources ($\boldsymbol{E}_{\text{su}}$, $\boldsymbol{E}_\sigma$, $\boldsymbol{E}_\mu$) into Eq. (4) and (12) and rearranging terms we arrive at a system of linear algebraic equations for both σ and $\boldsymbol{\mu}$ [7,63],

$$\sum_j^{l,\text{ne}} K_{ij}\sigma_j + \sum_m^{i,\text{na}} \alpha_m \sum_a^{1,3} V_{im}^a \mu_{\text{m}}^a = -\sum_a^{1,3} E_{\text{su}}^a n_i^a, \tag{13}$$

$$a_m \sum_i^{1,\text{ne}} U_{mi}^a \sigma_i + \alpha_m \sum_{j\neq m}^{1,\text{na}} \sum_a^{1,3} W_{mj}^d \mu_{mj}^a \mu_m^a = -\alpha_m E_{\text{su}}^{0a}, \tag{14}$$

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 indices are explicit. More detailed expressions for the coefficients of these equations can be found in refs. [7,63]. σ_i and $\boldsymbol{\mu}_i$ can be found from simultaneous solution of Eqs. (13) and (14).

A number of approximations, differing in interactions that are explicitly involved in the polarization, have been suggested within this scheme. For example, induced dipoles may or may not be allowed to interact with each other [66], or interact only for non-bonded atoms [61], and initial charges on atoms may or may not be allowed to induce dipoles on other atoms. Here we consider only a simple model in which polarization of the solute is induced only by the reaction field of the solvent, atomic polarizabilities are additive, and atomic charges do not interact with induced dipoles [7,63]. (This simple model is similar in spirit with 'polarizable electropole model' of water [60] in which all charge distribution of the water molecule is represented by a single point dipole. This dipole polarizes bulk water, which in its turn induces a point dipole on the water molecule. The original and the induced dipoles are located at the same point and thus do not interact. In our simple model the original and induced dipoles are just distributed over all atoms in the form of atomic charges and additive induced atomic dipoles). In this simple model the total electrostatic contribution to the free energy of hydration, $\Delta G_{\text{el}}^{\text{tot}}$, can be calculated as

$$\Delta G_{\text{el}}^{\text{tot}} = \tfrac{1}{2}\left(\sum_m q_m^{\text{su}} \sum_j \sigma_j S_j / |r_m - r_j|\right). \tag{15}$$

Additive atomic polarizabilities, α_m, are taken from ref. [67]. $\Delta G_{\text{el}}^{\text{tot}}$ calculated classically from Eq. (15) can be compared to G_k^{el} calculated with the hybrid quantum method (Eq. (3)).

2.5. *Density functional theory*

The density functional (DFT) calculations in this study are carried out using the deMon program [68]. deMon uses Gaussian basis sets to expand DFT orbitals, and to fit the density and exchange–correlation potentials. There is an extensive literature published in this field [68–77].

The DFT method is based on the Hohenberg, Kohn and Sham formalism [1,73,74]

$$E[\rho] = T[\rho] + U[\rho] + E_{xc}[\rho], \tag{16}$$

where the total energy is a functional of the electron density, ρ, T is the kinetic energy of the non-interacting electrons, U is the classical Coulomb electrostatic energy and E_{xc} includes all the many body exchange and correlation contributions to the energy. In the local approximation (LSD) [74] E_{xc} can be expressed as follows

$$E_{\text{xc}} = \int_v \rho(r)\epsilon_{\text{xc}}(\rho)\,\mathrm{d}V, \tag{17}$$

where ϵ_{xc} is the exchange–correlation energy per electron in a correlated electron system of uniform density. DeMon uses the form of ϵ_{xc} given by Vosko, Wilk and Nusair (VWN) [75]. To account for non-uniform electron density present in molecular systems the non-local (or gradient) corrections, using the density gradient, can be applied. In this study non-local corrections (NLSD) will be computed using the exchange and correlation functions of Becke and Perdew (BP) [71]. It has been demonstrated [71,72,76] that BP approach leads to accurate bond energies. This, however, does not guarantee that DFT leads to accurate values of other properties important for calculations of hydration energies.

Several types of basis sets have been used to establish the accuracy of deMon calculations [76]. deMon basis sets have been optimized for atomic DFT calculations. The standard orbital DZVD basis set has a notation (621/41/1) and (41) for first row and hydrogen atoms respectively [76] (this notation indicates that e.g., for carbon there are three s-, two p-, and d-type functions; the first s-type function has six primitive Gaussians, the second has two Gaussians, and the third function is not contracted). This basis set corresponds in the number and type of basis functions to the popular Hartree–Fock 6–31G* basis set [77]. DZVP set differs from DZVD basis set by an additional polarization p-type function for hydrogen. The notation of this hydrogen basis set is (41/1), and thus, the DZVP basis set has the same size as Hartree–Fock 6-31G ** basis set [77]. Both, DZVD and DZVP, basis sets are of the double-zeta split valence quality. In addition a triple-zeta, TZVP, basis set was used. The notation for this set is (7111/411/1) [76]. In addition to orbital basis sets, described above, deMon calculations require auxiliary basis sets, A1 and A2 [75], to accurately represent electron density and exchange–correlation potentials. The exchange and correlation terms cannot be evaluated analytically, and should be synthesized on a grid [76,77]. A fine grid with with 832 and 2968 points per atom has been used in this study for the single point and gradient calculations respectively.

Polarization and diffuse orbitals play an important role in molecular calculations [77]. We investigated the effect of diffuse functions by adding diffuse orbitals to s-, p- and d-manifolds for both first row and hydrogen atoms. The following basis sets, with diffuse orbitals, were used: DIF1, DIF2, DIF3 and DZVPD. These sets, together with the standard basis sets are listed below (basis sets for first row atoms are followed by the hydrogen basis sets):

DZVD: (621/41/1), (41)
DZVP: (621/41/1), (41/1)
DZVP2: (721/51/1), (41/1)
TZVP: (7111/411/1), (41/1)
*DIF*3: (621/41/11 +), (41/11 +)
*DIF*1: (6211 + /411 + /1), (411 + /1)
*DIF*2: (6211 + /411 + /11 +), (411 + /11 +)
DZVPD: (621/41/11 +), (41/1).

