Solvation of complex molecules in a polar liquid: An integral equation theory

Dmitrii Beglov and Benoît Roux
Centre de Recherche en Calcul Appliqué (CERCA), Departments of Physics and Chemistry, Université de Montréal, C.P. 6128, succ. Centre-Ville, Canada H3C 3J7

(Received 1 November 1995; accepted 23 February 1996)

A statistical mechanical integral equation theory is developed to describe the average structure of a polar liquid around a complex molecular solute of irregular shape. The integral equation is formulated in three-dimensional Cartesian coordinates from the hypernetted chain (HNC) equation for a solute at infinite dilution. The direct correlation function of the pure solvent used in the theory is taken from the analytical solution of the mean spherical approximation (MSA) equation for a liquid constituted of nonpolarizable hard spheres with an embedded dipole at their center. It is demonstrated explicitly that, in the limit where the size of the solvent particles becomes very small, the present theory reduces to the well-known equations for macroscopic electrostatics in which the solvent is represented in terms of a dielectric continuum. A linearized version of the integral equation corresponds to a three-dimensional extension of the familiar MSA equation. This 3D-MSA integral equation is illustrated with numerical applications to the case of an ion, a water molecule, and *N*-methylacetamide. The numerical solution, obtained on a discrete three-dimensional cubic grid in Cartesian coordinates, yields the average solvent density and polarization density at all the points x,y,z around the solute. All spatial convolutions appearing in the theory are calculated using three-dimensional numerical fast Fourier transforms (FFT). © *1996 American Institute of Physics.* [S0021-9606(96)51520-5]

I. INTRODUCTION

Computer simulations of atomic models with explicit water represent the most detailed method to study the influence of solvent on the conformation of complex biological molecules.[1] Nevertheless, this computationally intensive methodology often suffers from statistical uncertainties due to finite sampling. In particular, the average properties of nonuniform systems, such as the solvent structure around a single biomolecule, converge slowly and are difficult to obtain accurately.[2] It thus remains worthwhile to develop alternative theoretical methods to account for the solvation of biomolecules by treating the influence of solvent implicitly.

A popular approach to treat the influence of solvation consists in representing the polar solvent in terms of a structureless dielectric continuum.[3] In the continuum approximation, the dielectric constant corresponding to the interior of the solute and the bulk solvent are assigned different values and the average electrostatic potential resulting from the solvent reaction field is calculated by solving the Poisson equation for macroscopic media.[3] The approximation was first introduced by Born to describe the solvation of spherical ions,[4] and later extended by Kirkwood to the case of an arbitrary charge distribution embedded inside a spherical cavity.[5] At the present time, the Poisson equation can be solved routinely using numerical methods in the case of molecules of irregular complex shapes.[3] Numerical calculations show that the continuum approximation can be very effective, for example, in the calculation of solvation free energies for a variety of molecules.[3,6] Nevertheless, it fails to provide information about the solvent at the molecular level. For example, the discreteness and granularity of the solvent particles is ignored completely.[7] Furthermore, the continuum macroscopic electrostatics is based on the assumption that, at any point **r**, the average polarization density $\mathbf{P}(\mathbf{r})$ and the total electrostatic field $\mathbf{E}^{\mathrm{tot}}(\mathbf{r})$ are colinear and their relative magnitude is governed by the dielectric constant.

Alternative approaches attempting to address the limitations of continuum macroscopic electrostatics have been proposed. In the protein dipole Langevin dipole (PDLD) method,[8] the average solvent polarization is determined on a three dimensional lattice of molecular scale meant to represent the discreteness of the solvent particles. To account for saturation of the solvent polarization, the magnitude of the polarization vectors induced at the lattice point is determined self-consistently from the total local electrostatic field using the nonlinear Langevin polarization function. The dimension of the lattice and the polarizability assigned to each lattice point are adjusted empirically to yield the correct dielectric constant for the bulk fluid. The method was shown to be effective in calculating the electrostatic contributions to the free energy of various solutes.[8] However, variations of the electrostatic potential on a microscopic length scale are neglected because the induced polarization vectors are related to the *local* electrostatic field, as in the macroscopic electrostatic approximation. Mean-field models incorporating microscopic considerations have also been proposed to describe the hydration and solvation forces due to water.[9,10] The structure of liquid water was modeled as ice with a large number of Bjerrum defects, i.e., a random network of hydrogen-bonds, with many of the bonds strained or broken. To take into account the molecular nature of water, a correlation length ξ associated with the decay of a perturbation in

 0021-9606/96/104(21)/8678/12/$10.00 © 1996 American Institute of Physics

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

the polarization density, was introduced. Following these ideas, differential equations were derived to describe the polarization of water near a planar wall. The approach described well the hydration forces between two plates and seemed promising. However, it remained difficult to derive appropriate boundary conditions for the differential equation in the case of complex solutes with irregular shape.

