Application of the Linearized MD Approach for Computing Equilibrium Solvation Free Energies of Charged and Dipolar Solutes in Polar Solvents

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · NOVEMBER 2002

Impact Factor: 3.3 · DOI: 10.1021/jp021396z

Application of the Linearized MD Approach for Computing Equilibrium Solvation Free Energies of Charged and Dipolar Solutes in Polar Solvents

M. V. Vener,[†] I. V. Leontyev,[†] Yu. A. Dyakov,[†] M. V. Basilevsky,*[,†] and M. D. Newton*[,‡]

Karpov Institute of Physical Chemistry, ul. Vorontsovo Pole 10, Moscow 105064, Russia, and Department of Chemistry, P.O. Box 5000, Brookhaven National Laboratory, Upton, New York 11973

Received: June 12, 2002; In Final Form: September 30, 2002

The linearized MD technique is developed in order to treat systematically free energies of simple charged and dipolar solutes in water. The solvent electrostatic response field in the solute region is modeled by averaging instantaneous fields found in a MD computation for solvent configurations confined within a cavity that conforms to the real shape of the solute particle. At this stage, all electrostatic interactions are explicitly treated inside the cavity. The solvent in the external region (outside the cavity) is modeled in terms of a standard continuum theory. For nonspherical cavities, the present approach is more accurate than the field computation employed at the preceding MD stage, where spherically truncated Coulomb potentials are modified by the reaction field corrections. We considered two different linearization schemes based on a computation of either the average response field or of its fluctuations. Only the first algorithm proved to be successful. For a series of single-charged monatomic cations and anions, it provides free energies that deviate by few percent from those found in full MD computations. The results are stable relative to a separation of the whole space occupied by the solvent into explicit solvent region (inside the cavity) and the continuum region (outside the cavity). The two-site dipolar dumbbell system was also studied in the range of intersite separation D within 2 Å $< D <$ 10 Å. At the stage of the field computation, three different types of its solvation shell were considered: spherical and bispherical cavities and periodic solvent environment monitored in terms of Ewald method. Free solvation energies are the same (within 1 kcal/mol) for all three models. A smooth dependence of the mean field potential is observed as a function of separation D but its asymptotic value differs by 4 kcal/mol from the free energy computed for the isolated ion pair. The results generally agree with those obtained in the literature in terms of full MD simulations.

1. Introduction

Theoretical studies of polar solvation effects at a molecular level attract now a considerable attention owing to numerous applications in many fields of chemistry and biochemistry. A direct implementation of molecular dynamics (MD) simulations for this purpose seems natural but consumes a significant amount of computational resources, especially if one keeps in mind as a perspective a systematic studies of complex chemically or biologically interesting objects. Therefore, attempts to simplify standard MD technologies without reducing seriously their predictive ability seem to be of significant importance.

Our main concern will be molecular level treatment of charged or strongly polarized solute particles in polar solvents. Electrostatic interactions dominate in such systems. The experience of earlier phenomenological theories as well as numerous simulations suggests the idea that electrostatic solvation effects can be well understood within the linear response approximation.[1] The linearized MD approach has been formulated[2,3] based on the Gaussian fluctuation model and proved to give rather satisfactory results in several test computations.[4−7] This method bridges between straightforward molecular level simulations and implicit continuum solvent models which, with relative success, are widely used in applications.[8,9] The objective of the present work is to develop the linearization technique systematically in order to use it in future as a regular computational tool.

It is generally clear that the Gaussian fluctuation model is closely related to the standard linear response approximation of statistical physics.[10] We show below that the reformulation of the pertaining theory as a linearized MD approach reproduces basic equations of the Gaussian model[2,3,6,7] and, additionally, brings some new insight on their significance and range of validity. In particular, a perspective of treating nonequilibrium solvation phenomena is opened in a full analogy with nonequilibrium extensions of the linear continuum theories.[11]

The essence of the linear response method is the underlying relation $\Phi = \hat{K}\rho$ where $\rho(r)$ and $\Phi(r)$ represent, respectively, the solute charge density ("the force") and the medium response potential field. The both variables are considered as continuous functions of space points r (although a discrete representation of ρ is always available). The linear integral operator $\hat{K}$ performs a connection between fields ρ and Φ under equilibrium conditions. The corresponding formulation of the continuum medium theory is well-known,[11−15] it is borrowed here to treat the same subject at a microscopic level.

The purpose of the present paper is to prepare a methodological background for the following advanced treatment. We deal here with a totally nonpolarizable solvent model and mainly address to a computation of equilibrium solvation energies. The test computations considered below were performed with ions in water solution and a model dipolar two-site system with strongly separated charged centers. The electrostatic contribution comprises about 90% of total solvation free energy changes in

[†] Karpov Institute of Physical Chemistry.
[‡] Brookhaven National Laboratory.

10.1021/jp021396z CCC: $22.00
Published on Web 11/19/2002

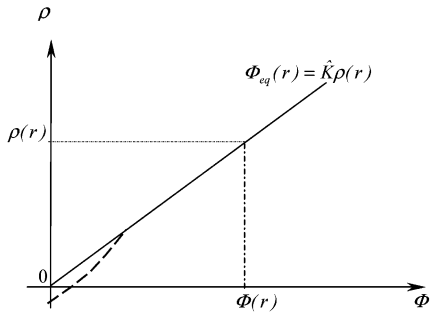

Figure 1. Free energy as a functional of field variables $\Phi(r)$ and $\rho(r)$. This scheme illustrates the full coordinate space and covers both equilibrium and nonequilibrium cases. The equilibrium condition (eq 6) holds along the straight line in a totally linear model. The dashed line shows the actual small change of the equilibrium integration path [5,16] which can be described by eqs 16−18. When the response kernel depends on ρ, $\Delta F(\Phi = 0) = C(\rho)$ along the ordinate axis; see eq 16.

such systems. Therefore, they are quite suitable for studying the effects addressed in the present work.

The possibility of a linearization of the MD treatment of polar solvation free energies was thoroughly studied by Aqvist and Hansen.[16] They revealed a nontrivial nonlinear effect due to a formation of solvent structures around the solute at weak values of its electrostatic field. The structures are destroyed by stronger fields and then the linear relation between the solute charge and the solvent response is regenerated with a good accuracy. This phenomenon is typical for ionic systems[5] whereas for dipolar solutes the situation seems to be ambiguous. Computations in the present study were performed with the similar purpose but with different techniques and objects. The principal results agree with the earlier conclusions.[5,16]

2. The Outline of the Linear Response Approach

We represent the solvation energies in terms of a free energy functional (FEF), denoted as ΔF, defined as $\Delta F(\Phi,\rho) \equiv F(\Phi,\rho) - F(0,0)$, i.e., the charging free energy relative to $F(0,0)$, where the functional variables are a variable medium response field $\Phi(r)$, a collective "solvent coordinate", and the charge density $\rho(r)$, the "solute coordinate":

$$\Delta F(\Phi,\rho) = \int \rho(r)\Phi(r)\, \mathrm{d}^3r + \frac{1}{2}\int L(r,r')\Phi(r)\Phi(r')\, \mathrm{d}^3r\, \mathrm{d}^3r' \tag{1}$$

Equation 1 covers all combinations of variables ρ and Φ, both equilibrium and nonequilibrium. The solvation free energy contains also a nonelectrostatic component which is included in the definition of the free energy of the reference state (0,0). As needed for clarity below, the electrostatic ΔF given by eq 1 will be denoted as ΔF_{el} and the nonelectrostatic contribution denoted as ΔF_{nel}.

The expression for ΔF in eq 1 corresponds to the linear response situation, where the symmetric response kernel

$$L(r,r') = L(r',r) \tag{2}$$

does not depend on ρ. In the more general situation, ΔF would include an additional term (the quantity $C(\rho)$ introduced below), constant with respect to Φ, but dependent on ρ. The two contributions to ΔF (eq 1) are, respectively, the interaction (linear) and self-energy (quadratic) functionals.

The representation of the FEF according to eq 1 is illustrated schematically in Figure 1. For fixed ρ, eq 1 defines a manifold of solvent configurations, generally nonequilibrium. The equilibrium configuration satisfies the stationary condition

$$\delta\Delta F(r) = [\rho(r) + \int L(r,r')\Phi(r')\, \mathrm{d}^3r']\delta\Phi = 0 \tag{3}$$

where $\delta\Delta F$ is the variation of the FEF, a function of r. Its inversion results in the basic linear response equation for the equilibrium field:

$$\Phi(r) = \int K(r,r')\rho(r')\, \mathrm{d}^3r' \tag{4}$$

The kernels $K(r,r')$ and $L(r,r')$ are related by the integral equation:

$$\int K(r,r'')L(r'',r')\, \mathrm{d}^3r'' = -\delta(r - r') \tag{5}$$

The contracted representation of eqs 4 and 5 is derived in terms of relevant integral operators $\hat{K}$ and $\hat{L}$ with symmetric kernels $K(r,r')$ and $L(r,r')$:

$$\Phi = \hat{K}\rho$$
$$\hat{K}\hat{L} = \hat{L}\hat{K} = -I \tag{6}$$

Here I represents the unit operator.