The DZVPD basis set corresponds in the number and type of basis functions to the popular 6-31G** + Hartree–Fock basis set. The DIF2 is the largest basis set used in the present work.

3. Results and discussion

3.1. In vacuo calculations

The first question we pursued concerns a possibility of accurate prediction of vacuum dipole moments with DFT calculations employing basis sets with diffuse orbitals. Results of our calculations of vacuum dipole moments of water and ammonia with such basis sets are listed in Table 2. It can be seen from Table 2 that vacuum dipole moments of water and ammonia calculated with DIF3 and DIF2 basis sets agree within a few hundredth of debye and are very similar to available experimental values of the corresponding dipole moments. The dipole moment calculated with the intermediate size DIF1 basis set is about 0.2 Debye larger and further from the experimental value. Comparison of DIF basis sets suggests

Table 2
Vacuum dipole moments (D) of water and ammonia calculated with different basis sets

Molecule	Basis set/potential	DIF3/VWN	DIF3/BP	DIF1/VWN	DIF1/BP	DIF2/VWN	DIF2/BP	Exp. [a]
H_2O		1.910	1.873	2.180	2.122	1.940	1.896	1.85
NH_3		1.621	1.606			1.568	1.571	1.47

[a] Experimental dipole moments in gas phase from ref. [78].

that the diffuse d-orbital that is present in DIF3 and DIF2 basis sets but absent in DIF1 basis set is responsible for this difference between the calculated dipole moments. It is gratifying that the agreement between the experimental dipole moment of water and its value calculated here with DIF3 basis (a few hundredth of Debye) is much better than that obtained with traditional SCF methods reported in ref. [21] (≈ 0.2–0.3 D), or in CI MP2 calculations (≈ 0.2 D) [79]. The dipole moment of ammonia calculated with diffuse polarization functions (Table 2) is also in good agreement with its experimental value [78] (the difference is only ≈ 0.1–0.15 D). This suggests that DFT calculations with diffuse polarization functions on d-orbitals can provide reasonable dipole moments for small molecules in the gas phase.

The similarity to each other and to experimental values of vacuum dipole moments calculated with DIF2 and smaller DIF3 basis sets raises a question whether vacuum energies calculated with these basis sets are also similar. Results of our calculations of vacuum energies of water and ammonia with such basis sets are listed in Table 3. It can be seen from data in table 3 that total energies of water and ammonia molecules in vacuum agree within 2.0 kcal/mol in calculations with basis sets of increasing size including diffuse orbitals. The agreement is similar in subsets using local (VWN) and non-local (BP) auxiliary functions. Comparison of data in Tables 2 and 3 show that, while differences between the total energies obtained in local and non-local calculations are rather large as one would expect (hundreds of kcal/mol), corresponding differences in vacuum

Table 3
Vacuum energies [a] of water and ammonia obtained with different basis sets including diffuse functions

Molecule	Basis set/potential	DIF3/VWN	DIF3/BP	DIF1/VWN	DIF1/BP	DIF2/VWN	DIF2/BP
H_2O		−75.8917	−75.4555	−75.8922	−76.4555	−75.8947	−76.4584
NH_3		−56.0942	−56.5790			−56.0967	−56.5813

[a] Energies in hartree (1 hartree = 627.51 kcal/mol)

Table 4
Vacuum energies of molecules obtained with different basis sets (hartree)

	DZVD/A1 = 631G*		DZVP2/A2 = 631G**		TZVP/A2 = 6311G**		DZVPD = 631G** +	
	VWN	BP	VWN	BP	VWN	BP	VWN	BP
H_2O	−75.8803	−76.4437	−75.8986	−76.4623	−75.8990	−76.4638	−75.8912	−76.4550
CH_3OH	−114.8159	−115.7525	−114.8194	−115.7564	−114.8333	−115.7727	−114.8218	−115.7600
NH_3	−56.0859	−56.5701	−56.0919	−56.5764	−56.0997	−56.5850	−56.0940	−56.5789
CH_3NH_2	−95.0207	−95.8763	−95.0250	−95.8909	−95.0366	−95.8946	−95.0268	−95.8845
CH_3NH_3+	−95.3683	−96.2322	−95.3720	−96.2368	−95.3821	−96.2493	−95.3732	−96.2395
FCH_2CO_2-	−325.4525	−327.8560	−325.4500	−327.8537	−325.4966	−327.9057	−325.4577	−327.8641
CH_2CO_2-	−226.7957	−228.5872	−226.7945	−228.5867	−226.8286	−228.6243	−226.8010	−228.5955

Table 5
Vacuum dipole moments of polar molecules optimized with different basis sets

	DZVD/A1 = 631G *		DZVP2/A2 = 631G **		TZVP/A2 = 6311G**		DZVPD = 631G** +		Exp. [a]
	VWN	BP	VWN	BP	VWN	BP	VWN	BP	
H_2O	2.278	2.216	2.160	2.095	2.155	2.096	1.939	1.887	1.86
CH_3OH	1.826	1.814	1.760	1.766	1.814	1.815	1.622	1.626	1.70
NH_3	1.965	1.977	1.844	1.815	1.818	1.807	1.614	1.587	1.47
CH_3NH_2	1.652	1.684	1.514	1.583	1.490	1.556	1.374	1.397	1.31

[a] Experimental dipoles moments (in Debyes) in gas phase are from ref. [78].

dipole moments obtained with DIF3 or DIF2 basis sets are small (a few hundredth of debye). Note also that while the dipole moments of water calculated with DIF3 and DIF1 basis sets differ by 0.25 D, the corresponding energies are practically the same. Thus, accuracy (or an increase in it) in calculations of one observable does not guarantee a similar accuracy (or an increase in it) in another observable.