Integral equation theories of liquid solutions provide a rigorous statistical mechanical approach to incorporate the microscopic nature of the solute and solvent in a description of solvation phenomena.[11] However, because integral equation theories based on full molecular distribution functions can rapidly become too complex for attempting numerical calculations, it is necessary to introduce some approximations. The reference interaction site model (RISM) is an approximate integral equation theory constructed on the basis of reduced site–site radial distribution functions.[12,13] Because it reduces to a set of one-dimensional radial integral equations, the approach is computationally tractable. The RISM theory,[12] extended for polar systems,[14] has been used to investigate the influence of water on the conformational equilibrium of flexible molecules such as *n*-butane,[15] dipeptides,[16–19] and tripeptides.[20] Experience has shown that RISM often succeeds remarkably in reproducing the main features of solvation in dense liquids. However, the RISM theories that are currently available are not appropriate for large molecular solutes; e.g., shielding of buried sites from the solvent is only partially accounted for.[18] In addition, only reduced site–site radial distribution functions appear naturally within the RISM framework. Consequently, no information is available about important quantities, such as **P(r)**, the average solvent polarization density at a point **r** in the neighborhood of the solute. Theories such as the reference hypernetted chain equation (R-HNC) provide a much more complete representation of the solvent distribution functions.[21] For example, the R-HNC equation has given insight on the structure of the double layer by considering a large spherical macroion solute immersed in an ionic solution mixture.[22] However, numerical solutions of the R-HNC equation are usually based on an expansion of the correlation functions in terms of rotational invariants.[21,23,24,25] Although such an expansion provides an effective method to solve the integral equation when the shape of the solute substrate is spherical or nearly spherical,[21,23,24,25] this numerical approach may not converge rapidly when the shape of the substrate is markedly nonspherical as in the case of biomolecular solutes. Recently, a simpler approximation for describing the distribution function of water around a globular solute was proposed.[26,27] The water distribution function was assumed to be the product of a position-dependent part and an orientational factor. The spatial part was calculated from the hypernetted chain (HNC) equation while the orientational factor was approximated by a Boltzmann distribution corresponding to the interaction of a water dipole with a local effective orienting field. Using arguments similar to those used in the Onsager theory of polar liquids,[28] the effective orienting field was approximated by the total electrostatic field multiplied by a prefactor equal to $3\epsilon/(2\epsilon+1)$. To account for microscopic interactions, it was assumed that the dielectric constant in the prefactor was position-dependent. A function $\epsilon(r)$ was extracted from the orientational distribution function of water molecules around an ion based on molecular dynamics simulations.[27] Thus, solvent structure on microscopic lengthscales is included, at least in an average sense, through the calculated position-dependent microscopic dielectric constant, $\epsilon(r)$. The method was successful in calculations of the water distribution around frozen water clusters and in estimating the solvation energy of microscopic ions.[26,27] However, the approach may be difficult to generalize precisely due to the assumptions concerning the existence of $\epsilon(r)$. For instance, it was noted that the position-dependent dielectric constant represents the cumulative screening between water and an ion and is not meant to be substituted in the Poisson equation as in continuum electrostatic approximations.[27] Thus, there remains a need for a rigorous and numerically tractable statistical mechanical integral equation theory to describe the solvation of complex biomolecules in polar liquids.

In a previous paper, we presented a numerical solution of the hypernetted chain (HNC) equation in three-dimensional Cartesian coordinates, the 3D-HNC, to obtain the average solvent structure around solid substrates.[29] The goal of the present work is to extend this approach to the case of polar liquids. In Sec. II, an integral equation theory is formulated to describe the solvation of a complex molecule of irregular shape in a polar liquid. The integral equation is derived from the HNC theory for a solute at infinite dilution and it is assumed that the solvent is constituted of hard spheres with embedded dipoles. This model is clearly a gross simplification of liquid water since multipoles of higher order are ignored. However, such an approximation cannot be avoided at this point to obtain a numerically tractable theory. The pure solvent direct correlation function used in the theory is taken from the analytical solution of the mean spherical approximation (MSA) equation. The solvent direct correlation function is separated in terms of long and short range components. This separation leads to a pair of coupled integral equations involving the average density distribution and the polarization density. A particular attention is given to the electrostatic properties due to the polar nature of the solvent. In particular, in Sec. III it is demonstrated rigorously that, in the limit where the size of the solvent particles becomes very small, the present theory reduces to the familiar equation of continuum macroscopic electrostatics.[3,30] A linearized version of the integral equation corresponds to a three-dimensional extension of the familiar MSA equation (3D-MSA). In Sec. IV, a numerical method is described to solve the 3D-MSA integral equation in the Cartesian coordinates using a discrete three-dimensional cubic grid. The solution of the 3D-MSA integral equation yields the average density, the average polarization density, and the average electrostatic potential at the point x,y,z. All spatial convolutions appearing in the integral equation are calculated using three-dimensional numerical fast Fourier transforms. In Sec. V, the method is illustrated with numerical calculations applied to the case of an ion, a water molecule, and a molecule

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

of *N*-methylacetamide immersed in liquid water. In Sec. VI, the paper is concluded with a brief summary of the main results and an outlook for further developments.