It is expedient to introduce special notations for the interaction and self-energy components of the FEF:

$$W(\Phi,\rho) = \langle\Phi|\rho\rangle = \int \rho(r)\Phi(r)\, \mathrm{d}^3r$$
$$S(\Phi,\rho) = \frac{1}{2}\langle\Phi|\hat{L}|\Phi\rangle = \frac{1}{2}\int L(r,r')\Phi(r)\Phi(r')\, \mathrm{d}^3r\, \mathrm{d}^3r' \tag{7}$$

The well-known important relation between these functionals which is valid in the equilibrium case reads:

$$\frac{1}{2}W_{\mathrm{eq}}(\rho) = -S_{\mathrm{eq}}(\rho) = \frac{1}{2}\langle\rho|\hat{K}|\rho\rangle \tag{8}$$

It follows immediately from eqs 6 and 7. Using this result it is easy to find the quadratic dependence

$$\Delta F(\Phi,\rho) = \Delta F_{\mathrm{eq}}(\rho) + \frac{1}{2}\langle\Phi - \Phi_{\mathrm{eq}}|\hat{L}|\Phi - \Phi_{\mathrm{eq}}\rangle \tag{9}$$

where

$$\Delta F_{\mathrm{eq}}(\Phi,\rho) = \frac{1}{2}\langle\rho|\hat{K}|\rho\rangle \tag{10}$$

We have now to study the dependence of the FEF on the solute coordinate $\rho(r)$. It can be derived from a standard definition of the equilibrium free energy variation:[17] $\delta\Delta F = \int \mathrm{d}^3r\Phi(\rho)\delta\rho$ where $\Phi(\rho)$ and $\delta\rho$ are functions of r. In the form of a functional integral (denoted as $\int$ (...) $\delta\rho$) over a set of equilibrium configurations $\Phi(\rho)$, determined by eqs 4−6, one obtains

$$\Delta F = \int \mathrm{d}^3r \int_{\rho_0}^{\rho} \Phi(\rho)\delta\rho \tag{11}$$

where ρ_0 is some reference density. The functional (Feynman) integral is a superposition of ordinary path integrals between ρ_0 and ρ. Because the equilibrium free energy is a function of state,[17] its change ΔF between two points (which are functions of r) is independent of the path in r-space connecting the points. Therefore, the contribution of each path is the same. This reduces eq 11 to an ordinary integral and gives rise to the thermodynamic integration procedure.

In eq 1, ΔF is defined relative to the reference ($F(0,0)$) where $\rho_0 = 0$. Using this choice, and suppressing any ρ dependence

of the kernel $K(r,r')$ in eq 4, we derive from eq 11 the array of equilibrium free energies (see Figure 1):

$$\Delta F_{eq}(\rho) = \frac{1}{2}\langle\rho|\hat{K}|\rho\rangle = \frac{1}{2}\int d^3r \int \rho(r)K(r,r')\rho(r')\, d^3r' \quad (12)$$

As a concluding step of the reasoning, we must reveal the connection of the fully phenomenological treatment given above to a microscopic background as suggested by the statistical theory. This is supplied by the Kirkwood expression[10,18,19] for the susceptibility kernel $K(r,r')$:

$$K(r,r') = -\frac{1}{k_B T}\langle\Delta\Phi(r)\Delta\Phi(r')\rangle_{\rho,T} \quad (13)$$

Here $\langle ... \rangle_{\rho,T}$ means the thermal ensemble average with given ρ and temperature T and $\Delta\Phi$ representing a fluctuation of the microscopic field $\tilde{\Phi}(r)$:

$$\Delta\Phi(r) = \tilde{\Phi}(r) - \bar{\Phi}(r) \quad (14)$$

The average field $\bar{\Phi}(r)$ is identified with the equilibrium value in eqs 4−6:

$$\bar{\Phi}(r) = \langle\tilde{\Phi}(r)\rangle_{\rho,T} = \Phi_{eq}(r|\rho) \quad (15)$$

The results, (12) and (13), become now fully equivalent to the alternative treatment based on Gaussian approximation for fluctuations $\Delta\Phi(r)$.[2,6] The more general free energy expression of ref 2 is readily recovered in terms of the present reasoning when one considers an alternative choice for $\rho_0(r)$, instead of the present case $\rho_0(r) = 0$, for the reference state of the solute in the functional integral (11). This analysis is limited to the case where $K(r,r')$ is ρ-independent.

To circumvent this limitation, one has to add a Φ-independent term, $C(\rho)$, to eq 1:

$$\Delta F(\Phi,\rho) = C(\rho) + \int \rho(r)\Phi(\rho)\, d^3r + \frac{1}{2}\int L(r,r')\Phi(r)\Phi(r')\, d^3r\, d^3r' \quad (16)$$

Because $C(\rho)$ does not depend on Φ and vanishes in the variational procedure, expression 16 does not change the results in eqs 3−9. Equation 10 changes as

$$\Delta F_{eq}(\rho) = C(\rho) + \frac{1}{2}\langle\rho|\hat{K}(\rho)|\rho\rangle \quad (17)$$

Alternatively, we can find ΔF_{eq} using the general expression 11. With $\rho_0 = 0$, it estimates $C(r)$ as

$$C(\rho) = \int_0^\rho \langle\delta\rho|\hat{K}(\rho)|\rho\rangle - \frac{1}{2}\langle\rho|\hat{K}(\rho)|\rho\rangle \quad (18)$$

This is done in such a way the nonlinearities arising in weak fields $\Phi(r)$ in water solvent[5,16] can be described in terms of the quadratic relative to Φ FEF (eq 16). For instance, the different slopes of the ρ-dependencies $\Phi_{eq}(r_i|\rho)$ at the solute charge sites r_i for anions and cations with the same geometry, LJ parameters, and absolute values of partial charges[5] can be described as a change of the ρ-dependent response kernel $K(r,r')$ in the vicinity of point $\rho = 0$. Another effect of this sort is a small nonzero value $\Phi_{eq}(r|\rho) = \hat{K}(r)\rho(r) \neq 0$ revealed when $\rho \to 0$. It is easy to see that $K(r,r')$ must have a pole in this case although $\Phi_{eq}(r)$ is regular around $\rho = 0$. Expression 18 has no singularities and, with a given $\hat{K}(\rho)$, can be obtained by thermodynamic integration.

3. The Practical Implementation

We discuss now a computational procedure for equilibrium solvation free energies (the so-called charging energies) based on generating microscopic fields $\tilde{\Phi}(r)$ and their fluctuations $\Delta\Phi(r)$ in terms of a MD simulation. For doing this, starting from eq 10, alternative algorithms can be devised. The first algorithm (called algorithm I below) focuses on the average field $\bar{\Phi}(r)$ and represents the solvation energy as

$$\Delta F_{eq}(\rho) = \frac{1}{2}\langle\bar{\Phi}|\rho\rangle \quad (19)$$

This result follows from eqs 7 and 8. Using eq 19 requires the computation of a single MD trajectory (for an ensemble with $\rho = \rho(r)$). A single trajectory is also required when we use eq 10; here $K(r,r')$ is computed for an ensemble with an arbitrary choice of ρ between $\rho = 0$ and $\rho = \rho(r)$. Note that the linear model (eq 4) requires that $\bar{\Phi}(r)$ is negligible for $\rho = 0$. This condition is slightly violated in a real MD computation. Illustrations are given and discussed below.

Equation 10 suggests an alternative representation of ΔF_{eq}, which shows that the response kernel remains the key quantity of a computation: it is concealed in the notation $\bar{\Phi} = \hat{K}\rho$. This approach uses an algorithm (called "algorithm II" below) based on eq 10 and on the susceptibility kernel $K(r,r')$, obtained as a correlation function from MD simulation according to eq 13. In computational practice, these two algorithms are quite different. By comparing their results, one can judge on the validity of the linear approach in general as well as on the quality of the underlying computational techniques applied in performing the averages in eqs 13 and 15.

The choice of $\rho(r)$ in these simulations should be especially emphasized. When computing $\bar{\Phi}(r)$ in terms of eq 15, $\rho(r)$ is taken as the real solute charge density. The same is true in free energy computations according to eq 19 (algorithm I) or eq 10. On the other hand, in the implementation of eq 13 (algorithm II), the value of ρ is not completely specified. According to the fully linear approach, the kernel has no ρ dependence. The choice $\rho = 0$ as well as $\rho = \rho(r)$ (the solute charge density) and also intermediate choices are equally well rationalized. This ambiguity can be resolved by going beyond the linear model and assuming some functional ρ-dependence for $K(r,r')$ as suggested in eqs 16−18.