It is always important to cut the computation time (e.g., by reducing the basis set size) without losing accuracy. We guessed that hydrogen wave functions may not need diffuse d-orbitals to accurately represent their electron density distributions in contrast to heavy atoms with much larger number of electrons. Comparison of the calculated vacuum energies and dipole moments for water and ammonia in Tables 2 and 3 to the corresponding values in Table 4 and 5 supports this guess. It suggests that calculations with DZVP basis set modified by placing diffuse functions on d-orbitals of heavy atoms alone (this builds a new DZVPD basis set) yield practically the same results as DIF3 basis set. Use of DZVPD basis set decreases the computational time compared to that required for computations with DIF1-DIF3 basis sets. Therefore, we used DZVPD basis set in calculations for molecules larger than water and ammonia.

Further insights into effects of the use of different basis sets on calculated energies and dipole moments can be obtained from computations for a larger number of molecules. Analysis of total vacuum energies in Table 4, containing results for four dipolar and tree ionic molecules, shows that calculations with TZVP/BP basis set always yield

Table 6
Hydration energies, solution dipole moments, and final total energies of polar molecules in water calculated with different basis sets

Basis set	DZVD/A1 = 631G*		DZVP2/A2 = 631G**		TZVP/A2 = 6311G**		DZVPD = 631G**$^{+}$	
Potential	VWN	BP	VWN	BP	VWN	BP	VWN	BP
(a) electrostatic contribution to hydration free energies (kcal/mol)								
H_2O	−11.971	−11.191	−11.039	−10.329	−11.045	−10.342	−9.708	−9.136
CH_3OH	−5.909	−5.438	−5.184	−4.747	−5.104	−4.756	−5.619	−5.069
NH_3	−10.239	−9.763	−9.289	−8.721	−9.269	−8.745	−8.202	−7.761
CH_3NH_2	−5.914	−5.963	−5.130	−5.652	−5.058	−5.174	−4.572	−4.855
(b) solution dipole moments (D)								
H_2O	2.783	2.718	2.687	2.616	2.688	2.621	2.555	2.492
CH_3OH	2.240	2.222	2.158	2.161	2.235	2.231	2.090	2.088
NH_3	2.532	2.539	2.388	2.362	2.360	2.359	2.366	2.334
CH_3NH_2	2.078	2.151	1.904	2.049	1.891	2.013	1.837	1.920
(c) increase in dipole moments due to hydration								
H_2O	0.510	0.502	0.527	0.520	0.533	0.525	0.616	0.605
CH_3OH	0.414	0.408	0.398	0.395	0.421	0.416	0.468	0.462
NH_3	0.547	0.562	0.544	0.547	0.542	0.552	0.752	0.747
CH_3NH_2	0.426	0.467	0.390	0.466	0.401	0.457	0.463	0.523
(d) total energies in solution in hartree (do not include non-electrostatic solute–solvent interactions)								
H_2O	−75.9220	−76.4828	−75.9372	−76.4985	−75.9377	−76.5002	−75.9259	−76.4878
CH_3OH	−114.8361	−115.7711	−114.8371	−115.7727	−114.8508	−115.7891	−114.8414	−115.7776
NH_3	−56.1222	−56.6050	−56.1249	−56.6075	−56.1325	−56.6161	−56.1244	−56.6077
CH_3NH_2	−95.0411	−95.0411	−95.0425	−95.9106	−95.0539	−95.9127	−95.0428	−95.9017
(e) change in total energy upon hydration in kcal/mol (does not include non-electrostatic solute–solvent interactions)								
H_2O	−26.168	−24.536	−24.223	−22.717	−24.285	−22.842	−21.775	−20.583
CH_3OH	−12.676	−11.672	−11.107	−10.229	−10.982	−10.291	−12.300	−11.044
NH_3	−22.779	−21.901	−20.708	−19.516	−20.583	−19.516	−19.077	−18.073
CH_3NH_2	−12.802	−13.115	−10.982	−12.362	−10.856	−11.358	−10.040	−10.793

the lowest total vacuum energies (TZVP energies are lower than the next best ones by 2–7 kcal/mol). Total vacuum energies calculated with DZVPD/BP basis set (that includes diffuse orbitals) are in most cases lower than those calculated with double-zeta basis sets without such orbitals. However, for water and methylammonia energies calculated with DZVP2/BP are lower than those calculated with DZVPD/BP (note that DZVPD basis set does not contain some basis functions that are present in DZVP2 basis set, and thus variational principle is not violated). On the other hand data for four dipolar molecules in Table 5 show that vacuum dipole moments for water, ammonia and methylammonia calculated with DZVPD/BP basis set (that includes diffuse orbitals) agree with experimental dipole moments significantly better than those calculated with any other basis set. Vacuum dipole moments calculated for methanol with all basis sets are within 0.13 D from experimental value, while its DZVP2 and DZVPD calculated values deviate from experimental value by ±0.07 D. Further analysis of the results in Table 5 also shows that use of diffuse functions decreases calculated dipole moments by 0.1–0.2 D for all molecules studied, while use of the triple-zeta basis set (TZVP) not necessarily reduces dipole moments towards their experimental values compared with corresponding double zeta (DZVP2) results (e.g., for methanol).

It can also be noted that while for each set the energies calculated with local (VWN) and non-local (BP) potentials (Table 4) differ by hundreds of kcal/mol thus confirming importance of NLSD corrections for molecular energetics [71,77], the dipole moments in corresponding calculations differ very little (less than 0.07 D). This agrees with a recent study [80] suggesting that the electron density is not significantly effected by the NLSD potential.