4. The Mixed Discrete/Continuum Model for Treating Coulomb Interactions

When computing the electrostatic component of the solvation free energy one encounters the problem of treating long-range Coulomb forces in a MD simulation, a theme which is being extensively discussed in the recent literature[5,6,20,21] and seems still to be unresolved. To minimize the associated inaccuracy, we separate the computation into two steps. During the first step, the MD trajectory is calculated by the standard techniques (we applied the GROMOS96 package[22]) invoking conventional reaction field prescriptions for cutting off the Coulomb interaction potentials. The second step focuses entirely on computations of microscopic fields $\tilde{\Phi}(r)$, their fluctuations $\Delta\Phi(r)$, and averaging over the ensemble obtained in the first step. Here the Coulomb potential must be treated as accurately as possible. Therefore, instead of invoking conventional truncation

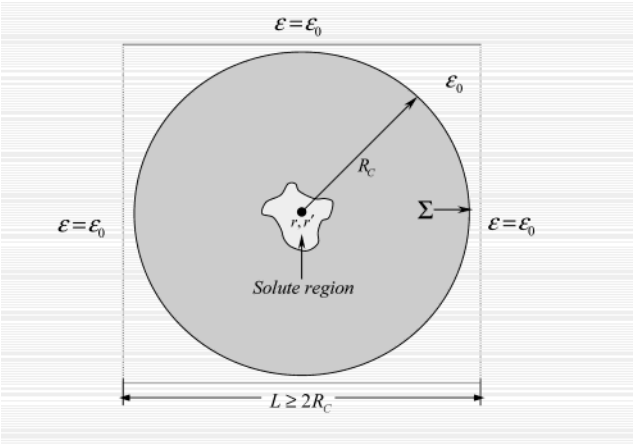

Figure 2. Computation of the microscopic field $\tilde{\Phi}(r)$. The ensemble of solvent particles, surrounding a solute molecule, is simulated within a cubic MD cell with edge L. Those particles contained within a sphere of radius R_c are treated explicitly, their instantaneous response field is given by eq 17. The solvent response outside boundary surface Σ is treated in terms of a continuum model with static dielectric permittivity $\epsilon = \epsilon_0$. In the internal region (inside Σ) $\epsilon = 1$. The case of a cubic cell and a spherical boundary surface is given here as an illustration. A general case with a rectangular cell and nonspherical boundary is considered in section 7.

schemes,[5,6,21−24] we worked in terms of the model which is illustrated by Figure 2. The sphere S of radius R_c is immersed in the cubic MD cell with edge $L = 2R_c$. The solute particle is fixed at the center of the sphere. The surface S divides the whole space into an external part ($|r| > R_c$), which is treated as a continuum with dielectric constant ϵ_0, and the internal part ($|r| < R_c$) where solvent particles are considered explicitly. The microscopic solvent response field is calculated as

$$\tilde{\Phi}(r) = \sum_\nu q_\nu \left[\frac{1}{|r - r_\nu|} + \Psi_\nu(r, r_\nu) \right], \quad \text{where } |r|, |r_\nu| < R_c \tag{20}$$

Here q_ν are partial charges of solvent particles, such that

$$\sum_\nu q_\nu = 0 \tag{21}$$

They are located at points r_ν inside the sphere Σ. These points are determined by an instantaneous solvent configuration generated in terms of the MD trajectory. Point r, at which the total microscopic response is calculated, is located in a relatively small solute region in the vicinity of the center of the sphere.

Note that invoking the spherical boundary between the explicit solvent region and the continuum is not at all obligatory. This is, however, convenient when solutes are monatomic ions (as in the major part of computations performed in the present article) or when solute charge distributions are not strongly anisotropic. For the case of strongly nonspheric solutes we will change accordingly the shape of the boundary surface. In the present paper we follow this general strategy in section 7.

The first term of eq 20 is the usual Coulomb potential, the second one (Ψ_ν) represents the continuum response field due to the unit charge located at r_ν. This is the polarization potential promoted by the interaction of discrete solvent charges with the outer continuum region. Because the boundary Σ is spherical, the standard Kirkwood approach (in terms of spherical functions[25]) is available for computing Ψ_ν. For the general nonspherical boundary Ψ_ν is computed by solving Poisson equation in terms of the PCM method.[8] The two components of eq 20 will be termed below as the "direct potential" (the first, Coulomb term) and the "interference potential" (the second, continuum polarization term).

The average field $\bar{\Phi}(r)$ is now computed according to eq 15. This gives rise to the first algorithm for the free energy computation (see section 3). When $\rho \neq 0$, we have to add to the averaged solvent field the continuum response field due to the solute charge density: the second term in eq 20, yielding

$$\bar{\Phi}(r) = \langle \tilde{\Phi}(r) \rangle_{T,\rho} + \int \rho(r') \chi(r,r')\, \mathrm{d}^3 r' \tag{22}$$

Here $\chi(r,r')$ is Green's function of the continuum theory applied to the external region. The corresponding field component can be readily computed by standard methods of the continuum theory. For nonspherical solutes either Kirkwood[25] or PCM[8] approaches are appropriate. We retain the same notation as in eq 15 for this continuum corrected field.

The special feature of such a computation of the microscopic field, distinguishing it from the usual Coulomb potential cutoff schemes, is that it employs the exact (unperturbed) Coulomb potentials. The continuum approximation outside the boundary surface does not impose any restriction on the potentials: with the true Coulomb forces, this scheme is well justified when R_c (or, in a nonspherical case, a characteristic size of the explicit solvent region) becomes much larger than the characteristic size of a solvent particle. Procedures of this sort have been mentioned in the literature.[26]

The situation is very different for the case of algorithm II, based on the explicit evaluation of field fluctuations, eq 14. The continuum response promoted by the solute charge does not contribute to fluctuations. In the expression (13) for the response kernel, the explicit continuum correction, similar to that in eq 22, disappears. Thereby, the long-range tails of Coulomb potentials enter computations of the field only implicitly, via interference terms and via the MD trajectory that is employed for preparation of the statistical ensemble of solvent configurations. This imposes rigid requirements upon the accuracy of treating Coulomb potentials in the course of those stages of a computation, which constitute its most cumbersome ingredient. For monatomic systems, the interference terms vanish identically, and accurate treatment of Coulomb forces during MD trajectory computations becomes especially important. Problems of explicit utilization of response field fluctuations as a tools for free energy computations have been mentioned in the literature[16] (see also the discussion in section 6D).

5. Classification of the Free Energy Contributions

Let us now combine field equations 20 and 22 and insert them in the energy expression 19 corresponding to the algorithm I. This provides a useful representation of the total electrostatic energy ΔF_{el}(alg. I) which is invoked in the forthcoming text:

$$\Delta F_{el}(\text{alg. I}) = \Delta F_d + \Delta F_{int} + \Delta F_c \tag{23}$$

According to this classification, the first term on the right-hand part is the "direct contribution" due to the pure Coulomb field of eq 20. The second term is a contribution from the continuum response Ψ_ν promoted by solvent charges; see eq 20. It is called the "interference contribution". The third term is the "continuum contribution" due to the continuum response promoted by solute charges. The first two terms represent microscopic effects averaged over the MD ensemble. The last term follows from a macroscopic continuum computation. Note that the interference term is absent in purely spherical (Born) systems.

In the case of algorithm II, eqs 20 and 22 are inserted in the energy eq 12. From the explicit form (eq 13) of the response kernel, it follows that the corresponding electrostatic free energy ΔF_{el}(alg. II) equals

$$\Delta F_{el}(\text{alg. II}) = \Delta F_{d} + \Delta F_{int} \tag{24}$$

Again, the direct and interference terms originate from the Coulomb and continuum field components of eq 17. The solute continuum term vanishes because the solute response does not contribute to fluctuations. Therefore, the interference term may be defined as $\Delta F_{int} = \Delta F_{el}(\text{alg. II}) - \Delta F_{d}$.

6. Computations for Single-Charged Ions in Water

Computations of the free energy (ΔF) for monatomic ions in water are studied in a number of papers. Several methodological problems have been discussed, namely, the estimates of effective Born radii,[27,28] a treatment of long-range electrostatic interactions in numerical simulations,[4,7] including the problem of finite cutoff radii,[29−31] etc. The solvent structure around the solute ion is often addressed.[32,33] Hydration ΔF values for a few single-charged ions obtained by different simulation procedures (Monte Carlo, MD, and RISM) are scattered in the literature.[34−37] Being nonsystematic, these data are insufficient for inferring a unique picture of the relative efficiency of simulation methods and other methodological problems like the truncation of potentials, the validity of parameters entering atom−atom interaction potentials (water−ion interaction potential parameters in the hydration case) etc. Below we report computational results for the equilibrium solvation energies of single-charged cations and anions in water. The simulations reproduce satisfactorily the experimental trends separately in series of cations and anions, as discussed below.

We have computed ΔF values for family monatomic anions and cations for relatively large MD cell (1000 water molecules) and using a large values of the cutoff radii, in comparison with refs 5 and 38−40. To the best of our knowledge, there exist three different sets of the LJ parameters for halogen anion and alkali-metal cations.[5,38−40] It should be noted that these sets were not calibrated to reproduce the solvation energy of monatomic ions. We calculated two components of ΔF using Berendsen parameters[39] and then investigated the influence of LJ parameter variations on the ΔF_{el} value using two alternative parametrizations.

Computations followed the strategy outlined in section 4. The calculation of the MD trajectory and the following calculation of the response field, to be substituted in algorithms 1 and 2, were totally separate and considered as two independent procedures. The statistical ensemble obtained at the MD stage was used for averaging either the instantaneous field $\tilde{\Phi}$ (eq 20) or its fluctuations (eq 14). The treatment of electrostatic potentials in the field calculations is, in principle, more accurate than during the MD stage. Coulomb interactions are considered completely (without cutoffs) within the cavity bounding the explicit solvent region (see boundary surface Σ in Figure 2). Although for spherical solutes the advantage of the report, more accurate determination of the potential is not relevant, it becomes significant for the nonspherical case studied in section 7.

6A. Computational Techniques. *MD Simulations.* Our MD simulations used GROMOS96[22] for NVT ensemble. Water (the SPC and SPC/E models) was considered as a solvent. The SHAKE algorithm was employed to constrain the geometry of each solvent molecule. The dynamical equations were integrated by means of Verlet algorithm. Weak coupling to an external temperature bath of 300 K with a coupling constant 0.4 ps for water was used to maintain temperature in all simulations. The procedure includes several cutoff radii:[22] those for pairwise atom−atom van der Waals interactions, separately for Coulomb interactions and also those introduced in terms of the reaction field approach complementing the truncation of Coulomb forces. All three types of cutoff radii were assumed to be equal; the corresponding parameter is denoted as R_{MD} below. The list of pair interactions has been regenerated after every 10 steps for algorithm I and after *every single* step for algorithm II. The time step was 2 fs for algorithm I and 1 fs for algorithm II. For each ion, a 100 ps (200 ps for algorithm II) trajectory was carried out to equilibrate ion−solvent system, followed by as least 500 ps (300 ps for algorithm II) of data collections for analysis. SPC water−ion interaction parameters for Na^+, K^+, F^-, Cl^-, and Br^- were recommended by Straasma and Berendsen,[39] but we considered also some alternative parametrizations.[43,44] The latter papers provide parameters for all alkali-metal ions and the halides.

The Mean Response Field and Fluctuations. Computations involved averages for the mean potential $\langle\tilde{\Phi}(r)\rangle$ and for the correlation function $\langle\Delta\Phi(r)\Delta\Phi(r')\rangle$. These averages used every *seventh* solvent configuration steps for algorithm I and *every single* configuration step for algorithm II on each MD trajectory. The ion center of mass was fixed at the center of the cubic MD cell. The radius of the spherical boundary of the explicit solvent region is denoted as R_c (see Figure 2). Following the prescription of ref 29, we always considered values $R_c \le R_{MD}$. A common practice of evaluating electrostatic interactions considering whole (solvent) molecules can lead to spurious dependence on the choice of the center of a molecule.[45] In particular, when the solute is an uncharged Lennard-Jones (LJ) sphere, the mean potential on the solute may have nonzero (negative or positive) values for exactly the same force field parameters.[46] Our calculations for the uncharged LJ spheres showed that the smallest nonzero values of the potential one can get using the geometric center of the water molecules. If the center of charge of the water molecules is used, the potential varies from ∼−2 to ∼−3 kcal/mol $(e)^{-1}$, depending on the LJ parameters of the uncharged sphere.

6B. Convergence of the Simulation Procedure. We discuss here only the results found for the electrostatic part of the free energy (ΔF_{el}). For a monatomic ion in a spherical cavity the general prescriptions of section 5 for the first algorithm (algorithm I) suggest

$$\Delta F_{el} = -\frac{Q}{2}\left[\langle\tilde{\Phi}(0)\rangle_{\rho,T} + \frac{Q}{R_c}\left(1 - \frac{1}{\epsilon_0}\right)\right] \tag{25}$$

Here the response field $\tilde{\Phi}$ is computed at the ion center and averaged over the ensemble with given charge distribution ρ ($\rho = Q\delta(r = 0)$) and temperature (T). This field is created by discrete solvent particles contained inside the spherical cavity with radius R_c. Interference terms completely vanish at $r = 0$ due to the condition $Sq_n = 0$. The continuum dielectric constant in the Born correction term is $\epsilon_0 = 70$ for water.[47] The second algorithm (algorithm II) suggests

$$\Delta F_{el} = -\frac{Q^2}{2k_BT}\langle\Delta\Phi(0)\,\Delta\Phi(0)\rangle_{\rho,T} = -\frac{Q^2}{2}K \tag{26}$$

We introduced here a special notation

$$K = \frac{1}{k_BT}\langle\Delta\Phi(0)\,\Delta\Phi(0)\rangle_{\rho,T} \tag{27}$$

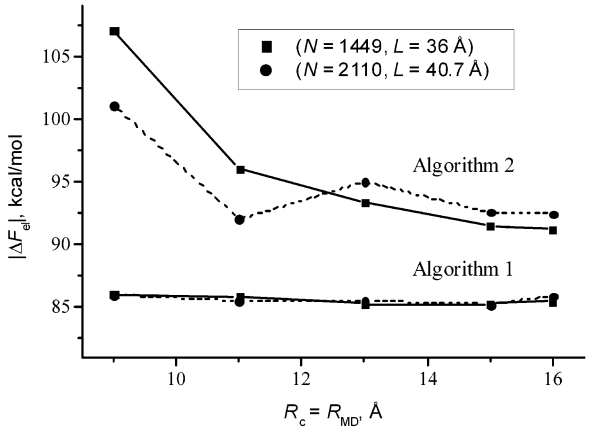

Figure 3. ΔF_{el} values for Cl^- in water calculated for two cells, $N = 1449$ and $L = 36$ Å (given as a solid line) and $N = 2110$ and $L = 40.7$ Å (given as a dotted line); see section 6B, as a function of cutoff radii ($R_c = R_{MD}$) varying from 9 to 15 Å.

for the absolute value of the response kernel (eq 13) at $r = r' = 0$. Because, in the framework of the linear response,

$$\langle \tilde{\Phi}(0) \rangle_{\rho,T} = Q\,K \tag{28}$$

we observe that algorithms I and II suggest two alternative ways for evaluating K (eqs 8 and 27, respectively) in terms of the average field and its dispersion.

The choice of radius R_c, which separates discrete and continuum solvent regions, is of key importance. In the present set of computations we assumed $R_c = R_{MD}$, and it was desirable to take R_c as large as possible. On the other hand, R_c cannot exceed $L/2$, where L is the edge of the MD cubic sell (see Figure 2). Increasing R_c from 10 to 20 Å increases the number N of explicit solvent (water) particles, involved in a simulation, from $N \approx 200$ to $N \approx 2000$. We extracted the acceptable R_c values, combining good convergence with a reasonable computational effort, from dependence ΔF_{el} on R_c. This test was performed for Cl^- in water with the parameters from ref 39 for two cells, ($N = 1449$, $L = 36$ Å) and ($N = 2110$, $L = 40.7$ Å). The results are illustrated by Figure 3.

The following conclusions are deduced from these computations:

(a) Algorithm I results in an almost converged value of ΔF_{el} even for small R_c, as was pointed by several authors.[3,26,39] The convergent limit is practically independent of the size of the pertaining MD cell and the number of explicit solvent particles involved in a simulation under the condition $L/2 \geq R_c$. The absolute value of the electrostatic hydration energy is close to the experiment (83[41] or 89[42] kcal/mol for Cl^- in water).

(b) The behavior of the $\Delta F_{el}(R_c)$ computed by algorithm II is irregular and the good convergence was not reached. The absolute values of ΔF_{el} are systematically larger than those obtained by algorithm I. It should be recalled, that the dependence of ΔF_{el} on R_c, computed by algorithm II, was obtained using the simulations for which the list of pair interactions has been regenerated after *every* step. In the conventional MD simulations, when this list is generated after 5 or 10 steps,[22] algorithm II leads to large scattering of the ΔF_{el} values, although this conventional prescription is quite sufficient as applied to algorithm I.