All this suggests that neither the triple-zeta basis sets nor the double-zeta basis set with diffuse functions can guarantee a convergence for all physical observables of interest, and thus even larger basis sets may be required for this purpose. Electrostatic part of the free energy of hydration, ΔG_k^{el} (Eq. (3)), includes contributions both from the change in the internal energy and from the interaction of the dipole with the reaction field. Therefore, an accuracy (or its lack) in both characteristics can influence results of computations of hydration energies. However, because correct dipole moments are expected to be a major factor in accurate calculation of reaction fields of polar molecules, we decided to stop at DZVPD/BP level calculations in this study and to leave studies with TZVP/BP and larger basis sets for future.

3.2. Solution calculations

3.2.1. Uncharged dipolar molecules

Results of computations of hydration of four uncharged molecules with hybrid approach utilizing different basis sets and auxiliary functions are given in Table 6(a)–6(e). These results can be analyzed similarly to the *in vacuo* results to find if a change in the basis set leads to similar effects for different molecules in water. However, solution calculations for dipolar molecules produce more data of interest than *in vacuo* calculations, because not only dipole moments and total energies in solution but also changes in their values upon vacuum to water transfer could be informative.

We start with calculated electrostatic contributions to the free energies of hydration, ΔG_k^{el} (Eq. (3)), of dipolar molecules in Table 6(a). Analysis of the corresponding results shows that depending on the basis set used their values for individual uncharged polar molecules vary within ≈ 2 kcal/mol in both local (VWN) and non-local (BP) calculations. With exception of methanol, for which the smallest electrostatic contribution to hydration energy is calculated with TZVP basis set, DZVPD calculations lead to the smallest absolute value of the electrostatic part of the free energy of hydration. This would agree with a leading role of the magnitude of the vacuum dipole (which is always the lowest in DZVPD calculations) in determining the magnitude of hydration energies of individual molecules. TZVP results fall out of line for unclear reasons.

Solution dipole moments from the calculations with different basis sets vary within 0.24 D (Table

6(b)). This variation constitutes ≈ 30% of the maximum *increase in the dipole moments due to the solute polarization by the solvent* (i.e. induced dipole moments) shown in Table 6(c). These increases in dipole moments vary between 0.39 and 0.75 D, and are largest in calculations with DZVPD basis set including diffuse functions on d-orbitals (Table 6(c)). Thus, inclusion of diffuse functions leads to lowest values of dipole moments in vacuum (Table 5) and to their largest increase in solution (Table 6(c)). A comparison of hydration energies in Table 6(a) and induced dipoles in Table 6(c) also shows that while inclusion of diffuse functions leads to the largest induced dipoles (i.e., increases in dipole moments), its leads (with exception of methanol) to the smallest hydration energies. Thus the magnitude of hydration energy correlates rather with the magnitude of the vacuum dipole moment than with its increase due to the hydration (note the methanol deviation from this rule).

Calculated total molecular energies of four polar uncharged molecules in solution and their increments upon the hydration are shown in Table 6(d) and 6(e). A comparison of the hydration energies from Table 6(a) to the corresponding increments in the total energy upon hydration from Table 6(e) shows that increases in the total energy due to hydration are always more than twice larger than the electrostatic contributions to the free energies of hydration shown in Table 6(a). In a classical electrostatic theory with linear response, solute–solvent interaction energy is always twice larger than hydration energy [47]. This observation may suggest either that the classical linear response theory is an oversimplification, or that DFT (by itself or within the hybrid approach) does not represent solute polarization accurately.

3.2.2. *Ionic molecules*

There is less meaningful data to analyze for hydration of ions than for hydration of uncharged polar molecules as dipole moments do not have a definite value for ions (see section 2). Calculated energy changes produced by hydration of ions and their total energies in solution are shown in Table 7(a)–7(c). While calculated electrostatic part of the free energy of hydration changes relatively little (< 1.5 kcal/mol) with the basis set used for the positive ion (methyl-ammonium), it changes quite dramatically (up to 5 kcal/mol) in calculations with diffuse functions for two negatively charged ions (Table 7(a)). The same trend is clear in the changes in total energies in solution (Table 7(c)) which, as for uncharged molecules, are more than twice the corresponding hydration energies from Table 7(a). Mulliken charges calculated for negative ions with DZVPD basis set including diffuse functions look rather

Table 7
Electrostatic hydration energies and total energies of charged molecules in water calculated with different basis sets

Basis set	DZVD/A1 = 631G*		DZVP2/A2 = 631G**		TZVP/A2 = 6311G**		DZVPD = 631G**$^+$	
Potential	VWN	BP	VWN	BP	VWN	BP	VWN	BP
(a) electrostatic contribution to hydration energies (kcal/mol)								
$CH_3NH_3^+$	−68.283	−68.447	−66.556	−68.660	−66.659	−68.740	−66.371	−68.371
$FCH_2CO_2^-$	−66.752	−67.031	−66.175	−66.356	−67.620	−67.394	−62.982	−60.330
$CH_3CO_2^-$	−68.837	−68.775	−68.298	−68.077	−69.279	−69.279	−64.146	−63.998
(b) total energies in solution in hartree (do not include non-electrostatic solute–solvent interactions)								
$CH_3NH_3^+$	−95.5871	−96.4514	−95.5850	−96.4566	−95.5955	−96.4695	−95.5857	−96.4585
$FCH_2CO_2^-$	−325.6680	−328.0727	−325.6634	−328.0679	−325.7152	−328.1238	−325.6598	−328.0581
$CH_3CO_2^-$	−227.0181	−228.8102	−227.0153	−228.8070	−227.0550	−228.8493	−227.0069	−228.8011
(c) change in total energy upon hydration in kcal/mol (does not include non-electrostatic solute–solvent interactions)								
$CH_3NH_3^+$	−137.31	−137.55	−133.66	−137.93	−133.91	−138.18	−133.35	−137.55
$FCH_2CO_2^-$	−135.23	−135.99	−133.91	−134.42	−137.18	−136.86	−126.82	−121.74
$CH_3CO_2^-$	−139.88	−139.94	−138.56	−138.24	−142.07	−141.19	−129.21	−129.02

unreasonable (see section 2) indicating that Mulliken analysis may fail for diffuse orbitals in negatively charged ions.