6C. Nonlinear Effects. According to the linear response approach the mean electrostatic field is proportional to Q, see eq 28, and the response kernel K is treated as Q-independent; the electrostatic energy is proportional to Q^2. The test of these trends was performed within the range $0 \leq |Q| \leq 1$. The

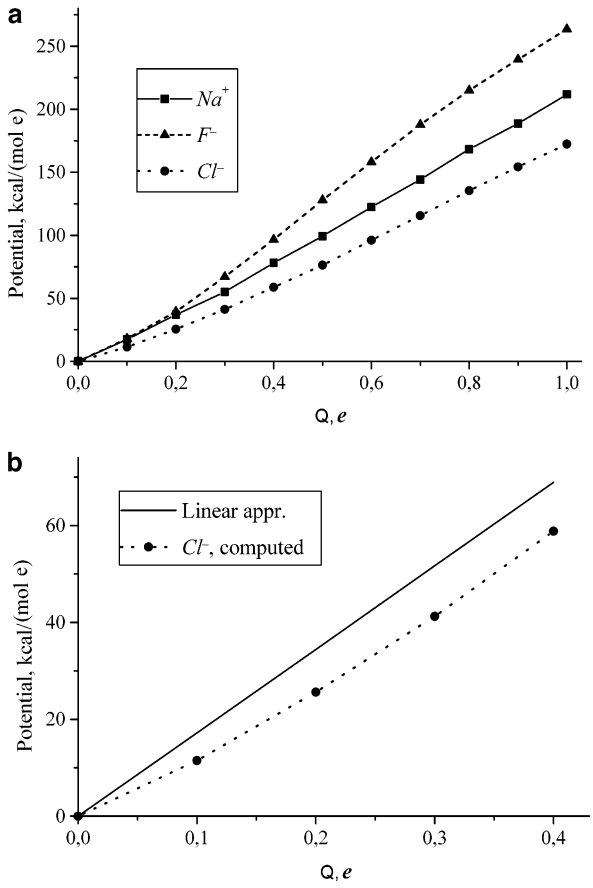

Figure 4. Average potential $\langle \tilde{\Phi}(0) \rangle_{\rho,T}$ (algorithm I) calculated for variable charge Q: (a) Na^+, F^-, and Cl^- in water when Q changes from 0 to $+1e$; (b) Cl^- in water when Q changes from 0 to $+0.4e$ (dotted line). Solid line corresponds to the linear approximation. The average potential at $Q = 0$ is assumed to be equal to 0; e means charge of the atomic unit.

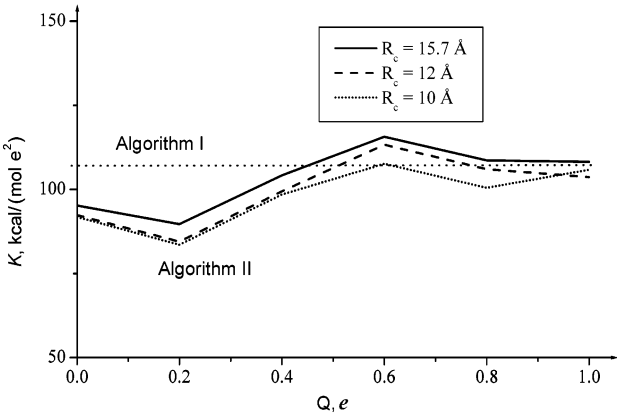

Figure 5. Values of kernel K for Na^+ in water, calculated using algorithm II (eq 27), when Q changes from 0 to $+1e$ for three different values of the cutoff radius R_c ($R_c \leq R_{MD}$). The K value, calculated using algorithm I (eq 28), is given as a dotted line.

illustrative examples are shown in Figures 4 and 5. According to our calculations, the linear approximation is valid within $\sim$ 3.5% of the total electrostatic free energy in the case Na^+ and $\sim$ 4% in the case of F^-. For Cl^- the nonlinear correction is larger ($\sim$ 7%). For the dipolar dumbbell system considered below (see section 7) the error depends on the separation of charges and constitutes less than 4.5%. The main conclusions are

(1) For the most cases mean field is practically linear in Q within the range $0 \leq |Q| \leq 1$ (see Figure 4a). More significant

deviations are found for the largest ion, namely Cl^-, at small values of Q only (see Figure 4b). This agrees with the results of refs[16,48,49] and certifies the validity of algorithm I for a majority of interesting cases. The origin of the observed nonlinearities has been discussed earlier.[16,48,50,51] The fact of their increase along with the ion size is remarkable.[16] It confirms the idea of solvent structure formation as an origin of the nonlinearity in weak solute fields. At stronger fields this effect disappears. For smaller ions their fields, with the same value of charge Q varying between 0 and 1, are stronger. This explains why the nonlinear effect is suppressed for them.

(2) The response kernel K found in terms of algorithm I is practically Q-independent. Its value as obtained by algorithm II fluctuates around a constant which is reached at $Q = 1$ (see Figure 5). This value is close to (but somewhat larger than) that of algorithm I. On the basis of the results of algorithm I, we conclude that the linear response is a quite reasonable approximation for single-charged ions. This is an empirical observation whereas its physical background may be a matter of special discussion.[16,50,52]

6D. Comparison of Algorithms I and II. Both algorithms may be, in principle, applied for a computation of ΔF_{el}. If performed quite accurately, the results are expected to coincide. On the other hand, observations summarized in sections 6B and 6C demonstrate that algorithm I is preferable. It requires much less computational effort, is simple and stable in regard to a variation of the MD cell size, cutoff parameters etc. These advantages are gained because the main effect of long-range interactions is explicitly taken into account by the Born correction in eq 25. According to algorithm II the long-range effects, as discussed at the end of section 4, appear implicitly and are smeared within a procedure preparing the MD ensemble. This accounts for a severe requirement in this case on a quality of a trajectory computation.

The basic problem with algorithm II seems to lie in the fact that the actual simulations result in a small nonzero value of the average field $\bar{\Phi}(0|0)$, eq 15, as discussed in refs 5,16,45,-46,51 and at the end of section 6A. The interpretation in terms of eqs 16−18 implies that K in eq 28 becomes singular at the point $Q = 0$ ($K \propto (1/Q)$). In the range of small absolute values of Q the linear approximation requiring that $K = \text{const}$ contains an inconsistency, and the basic expression 13 for the response kernel fails to be valid. On the other hand, as the average field $\bar{\Phi}(r)$ remains regular, the algorithm I has no problems of this sort.

6E. Nonelectrostatic Interactions. Free energy of creation of the uncharged LJ spheres ΔF_{nel}, was calculated using thermodynamic integration. For a given LJ sphere MD simulations were performed for a a control parameter λ,[53] varying between 0 and 1. For each value of λ the length of the MD run was 200 ps, with a step of 2 fs; the first 50 ps the system was equilibrated. The test calculations of creating neon[54] in a box of 262 SPC/E water molecules give $\Delta F_{nel} = 2.84$ kcal/mol, in accord with the experimental and the earlier MD data.[55]

At the next step the influence of the box size on the ΔF_{nel} value was investigated. Calculations of creating chlorine (Straasma and Berendsen parameters[39]) have been done for the two boxes containing 262 and 1000 SPC water molecules. The calculated ΔF_{nel} values were found equal to 6.16 and 5.58 kcal/mol, respectively. This shows the minor effect of the box size on the computation of nonelectrostatic interactions.

The ΔF_{nel} values calculated for uncharged LJ spheres, corresponding to the set of monatomic ions, is given in Table 1. The box of 262 SPC water molecules was used, solute/solvent LJ parameters were taken from ref 39. Obtained ΔF_{nel} values increase with increasing of van der Waals radius. Nonelectrostatic free energy contributions constitute from 2 to 8% of the total hydration energy.

TABLE 1: Comparison of the Experimental Hydration ΔF Values for Several Monatomic Ions with the Computed Values, Where Two Components of ΔF, Calculated in the Present Paper, Are Also Given (Units kcal/mol)

	exp		this work[a]		
ion	42	41	ΔF	ΔF_{el}	ΔF_{nel}
Na^+	−88.9	−89.7	−104.5	−106.8	2.3
K^+	−71.3	−72.7	−75.2	−81.3	6.1
F^-	−120.	−113	−129.1	−131.8	2.7
			−120.5[b]	−123.2[b]	
Cl^-	−89.2	−83.0	−80.1	−86.3	6.2
Br^-	−82.8	−76.8	−77.0	−83.6	6.6

[a] Berendsen LJ parameters were employed.[39] [b] On the basis of the GROMOS-96 LJ parameters suggested for hydrogen-bonded (F)−(H_2O). Analogous parameters are not available for the other ions.

A significant computational advantage of the present strategy, separating nonelectrostatic and electrostatic parts of a computation is evident. The nonelectrostatic run requires computing many trajectories (for the integration procedure), which can be performed in a small box. The electrostatic run requires a single trajectory in a large box. If we combine them together, the computation would effectively include many trajectories in a large box.