3.2.3. Comparisons with experiment and classical results

A comparison of calculated results with experimental data is possible for either free energies or enthalpies of hydration. Analysis of experimental values of these thermodynamic quantities for polar [81] and ionic [82] molecules shows that hydration entropies for these molecules are rather large and cannot be accounted for by the temperature dependence of the dielectric constant or of the cavity radius [7]. On the other hand, the only entropic contribution to the electrostatic part of hydration free energies in the Born theory is due to the temperature dependence of the dielectric constant of water; a substitution of the experimental value of this temperature dependence into the temperature derivative of the Born energy leads to hydration entropies that are only 1.77% of the total free energy produced by electrostatic calculations [53, and references therein]. Thus, we applied 1.77% correction to our calculated results to make them comparable to experimental hydration enthalpies [38,53]. Such a comparison also requires some account of non-electrostatic part of the enthalpy of hydration. This was done in Table 8 following the method of ref. [40].

Enthalpies of hydration calculated with DFT utilizing DZVP2 and DZVPD basis sets and combined with BEM and FDM for reaction field computations are listed in Table 8 along with the

Table 8
Experimental and calculated total hydration enthalpies ΔH_{hydr} (kcal/mol)

Model [a]	Molecule							
	H_2O	NH_3	CH_3OH	CH_3NH_2	$CH_3NH_3^+$	NO_2^-	$CH_3CO_2^-$	$FCH_2CO_2^-$
DZVP2/BP [b]	−11.3	−9.8	−9.0	−10.3	−74.4	−83.7	−74.7	−73.1
DZVP2/BP [c]	−11.6	−9.8	−9.2	−10.5	−76.2	−82.8	−72.8	−75.2
DZVPD/VWN [b]	−11.0	−9.2	−9.9	−9.2	−72.1			
DZVPD/BP [b]	−10.2	−8.8	−9.4	−9.5	−74.1			
DZVPD/BP [c]	−10.5	−8.8	−9.2	−9.7	−76.0			
class. contin. no solute polarizability [d]	−11.0	–	−10.2	–	−69.5 (−70.4)	−89.7 (−89.5)	−86.3 −73.1)	−84.0 (−72.3)
class. contin. with point polarizabil. [e]	−9.8	−8.5	−9.6	−9.1	−70.6 (−71.2)	−90.6 (−91.1)	−91.0 (−77.2)	−86.1 (−75.1)
non-polar contribution [f]	−0.9	−0.9	−4.2	−4.5	−4.5	−4.2	−5.4	−5.5
Exp. ΔH_{hydr} [g]	−10.0	−7.9	−10.2	−10.3	−67.9	−86.7	−101.1	−91.1

[a] Basis sets or classical models used; total hydration enthalpies include nonpolar contributions; electrostatic part of hydration enthalpies are calculated as electrostatic parts of hydration free energies (Table 6(a) and 7(a) times 1.0177 [38,53].

[b] Calculated with the boundary element based program HydrEn [7,38].

[c] Calculated with the finite difference based program DelPhi [8].

[d] Classical model (HydrEn) with no solute polarizability [7,38] and Mulliken charges from Gaussian 6-31G** calculations [38]: the values in parentheses are similar HydrEn calculations but with Mulliken charges from deMon DZVP2/BP basis set.

[e] Classical model (HydrEn) with point polarizabilities [67] on solute atoms [7]; vacuum charges are from ref. [38] except for FCH_2CO_2- where they were recalculated with GAUSSIAN 92 and 6-31G** basis set.

[f] From ref. [38].

[g] From ref. [7] for ions, ref. [81] for H_2O, NH_3, CH_3OH and CH_3NH_2.

corresponding experimental values. Analysis of data in Table 8 allows to seek answers (at least tentative for important questions: (a) does a use of a particular basis set lead to a better agreement with experiment? (b) Is there a significant difference between performances of the hybrid method utilizing BEM or FDM computations? (c) How purely classical calculations fare compared to the hybrid calculations?

Analysis of data in Table 8 shows that for four uncharged molecules enthalpies of hydration calculated with DZVPD/BP basis set containing diffuse functions agree with experimental data within 1 kcal/mol. For DZVP2/BP basis set (lacking diffuse functions) differences between calculated and experimental values reach 2 kcal/mol. Total hydration enthalpy is just to the sum of its electrostatic part (differing by only 1.77% from the values in Table 6(a) and of the non-polar contribution (tenth line in Table 8), that does not depend on the basis set in our calculations. Analysis of differences between electrostatic parts of hydration energies obtained with different basis sets in Table 6(a) shows that for water and ammonia these differences calculated with DZVD basis set become close to 3 kcal/mol. It can be noted, however, that for a larger molecule of methylammonia the best agreement with experimental value is achieved in calculations with DZVP2 basis set, and for methanol–with DZVD basis set (in the latter case the improvement is only 0.4 kcal/mol) (see Table 6(a)). The results for larger molecules can be due to our procedure for the normalization of dipole moments (see section 2) which is not uniquely defined for larger molecules. It can also be noted that deviations between the calculated and experimental hydration enthalpies do not follow the same pattern: while in DZVPD/BP computations calculated values for water and ammonia are somewhat larger than experimental values, they are smaller for methanol and methylammonia. Thus, while results of DZVPD calculations provide reasonable and consistent degree of agreement with experiment, the patterns of deviations between the calculated and experimental results currently obscure the answer to question a.