6F. Solvation Free Energies. Total solvation energies are composed of nonelectrostatic and electrostatic components: $\Delta F = \Delta F_{nel} + \Delta F_{el}$.The first component comprises the free energy change due to insertion of a hypothetical uncharged solute particle, described by the given solute/solvent LJ parameters, in the bulk solvent. The so obtained solute state is then involved in the charging process to give the second component. Both components and their sum for a set of ions in water are listed in Table 1. The computations of ΔF_{nel} were performed as described in section 6E. ΔF_{el} was computed by algorithm I ($R_c = 15$ Å) for a given monatomic ion in cubic MD cell ($N = 1000$, $L = 31$ Å, $R_{MD} = 15$ Å). The solute/solvent LJ parameters were borrowed from ref 39.

The electrostatic and nonelectrostatic parts should be combined when comparing with the experiment. Since ΔF_{nel} constitutes only from 2 to 8% of ΔF, the experimental trends in hydration energies are reproduced by ΔF_{el}. Moreover, the experimental trend of $|\Delta F|$ in the order $F^- > Na^+ \sim Cl^- > Br^- > K^+$ was found in the present computations. This seems to resolve the problem, discussed in the literature,[38−40] that free energy computations for cations and anions are incompatible. Given the uncertainty in the experimental values,[41,42] detailed quantitative comparisons are not warranted. Comparison of result based on the linear model and thermodynamic integration (see section 6C) suggests that the linear model somewhat overestimates the ΔF_{el} magnitudes.

ΔF_{el} values, obtained with the alternative solute/solvent LJ parameters for the SPC[43] and SPC/E[44] models, are given in Table 2. One can see that computed values of ΔF_{el} are very sensitive to the variation these parameters. Such a phenomenon can be explained in the following manner. The pertaining parameter variations change the mean distance between the ion and water molecules, which constitute the first solvation shell. This, obviously, modifies the value of the response field $\bar{\Phi}$.

It was not the objective of the present work to reproduce accurately experimental solvation energies. We only intended to show that the present original computational scheme is

TABLE 2: ΔF_{el} Values for Alkali Metal and Halide Ions Computed by Algorithm I (R_c = 15 Å) for Cubic MD Cell (N = 1000, L = 31 Å, R_{MD} = 15 Å) Where Three Different Sets of Solute/Solvent Lennard-Jones Parameters Were Used (Units kcal/mol)

ion	Berendsen params[a]	Lin−Jordan params[b]	Lee-Rasaiah params[c]
Li^+		−100.9	−117.9
Na^+	−106.8	−85.3	−96.6
K^+	−81.3	−75.1	−78.8
Rb^+		−69.6	−75.0
Cs^+		−65.2	−68.9
F^-	−131.8	−124.1	−121.4
Cl^-	−86.3	−92.9	−86.0
Br^-	−83.6	−90.6	−81.9
I^-		−75.6	−70.9

[a] Reference 39, the SPC model of water. [b] Reference 43, the SPC model of water. [c] Reference 44, the SPC/E model of water.

compatible with conventional simulation procedures and gives similar results. This is demonstrated in Tables 1 and 2. Better agreement with the experiment could be gained by specially fitting the solute/solvent LJ parameters; such prescriptions are sometimes invoked.[49] As indicated in Table 1, there is a significant difference in ΔF_{el} values obtained for F^- using, LJ parameters for hydrogen-bonded or non-hydrogen-bonded cases, with the former set giving closer correspondence with the experimental estimates. For the other ions, LJ parameters specifically adjusted for hydrogen-bonded situations are not available.[22]

Thereby, we conclude that reasonable estimates of ion hydration energies are available by using only their electrostatic component and computing it within the linear response approach by algorithm I (R_c = 15 Å) with a single MD trajectory (the time step is 2 fs, 100 ps for equilibration and 200 ps for data collections) obtained for a relatively large MD cell (N = 1000, L = 31 Å, R_{MD} = 15 Å). With the present nonpolarizable solvent model a better fitting of absolute values of ΔF_{el} could be gained in terms of a refined parametrization of nonelectrostatic interactions.

6G. Accuracy of the Computed Free Energies. The computed ΔF_{el} values incorporate two distinct sources of error (i.e., disagreement with the "ideal" computation); see, e.g., ref 57. MD simulations and subsequent calculations of the microscopic solvent response field $\tilde{\Phi}$ introduce a numerical error which is supplemented by a statistical error arising at the stage of averaging.[20] Our experience estimates the pertaining error in the free energy as 1 kcal/mol or less if one evaluates $\tilde{\Phi}$ applying algorithm I with a sufficiently long MD trajectory. The mixed discrete/continuum model, however, introduces an additional uncertainty (around 1 kcal/mol). It is caused by the dielectric permittivity, ϵ_0, the parameter used both in the MD and a field calculations. Our calculations for Cl^- show that variation of the ϵ_0 value from 50[22] to 80[39] reduces the ΔF_{el} (algorithm I, R_c = 15 Å) from 86.3 to 85.4 kcal/mol. Another ambiguity arises in algorithm I due to the ill defined treatment of explicit solvent molecules in the vicinity of the boundary of the cavity[45,46] (see a discussion at the end of the section 6A). A similar problem arises at the stage of computing interference fields Ψ_ν(eq 20) for solvent charges q_ν localized close to the boundary. The last two effects seem to be interrelated and are expected to cancel one another in a fully consistent "ideal" computation, which was not attained. Summing up, the mean deviation in the calculated ΔF_{el} values due to all these reasons is likely to be a few kcal/mol.

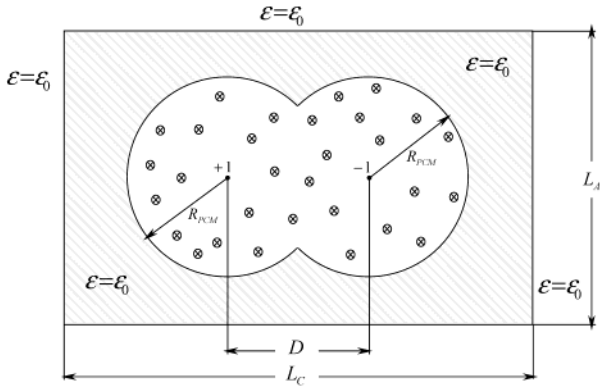

Figure 6. Scheme of the dumbbell computation by the PCM method.[8] The cavity is formed by two overlapping spheres with radius R_{PCM}. Dielectric constant is ϵ = 1 inside the cavity where explicit solvent particles (circled crosses) are located. A continuum solvent model with $\epsilon = \epsilon_0$ is accepted outside the cavity.

7. The Two-Site (Dumbbell) Solute Model

7A. Description of the Model and Motivation. This purely methodological computation was inspired by the earlier studies[48,57] which pointed to specific problems inherent to several truncation schemes of Coulomb potentials in MD simulations. For a dumbbell with a distance D between the two charge sites they were sometimes revealed as a striking inconsistency in the dependence on D of the mean force potential, $\Delta\Delta F_{el}$ (D). We shall demonstrate that this is not the case for the present methodology. $\Delta\Delta F_{el}$ (D) can be written as

$$\Delta\Delta F_{el}(D) = \Delta F_{el}(D) - 1/D + \Delta F_{el}(D = \infty) \quad (29)$$

where ΔF_{el} is the solvation free energy, $1/D$ describes the interaction between the two charge sites (ions) of the dumbbell and $\Delta F_{el}(D = \infty)$ corresponds to the solvation free energy of the isolated ions. ΔF_{el} is written as

$$\Delta F_{el} = \frac{1}{2}\sum_i Q_i \bar{\Phi}(r_i) \quad (30)$$

Here i = 1, 2 label the sites with charges Q_i (equal to ±1) and positions r_i; total average response field $\bar{\Phi}$ is given in eq 22.

Three different computational schemes were used to calculate the response field. They correspond to three different methods of constructing the solvation shell around the dumbbell solute. We considered cavities of different shapes with the boundaries Σ. According to algorithm I the explicit molecular treatment of the microscopic response field $\tilde{\Phi}(r)$ (see eq 20) was performed inside Σ. It was complemented by the continuum computation outside S. The first method corresponds to the Kirkwood computation with spherical cavity of radius R_c ($R_c > D$) centered at the middle point of the dumbbell particle (see Figure 2). The second method corresponds to the PCM cavity[8] with two spheres R_{PCM} around each dumbbell site (see Figure 6). The third method is Ewald summation for computing the field $\tilde{\Phi}(r)$ created by a periodic lattice of solvent particles. Exactly the same MD trajectory was used in all three cases. The field $\tilde{\Phi}(r)$ was averaged (with the same MD ensemble), complemented by a continuum correction for first two cases and inserted in eq 30.

In the present work, we also studied how the change of the spherical shape of the external cavity affects the computational results and estimated the contribution of interference effects to the solvation free energy, see eqs 23 and 24.

Calculations were performed for the Cl^-...Cl^+ system in water (the SPC model) with the parameters from ref 39 for a

TABLE 3: Electrostatic Part of the Solvation Energy ΔF_{el} Evaluated Using the Three Different Computational Schemes and the Different Terms of ΔF_{el} (Eq 30) for the Ion Pair $\{Cl^+, Cl^-\}$ as Functions of Separation D (with Energies in kcal/mol)

	three contributions to ΔF_{el} (the PCM model)								total ΔF_{el}				
	direct term			continuum term		interference term			PCM				
	alg. I [a]			alg. I[a]		alg. I[a]			alg. I[a]				
D, Å	9 Å	15 Å	alg. II 15 Å	9 Å	15 Å	9 Å	15 Å	alg. II 15 Å	9 Å	15 Å	alg. II 15 Å	Kirkwood alg I[b]	Ewald
2	28.15	28.77	30.92	0.71	0.16	−0.06	−0.07	−1.00	28.80	28.86	29.92	28.86	28.92
3	49.02	50.04		1.44	0.35	−0.15	−0.15		50.31	50.24		50.16	49.45
3.5	58.86	59.79		1.86	0.46	−0.28	−0.09		60.44	60.16		60.14	59.74
4	67.56	69.50	80.32	2.32	0.58	0.02	−0.03	−4.16	69.90	70.05	76.16	70.03	69.61
6	94.94	97.01	113.12	4.40	1.17	−0.61	−0.07	−8.28	98.73	98.11	104.84	98.07	97.56
8	104.43	109.32		6.70	1.87	−0.05	0.14		111.08	111.33		111.25	111.13
10	112.26	116.65	154.20	9.08	2.64	−0.66	0.34	−18.44	120.68	119.63	135.76	119.76	118.74

[a] PCM computations with 9 Å and 15 Å corresponding to the value of R_{PCM} radius. [b] Kirkwood computations with $R_C = R_{MD} = 15$ Å.

rectangular cell with edges $L_A = L_B = 30.5$ Å and $L_C = 40.5$ Å, $N = 1300$, see Figure 6. The value of D varied from 2 to 10 Å. The MD trajectory was computed in terms of the conventional reaction field methodology ($R_{MD} = 15$ Å, see section 6A). The list of pair interactions has been regenerated after every 10 steps for algorithm I and after *every single* step for algorithm II. The time step was 2 fs for algorithm I and 1 fs for algorithm II. For each D value, a 100 ps (200 ps for algorithm II) trajectory was carried out to equilibrate dumbbell-solvent system, followed by as least 200 ps (300 ps for algorithm II) of data collections for analysis. The free energies of separated ions Cl^- and Cl^+ were computed by the PCM method for the *same* rectangular cell, with the ion located in the center of the cell.

Kirkwood[18] and PCM[8] computations were standard. The Ewald computation followed prescriptions of refs 21 and 48, and the solute particle was withdrawn when the MD cell was replicated.[58]

7B. The Results. Table 3 shows the electrostatic part of the solvation energy, ΔF_{el}, as a function of D, calculated in terms of the PCM and Kirkwood models and the Ewald summation. Some additional data is given for the PCM calculations: the distribution of the components for the ΔF_{el}, see eqs 23 and 24, evaluated using algorithm I and algorithm II, and the influence of the cutoff radius R_{PCM} on the relative contributions of the direct and continuum terms. The following conclusions are deduced from Table 3:

•The PCM and Kirkwood models, with two different nonperiodic solvation shells, and the Ewald summation, with a periodic solvation shell, give practically the same ΔF_{el} values, within 1 kcal/mol. The value of the total solvation energy decreases monotonically with increasing of D.

•The cutoff radius R_{PCM} (the PCM model, algorithm I) influences significantly relative contributions of the direct and continuum terms of eq 23. The interference term is always negligible. The same trends are observed for Kirkwood computations with different cutoff radii R_c (they are not included in Table 3). Despite strong variations of partial contributions, the total ΔF_{el} value is quite stable to the variation of cutoff radii, R_{PCM} and R_c, from 15 to 9 Å. Some discrepancies appear when R_c becomes smaller than D, i.e., at $D = 10$ Å.

•In contrast to algorithm I, algorithm II (the PCM model) gives a relatively large value of the interference term. For large D, e.g., at $D = 10$ Å, its value is around 13% of the total ΔF_{el}. The response kernel (see eq 13) was computed with site charges $Q_i = \pm 0.5e$. Inclusion of the interference correction improves the agreement of the total energy with the result of a regular (algorithm I) computation. This observation is in accord with the discussion of drawbacks of algorithm II in section 4.

The basic conclusion is that algorithm I provides a stable and reliable procedure for computing electrostatic free energies.

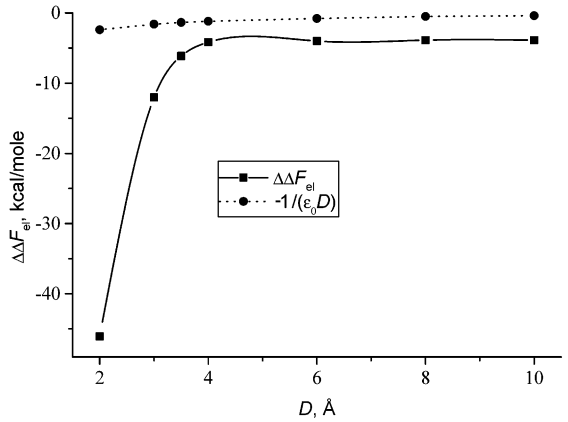

Figure 7. Potential of mean force, $\Delta\Delta F_{el}$ (D), see eq 29, for ion pair Cl^- ... Cl^+ in water computed using the PCM model (algorithm I). The dashed line is the dielectric continuum potential of mean force, $-1/\epsilon_0 D$, with $\epsilon_0 = 70$.

Its PCM and Kirkwood realizations are well adapted for treating both charged and strongly polar (with zero net charge) solutes. For solutes with shapes that deviate significantly from a sphere, the PCM procedure is preferable.

In Figure 7, we show the dependence of the potential of mean force, $\Delta\Delta F_{el}$, on D, see eq 29. It is calculated using the PCM model (algorithm I). Computed curve shows an asymptotic change of $\Delta\Delta F_{el}$ approaching a constant limit. The transition to the asymptotic behavior occurs in the region around $D = 4$ Å. (Note that the sum of van der Waals radii of the dumbbell spheric components is 3.75 Å). In terms of the continuum medium description, the electrostatic energy ΔF_{el} changes asymptotically as

$$\Delta F_{el}(D) = 1/D - 1/(\epsilon_0 D) + \text{const}$$

The significance of the term $1/(\epsilon_0 D)$ was mainly addressed and thoroughly discussed earlier.[48,57] We note that it ranges as $1.2 < 1/(\epsilon_0 D) < 0.5$ kcal/mol for 4 Å $< D <$ 10 Å. Because its magnitude is beyond the accuracy of a typical MD computation (see section 6G), we will avoid a discussion of the corresponding effects. On the coarse-grained energy scale the asymptotic free energy change is well described as

$$\Delta F_{el}(D) = 1/D + \text{const}$$

The computation gives the constant equal to $\Delta F_{el}(D = \infty) - 4$ kcal/mol. (The value $\Delta F_{el}(D = \infty) = -149$ kcal/mol is the free energy of separated ions Cl^- and Cl^+; see section 7A.) This constant is the same for PCM and Kirkwood models, but is different for the case of the Ewald summation. The discrep-

ancy between the continuum corrected and Ewald computations arises at the stage of treating isolated ions. Within our version of the Ewald procedure, there exists an ambiguty in the estimate of $\Delta F_{el}(D = \infty)$ in eq 29 requiring a consideration of charged solutes. The pertaining problems of the Ewald method are discussed in the literature.[6,7,21,31,59]

8. Conclusion

In the present work, we considered systematically whether a linearized MD theory can be applied as a tool for regular computations of electrostatic solvation free energies of strongly charged (or consisting of strongly charged fragments) solutes in polar solvents. It is shown that the linear approximation is valid within $\sim 5\%$ or less of the total electrostatic free energy. Similar observations and conclusions have been reported earlier.[16] Most significant nonlinear effects were observed for ionic systems where the charged site had a larger van der Waals radius (i.e., for the Cl^- ion, where the deviation due to the linear approach reaches 7.1%). Nonlinear effects systematically (by 3−4 kcal/mol in the present cases) exaggerate the solvation energies. Effects of such order of magnitude are within the uncertainty of conventional simulation techniques. They can be readily corrected and modified by a slight variation of internal parameters of the MD scheme.