A comparison of the values of hydration enthalpies calculated with the boundary element algorithm HydrEn and with finite difference algorithm DelPhi in Table 8 shows that they agree within 0.3 kcal/mol for dipolar molecules (leading to a negative answer to question (b) for these molecules).

Analysis of the values of hydration enthalpies calculated with classical model with atomic polarizabilities starting with experimental dipole moments (Table 8) shows that they are of a quality comparable to that obtained with DZVPD basis set (only the result for methyl ammonia is somewhat further from the experimental value). However, if this classical model would start with dipole moments calculated with DZVPD/BP basis set the result for ammonia would deviate from experiment by almost 2 kcal/mol. It can be also noted that while solution dipole moment of water calculated here (see Table 6(b)) with DZVPD basis set is 2.49 D (within 0.1 D of the experimental value of 2.4 D [7,83]), it is 2.67 D in the classical calculation [7] with experimental polarizabilities assigned to atoms. Thus despite similarities in the hydration energies calculated with the hybrid approach and with classical approach some physical characteristics of the polarized solute can be quite different. Thus the answer to question (c) is somewhat ambiguous. This conclusion leaves room for further comparative studies of physics of hybrid and classical models of hydration.

Results of our calculations for four ions are also listed in Table 8. For positively charged methyl ammonium hydration enthalpies calculated with DZVP2 and DZVPD basis sets are almost identical and deviate from experimental values by 10% (note that in deprotonation reactions for ammonium-like molecules in solution the cavity radius for nitrogen may need to be changed from the non-hydrogen-bonding to the hydrogen-bonding type (Table 1). Results of classical calculations for this molecule deviate from experiment by less then 5%. Hydration enthalpy calculated for nitrogen dioxide is within a few percent of the experimental value. The values yielded by the classical models for this molecule are of a similar quality. It is, however, interesting to note that for nitrogen dioxide a classical calculation with point polarizabilities and vacuum charges from DZVP2 calculation yield hydration energy 7.4 kcal/mol larger than that obtained in the hybrid calculation with DZVP2 basis set. This again may be indicative of some difference in the description of the physics in two approaches. It has been previously noted [7,38] that no classical calculation (including simulations) have produced an agreement with experimental hydration enthalpy of acetate ion (or its derivatives): for acetate deviations were over 10%. These deviations are even larger in our hybrid calculations: 25–30%. Classical calculations with Mulliken charges from DZVP2 calculations reproduce these hybrid results quite well, suggesting that large deviations from experiment are due to DZVP2 Mulliken charges. A comparison of these charges to Mulliken charges from Gaussian calculations (e.g., ref. [38]) shows that Mulliken charges from our DZVP2 calculations are much closer to those from Gaussian STO-3G calculations than from 6–31G** calculations. A reason for this is not currently clear.

Hydration enthalpies of ions calculated with the hybrid method using the boundary element program (HydrEn) and finite element program (DelPhi) agree within 2 kcal/mol. These differences between the results of two methods do not follow a clear pattern: for two ions hybrid-HydrEn energies are lower than hybrid-DelPhi energies, while it is reversed for two other ions. HydrEn version of the hybrid method gives correct ranking of hydration enthalpies for acetate and fluoroacetate while the DelPhi version reverses this rank order. This may be inconsequential because of the overall poor agreement of calculations with experiment for these ions, but it is useful to know that such an inversion can occur. Thus, for methylammonium and nitrogen oxide, for which classical calculations gave reasonable agreement with experiment, the hybrid method also gives comparably reasonable results. On the other hand, for acetate and fluoroacetate, for which classical calculations gave large errors, the hybrid method fails even more. This shows that for ions conclusive answers to our questions are not at hand yet, and that further studies are obviously required. A planned use of better point charge representations of molecular charge distributions may lead to a clarification of many problems found in this study.

Thus, our calculations, suggest that for uncharged polar molecules DFT calculations can provide values of vacuum dipole moments that are within ~ 0.1 D of their experimental values. For the same uncharged polar molecules our hybrid approach produced agreement with experiment close to the limit of ‘chemical accuracy’ (1 kcal/mol). This is very gratifying and should be tested in further calculations for larger number of compounds. Our calculations indicate that results can be sensitive to the calculated value of the dipole moment and to its particular representation by point charges. Thus both the use of diffuse functions in the basis sets and of accurate multipole expansions are likely to be important, and should be tried in future calculations. The use of accurate multipole expansions may be even more critical in calculations for non-spherical ions where our current results are not as good as for uncharged molecules. This work constitutes only first few steps on a possibly long road to a quantitative theory of hydration effects. However, the results obtained here represent a partial success and suggest that further improvement of the hybrid quantum approach described in this work may provide a basis for quantitative studies of chemical and biochemical reactions in solution.

Acknowledgement

We thank R. Fine and D. Salahub for helpful discussions.

References

[1] R.G. Parr and W. Yang, Density functional theory of atoms and molecules (Oxford Univ. Press, Oxford, 1989).

[2] D. Feller and E.R. Davidson, in: Reviews in computational chemistry, eds. K.B. Lipkowitz and D.B. Boyd (VCH, New York, 1991) p. 1.

[3] J.J.P. Stewart, in: Reviews in computational chemistry, eds. K.B. Lipkowitz and D.B. Boyd (VCH, New York, 1991) pp. 45–81.

[4] M.C. Zerner, in: Reviews in computational chemistry,

eds. K.B. Lipkowitz and D.B. Boyd, (VCH, New York, 1991) p. 313.

[5] M.P. Allen and D.J. Tildesley, Computer simulation of liquids (Clarendon Press, Oxford, 1989).

[6] C.L. Brooks III, M. Karplus and B.M. Pettitt, in: Advances in Chemical Physics, Vol. 71. Proteins: a theoretical perspective of dynamics, structure, and thermodynamics (Wiley, New York, 1988).