Therefore, the linear approach for treating electrostatic effects seems to be a reasonable approximation. It provides both conceptual and computational advantages. A significant computational benefit is gained due to a separation of a total free energy computation in two steps: one treating the nonelectrostatic component at full nonlinear level but without electrostatic interactions and the next one treating purely electrostatic interactions within the linear approach. This is demonstrated and discussed in section 6E. A related application to the study of vaporization energies of bulk alcohols has been reported.[60]

The linearization of electrostatic fields can be performed in terms of two algorithms which are quite different in their nature and realization. Only algorithm I, based on a microscopic computation of the response field $\tilde{\Phi}$, proved to be sufficiently stable and efficient. It is this representation of the linearization procedure which is usually discussed in the literature. Algorithm I is well adapted for introducing continuum corrections to $\tilde{\Phi}$ in order to treat properly long-range electrostatic forces. The underlying procedure, that separates a computation of the MD trajectory and the subsequent computation of $\tilde{\Phi}$, implies that the last step must be performed most accurately, with an explicit treatment (excluding any cutoffs) of electrostatic interactions, within a large cavity subject to a totally microscopic consideration. Advantages of this strategy are lost in algorithm II, based on a microscopic computation of fluctuations of the response field $\tilde{\Phi}$. A different level of accuracy accepted for the first (a MD computation) and the second (the response field computation) computational steps creates an internal inconsistency which hinders a straightforward application of the fundamental eq 13. This results in instabilities, slow convergence and large statistical errors observed in our tests of algorithm II. For its successful application the background computational scheme must be changed. On the other hand, the based on algorithm I linearized molecular/continuum procedure seems to provide a further perspective for computations of solvation free energies for polyatomic systems, including both equilibrium and nonequilibrium effects.

Acknowledgment. M.V.B., M.V.V., I.V.L. and Y.A.D. thank the Russian Foundation of Fundamental Research (Projects Nos. 02-03-33049 and 00-15-97295 and the Civilian Research and Development Foundation (CRDF) for the Independent States of the Former Soviet Union (Award No. RC2-2209) for financial support. The research at Karpov Institute was supported in part by the International Association for the promotion and cooperation with scientists from the New Independent States of the former Soviet Union (project INTAS-RFBR IR-97-620). The research at Brookhaven National Laboratory was carried out under contract DE-AC02-98CH10886 with the U.S. Department of Energy and supported by its Division of Chemical Sciences, Office of Basic Energy Sciences. The authors are indebted to Dr. Janos Angyan for a stimulating discussion of the linearized MD approach at the preliminary stage of the present work.

References and Notes

(1) *Simulation of Electrostatic Interactions in solution*; Pratt, L. R., Hummer, G., Eds.; AIP: New York, 1999.
(2) Levy, R. M.; Belhadj, M.; Kitchen, D. B. *J. Chem. Phys.* **1991**, *95*, 3627.
(3) Pratt, L. R.; Hummer, G.; Garcia, A. E. *Biophys. Chem.* **1994**, *51*, 147.
(4) Figueirido, F.; Del Buono, G.; Levy, R. M. *J. Phys. Chem. B* **1997**, *101*, 5622.
(5) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Phys. Chem.* **1996**, *100*, 1206.
(6) Levy, R. M.; Gallicino, E. *Annu. Rev. Phys. Chem.* **1998**, *49*, 531.
(7) Hummer, G.; Pratt, L. R.; Garcia, A. E.; Neumann, M. In ref 1, p 84.
(8) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
(9) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.
(10) Kubo, R.; Toda, M.; Hatshisume, N. *Statistical Physics II*; Springer: Berlin, 1985.
(11) Rostov, I. V.; Basilevsky, M. V.; Newton, M. D. In ref 1 p 331.
(12) Newton, M. D.; Friedman, H. L. *J. Chem. Phys.* **1988**, *88*, 4460.
(13) Basilevsky, M. V.; Chudinov, G. E. *Chem. Phys.* **1991**, *157*, 327, 345.
(14) Basilevsky, M. V.; Parsons, D. F. *J. Chem. Phys.* **1998**, *108*, 9107.
(15) Song, X.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 2594.
(16) Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512.
(17) Landau, L. D.; Lifshits, E. M. *Electrodynamics of continuous media*; Nauka: Moscow, 1982.
(18) Kirkwood, J. G. *J. Chem. Phys.* **1939**, *7*, 39.
(19) Angyan, J. *J. Math. Chem.* **1992**, *10*, 93.
(20) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1993.
(21) Figueirido, F.; Del Buono, G. S.; Levy, R. M. *J. Chem. Phys.* **1995**, *103*, 6133.
(22) Van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v.: Zürich, Switzerland, and Groningen, The Netherlands, 1996.
(23) Tironi, I. G.; Sperb, R.; Smith, P. E.; Van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451.
(24) Hummer, G.; Soumpasis, D. M.; Neumann, M. *Mol. Phys.* **1992**, *77*, 769.
(25) Tanford, C.; Kirkwood, J. G. *J. Am. Chem. Soc.* **1957**, *79*, 5333.
(26) Resat, H.; McCammon, J. A. *J. Chem. Phys.* **1996**, *104*, 7645.
(27) Hyun, J.-K.; Ichiye, T. *J. Phys. Chem. B* **199**7, *101*, 3596.
(28) Babu, C. S.; Lim, C. *J. Phys. Chem. B* **1999**, *103*, 7958 and references therein.
(29) Sakane, S.; Ashbaugh, H. S.; Wood, R. H. *J. Phys. Chem. B* **1998**, *102*, 5673.
(30) Resat, H.; McCammon, J. A. *J. Chem. Phys.* **1998**, *108*, 9617.
(31) Rick, S. W. In ref 1, p 114, and references therein.
(32) White, J. A.; Schwegler, E.; Galli, G.; Gydi, F. *J. Chem. Phys.* **2000**, *113*, 4668 and references therein.
(33) Raugei, S.; Klein, M. L. *J. Chem. Phys.* **2002**, *116*, 196 and references therein.
(34) Hirata, F.; Redfern, P.; Levy, R. M. *Int. J. Quantum Chem. Biol. Symp.* **1988**, *15*, 179.
(35) Tunon, I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-Lopez, M. F. *Chem. Phys. Lett.* **1995**, *241*, 450.
(36) Kalko, S. G.; Sese, G.; Padro, J. A. *J. Chem. Phys.* **1996**, *104*, 9578.
(37) Borodin, O.; Bell, R. L.; Bedrov, Y. Li, D.; Smith, G. D. *Chem. Phys. Lett.* **2001**, *336*, 292; *J. Chem. Phys.* **1998**, *109*, 10921.

(38) Chandrasekhar, J.; Spellmeyer, D. C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1984**, *106*, 903.
(39) Straatsma, T. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1988**, *89*, 5876.
(40) Migliori, M.; Corongiu, G.; Clementi, E.; Lie, G. C. *J. Chem. Phys.* **1988**, *88*, 7766.
(41) Marcus, Y. *Ion Solvation*; Wiley: Chichester, England, 1985; pp 107−108.
(42) Schmid, R.; Miah, A. M.; Sapunov, V. N. *Phys. Chem. Chem. Phys.* **2000**, *2*, 97.
(43) Lin, S.; Jordan, P. C. *J. Chem. Phys.* **1988**, *89*, 7492.
(44) Lee, S. H.; Rasaiah, J. C. *J. Phys. Chem.* **1996**, *100*, 1420.
(45) Hummer, G.; Pratt, L. R.; Garcia, A. E.; Berne, B. J.; Rick, S. W. *J. Phys. Chem. B* **1997**, *101* 3017.
(46) Åqvist, J.; Hansson, T. *J. Phys. Chem. B* **1998**, *102*, 3837.
(47) Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Chem. Phys.* **2001**, *115*, 1125.
(48) Del Buono, G. S.; Figueirido, F.; Levy, R. M. *Chem. Phys. Lett.* **1996**, *263*, 521.
(49) Cummins, P. L.; Gready, J. E. *J. Comput. Chem.* **1999**, *20*, 1028.
(50) Fonseca, T.; Ladanyi, B. M. *J. Phys. Chem.* **1991**, *95*, 2116.
(51) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Am. Chem. Soc.* **1997**, *119*, 8523.
(52) Hyun, J.-K.; Ichiye, T. *J. Chem. Phys.* **1998**, *109*, 1074.
(53) Mezei, M.; Beveridge, D. L. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1.
(54) Straatsma, T. P.; Berendsen, H. J. C.; Postma, J. P. M. *J. Chem. Phys.* **1986**, *84*, 6720.
(55) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025 and references therein.
(56) Shore, B. W. *J. Chem. Phys.* **1973**, *59*, 6450.
(57) Bader, J. S.; Chandler, D. *J. Phys. Chem.* **1992**, *96*, 6423.
(58) Lynden-Bell, R. M.; Rasaiah, J. C. *J. Chem. Phys.* **1997**, *107*, 1981.
(59) Hünenberger, P. H.; McCammon, J. A. *J. Chem. Phys.* **1999**, *110*, 1856.
(60) Martin, M. E.; Sanchez, M. L.; Olivares del Valle, F. G.; Aguilar, M. A. *J. Chem. Phys.* **2002**, *116*, 1613.