[7] A.A. Rashin, J. Phys. Chem. 94 (1990) 1725.

[8] K.A. Sharp and B. Honig Ann. Rev. Biophys. Biophys. Chem. 19 (1990) 301.

[9] G. King and A. Warshel, J. Chem. Phys. 91 (1989) 3647.

[10] A. Warshel and J. Åqvist, Ann. Rev. Biophys. Biophys. Chem. 20 (1991) 267.

[11] A. Warshel and S.T. Russell, Quart. Rev. Biophys. 17 (1984) 283.

[12] D. Rinaldi and J.L. Rivail, Theoret. Chim. Acta 32 (1973) 57.

[13] N.M. Karelson, T. Tamm, A.R. Katrizky, S.J. Cato and M.C. Zerner, Tetrahedron Comput. Methodol. 2 (1989) 295.

[14] A.R. Katrizky and N.M. Karelson, J. Am. Chem. Soc. 113 (1991) 1561.

[15] D. Rinaldi, J.L. Rivail and N. Rguini, J. Comput. Chem. 13 (1992) 675.

[16] A.M. Frisch et al., GAUSSIAN 92 (Gaussian, Inc., Pittsburg, PA, 1992).

[17] S. Miertus, E. Scrocco and J. Tomasi, Chem. Phys. 55 (1981) 117.

[18] J. Tomasi, G. Alagona, R. Bonaccorsi and C. Ghio, in: Modelling of structures and properties of molecules, ed. Z. Maksic, (Ellis Horwood, Chichester, 1987) p. 330.

[19] R. Bonaccorsi, R. Cammi and J. Tomasi, J. Comput. Chem. 12 (1991) 301.

[20] B.I. Tunon, E. Silla and J. Tomasi, J. Phys. Chem. 96 (1992) 9043.

[21] J.A. Grant, R.L. Williams and H.A. Scheraga, Biopolymers 30 (1990) 929.

[22] M.E. Colvin and C.F. Melius, Sandia Report, SAND93-8239, UC-411 Natl. Tech. Inf. Serv., US Dept. Commerce, Springfield, VA 22161 (1993).

[23] B. Wang and G.P. Ford, J. Chem. Phys. 97 (1992) 4162.

[24] G.P. Ford and B. Wang, J. Am. Chem. Soc. 114 (1992) 10563.

[25] V. Luzhkov and A. Warshel, J. Comput. Chem. 13 (1992) 199.

[26] D.T. Fox, N. Rosch and R.J. Zauhar, J. Comput. Chem. 14 (1992) 253.

[27] C.J. Cramer and D.G. Truhlar, J. Am. Chem. Soc. 113 (1991) 8305.

[28] C.J. Cramer and D.G. Truhlar, J. Am. Chem. Soc. 113 (1991) 8552.

[29] J. Gao and X. Xia, Science 258 (1992) 631.

[30] M.L.J. Drummond, Progr. Biophys. Mol. Biol. 47 (1986) 1.

[31] J.A.C. Rullmann, M.N. Bellido and P.Th. van Duijnen, J. Mol. Biol. 206 (1989) 101.

[32] J. Bajorath, D.H. Kitson, G. Fitzgerald, J. Andzelm, J. Kraut and A.T. Hagler, Proteins 9 (1991) 217.

[33] J. Bajorath, Z. Li, G. Fitzgerald, D.H. Kitson, M. Farnum, R.M. Fine, J. Kraut and A.T. Hagler, Proteins 11 (1991) 263.

[34] A. Alex and T. Clark, J. Comput. Chem. 13 (1992) 704.

[35] J.-Y. Liang and W.N. Lipscomb, Intern. J. Quantum. Chem. 36 (1989) 299.

[36] J.-Y. Liang and W.N. Lipscomb, Biochemistry 28 (1989) 9724.

[37] A. Warshel, Computer modeling of chemical reactions in enzymes and solutions (Wiley, New York, 1991).

[38] A.A. Rashin and K. Namboodiri, J. Phys. Chem. 91 (1987) 6003.

[39] E.R. Davidson, in: Reviews in computational chemistry, eds. K.B. Lipkowitz and D.B. Boyd (VCH, New York, 1991) p. 373.

[40] T.J. You and S.C. Harvey, J. Comput. Chem. 14 (1993) 484.

[41] A.A. Rashin and M.A. Bukatin, J. Phys. Chem. 95 (1991) 2942.

[42] A.A. Rashin and M.A. Bukatin, J. Phys. Chem. 97 (1993) 1974.

[43] D. Bashford, Curr. Opinions Struct. Biol. 1 (1991) 175.

[44] M.E. Davis and J.A. McCammon, Chem. Rev. 90 (1990) 509.

[45] S.C. Harvey, Proteins 5 (1989) 78.

[46] A.A. Rashin, J.R. Rabinowitz and J.A. Banfelder, J. Am. Chem. Soc. 112 (1990) 4133.

[47] A.A. Rashin, Progr. Biophys. Mol. Biol., 60 (1993) 73.

[48] A.A. Rashin and M.A. Bukatin, J. Phys. Chem. 98 (1994) 386.

[49] A.A. Rashin, L. Young and I. Topol, Biophys. Chem. this issue (1994).

[50] K.A. Sharp, J. Comput, Chem. 12 (1991) 454.

[51] DelPhi user guide, version 2.4 (Biosym Technologies, 1993).

[52] A.A. Rashin, J.R. Rabinowitz and J.R. Banfelder, J. Am. Chem. Soc. 112 (1990) 4133.

[53] A.A. Rashin and B. Honig, J. Phys. Chem. 89 (1985) 5588.

[54] A. Hinchliffe, Computational quantum chemistry (Wiley, New York, 1988).

[55] U.C. Singh and P.A. Kollman, J. Comput. Chem. 5 (1984) 129.

[56] A.J. Stone and S.L. Price, J. Phys. Chem. 92 (1988) 3325.

[57] W.A. Sokalski, A.H. Lowrey, S. Roszak, V. Lewchenko, H. Blaisdell, P.C. Hariharan and J.J. Kaufman, J. Comput. Chem., 7 (1986) 693.

[58] J.P. Rabinowitz and S.B. Little, Intern. J. Quantum. Chem. Quantum. Biol. Symp. 22 (1988) 721.

[59] G.G. Hall, J. Chem Soc. Faraday Trans. II 85 (1989) 251.

[60] P. Barnes, J.L. Finney, J.D. Nicholas and J.E. Quinn, Nature 282 (1979) 459.

[61] A. Warshel and S.T. Russel, Quart. Rev. Biophys. 17 (1984) 283.

[62] M. Sprik and M.L. Klein, J. Chem. Phys. 89 (1988) 7556.

[63] A.A. Rashin, Intern. J. Quantum Chem. Quantum. Biol. Symp. 15 (1988) 103.
[64] M.L.J. Drummond, J. Chem. Phys. 88 (1988) 5014.
[65] K. Sharp, A. Jean-Charles and B. Honig, J. Phys. Chem. 96 (1992) 3822.
[66] J. Applequist, J.R. Carl and K.-K. Fung, J. Am. Chem. Soc. 94 (1972) 2952.
[67] Young Kee Kang and Mu Shik Jhon, Theoret. Chim. Acta 61 (1982) 41.
[68] D.R. Salahub, in: Ab initio methods in quantum chemistry II, ed. K.P. Lawley (Wiley, New York, 1987).
[69] H. Sambe and R.H. Felton, J. Chem. Phys. 62 (1979) 1122.
[70] B.I. Dunlap, J.W.D. Connolly and J.R. Sabin, J. Chem. Phys. 71 (1979) 3396.
[71] A.D. Becke, in: ACS Symp. No. 394. The challenge of d and f electrons, eds. D.R. Salahub and M.C. Zerner (ACS, Washington, 1989).
[72] D.R. Salahub, R. Fournier, P. Mlynarski, I. Papai, A. St-Amant and J. Ushio, in: Density functional theory in chemistry, eds. J. Labanowski and J. Andzelm (Springer, Berlin, 1991) p. 77; J. Andzelm in: Density functional theory in chemistry, eds. J. Labanowski and J. Andzelm (Springer, Berlin, 1991). p. 151.
[73] P. Hohenberg and W. Kohn, Phys. Rev. B 136 (1964) 864.
[74] W. Kohn and L.J. Sham, Phys. Rev. A 140 (1964) 1133.
[75] S.H. Vosko, L. Wilk and M. Nusair, Can. J. Phys. 58 (1980) 1200.
[76] N. Godbout, D.R. Salahub, J. Andzelm and E. Wimmer, Can. J. Chem. 70 (1992) 560.
[77] J. Andzelm and E. Wimmer, J. Chem. Phys. 96 (1992) 1280.
[78] R.C. Weast, ed., CRC Handbook of Chemistry and Physics, 71 Ed., (CRC Press, Boca Raton, 1991).
[79] M.J. Frisch, J.E. Del Bene, J.S. Binkley and H.F. Schaefer III, J. Chem. Phys. 84 (1986) 2279.
[80] L. Fan and T. Ziegler, J. Chem. Phys. 94 (1991) 6057.
[81] A. Ben-Naim and Y. Marcus, J. Chem. Phys. 81 (1984) 2016.
[82] Y. Marcus, Ion solvation (Wiley, New York, 1985).
[83] D. Eisenberg and W. Kauzmann, The structure and properties of water (Clarendon Press, Oxford, 1969).

Discussion to the paper by A.A. Rashin, M. Bukatin, J. Andzelm and A.T. Hagler

Comments

By J. Gao

In the self-consistent reaction field (SCRF) approach, there are two philosophically different procedures to incorporate the energy penalty needed to polarize the medium. One is to introduce this effect self-consistently into the Schrödinger equation (Eq. (1)), while the other is to treat this term as a correction to the total "quantum mechanical" energy, as is done here in this paper (Eq. (2)). Could you provide some further analysis of the difference and similarity between the two procedures? Do they yield significantly different results if the same solute radius is used?

Responses by A. Rashin to Comments

To J. Gao

My response would likely have a relatively low content as I have not practiced alternative implementations of the self-consistent reaction field (SCRF) approach. A still very instructive older review of the subject is given by Drummond [1] and the latest one by Cramer and Truhlar [2]. I will not be able to match their detailed analysis here. One thing that looks attractive in the alternative approaches is an automatic inclusion of not only the electrostatic but also of the dispersion interactions with the medium. We may try at some point to incorporate something similar into our method. However, Drummond's review indicates that absolute values of dispersion contributions obtained within SCRF are usually overestimated a few times. To mix two unresolved problems together does not seem to be a very promising strategy now. The electrostatic influence is apparently larger, and we would like to understand it first using clear empirical approximations for other contributions.

I also suspect that either I do not clearly understand your question, or you have formulated it in a somewhat self-contradictory way. The Schrödinger equation is solved *only for the quantum part* of the entire system. This part has its own energy when isolated, and when perturbed. Energy penalty for polarizing the medium in our case is not a quantum problem by its nature, because its due mainly to the orientational an not the electronic polarization. This is inherently a classical statistical process. While the deformation of the quantum part in the medium is computed as a difference in solutions of the unperturbed and perturbed Schrodinger equation, what you describe as 'a correction' is just the energetic cost of the deformation/perturbation of the classical medium.

[1] M.L.J. Drummond, Progr. Biophys. Mol. Biol. 47 (1986) 1.
[2] C.J. Cramer and D.G. Truhlar, Rev. Comput. Chem. 6 (1994), in press.