II. INTEGRAL EQUATION THEORY

An isolated molecular solute is immersed in a polar liquid of nonpolarizable spherical particles with an embedded point dipole μ in their center. The solute–solvent interaction potential $U(\mathbf{r}_1,\omega_1)$, depending on the position in space $\mathbf{r}_1$ and the dipole orientation ω_1 (specified in terms of the spherical coordinates θ and ϕ), is composed of two parts. First, the solvent particles are expelled from the space occupied by the solute due to a harsh repulsive potential, $u^s(\mathbf{r}_1)$, arising from core exclusions. In addition, the orientation of the solvent dipoles is affected by the bare electrostatic potential, $\psi^{(0)}(\mathbf{r}_1)$, arising from the solute charge distribution. The total perturbing solute–solvent potential is given by,

$$U(\mathbf{r}_1,\omega_1)=u^s(\mathbf{r}_1)+\mu(\omega_1)\cdot\nabla_1\psi^{(0)}(\mathbf{r}_1), \tag{1}$$

where ∇_1 corresponds to differentiation with respect to $\mathbf{r}_1$. In computer simulations based on detailed atomic models,[1,2] the potentials $u^s(\mathbf{r}_1)$ and $\psi^{(0)}(\mathbf{r}_1)$ are generally constructed as a sum over radially symmetric functions such as Lennard-Jones 6-12 potentials,

$$u^s(\mathbf{r}_1)=\sum_n u_n^{6\text{-}12}(|\mathbf{r}_1-\mathbf{r}_n|), \tag{2}$$

and Coulomb potentials,

$$\psi^{(0)}(\mathbf{r}_1)=\sum_n \frac{q_n}{|\mathbf{r}_1-\mathbf{r}_n|}. \tag{3}$$

Due to the presence of the solute, the average structure of the polar liquid is perturbed from a uniform density $\overline{\rho}$. The full distribution function $\langle\rho(\mathbf{r}_1,\omega_1)\rangle$ provides all information about the average structure of the perturbed solvent. In characterizing the perturbed solvent structure around the solute, two reductions of the full distribution function are of particular interest. They are the average solvent density,

$$\langle\rho(\mathbf{r}_1)\rangle=\int \frac{d\omega_1}{4\pi}\langle\rho(\mathbf{r}_1,\omega_1)\rangle, \tag{4}$$

and the average polarization density,

$$\mathbf{P}(\mathbf{r}_1)=\int \frac{d\omega_1}{4\pi}\mu(\omega_1)\langle\rho(\mathbf{r}_1,\omega_1)\rangle, \tag{5}$$

where $d\omega_1=d\phi_1 d\theta_1\sin(\theta_1)$. The electrostatic potential due to the solvent (i.e., the reaction field) is,

$$\psi^{(\mathrm{rf})}(\mathbf{r}_1)=\int d\mathbf{r}_2\frac{d\omega_2}{4\pi}\mu(\omega_2)\cdot\left(\nabla_2\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\right)\times\langle\rho(\mathbf{r}_2,\omega_2)\rangle, \tag{6}$$

which, after integration, can be expressed in terms of the solvent polarization density,[30]

$$\psi^{(\mathrm{rf})}(\mathbf{r}_1)=\int d\mathbf{r}_2\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}[-\nabla_2\cdot\mathbf{P}(\mathbf{r}_2)]. \tag{7}$$

It follows that the total electrostatic potential at point $\mathbf{r}_1$

$$\psi^{\mathrm{tot}}(\mathbf{r}_1)=\psi^{(0)}(\mathbf{r}_1)+\psi^{(\mathrm{rf})}(\mathbf{r}_1), \tag{8}$$

obeys the Poisson equation,

$$\nabla_1^2\psi^{\mathrm{tot}}(\mathbf{r}_1)=-4\pi\rho^{(0)}_{\mathrm{elec}}(\mathbf{r}_1)+4\pi\nabla_1\cdot\mathbf{P}(\mathbf{r}_1), \tag{9}$$

where $\rho^{(0)}_{\mathrm{elec}}$ is the charge distribution of the solute.

The aim of a statistical mechanical theory is to determine the average distribution function $\langle\rho(\mathbf{r}_1,\omega_1)\rangle$, in response to the perturbing potential $U(\mathbf{r}_1,\omega_1)$. In the present theory, it is assumed that the distribution function of the polar liquid is given by the HNC integral equation for a single solute in the infinite dilution limit,[11]

$$\langle\rho(\mathbf{r}_1,\omega_1)\rangle=\overline{\rho}\exp\left[-\beta U(\mathbf{r}_1,\omega_1)+\int d\mathbf{r}_2\frac{d\omega_2}{4\pi}c(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2)\Delta\rho(\mathbf{r}_2,\omega_2)\right], \tag{10}$$

where $\beta=1/k_BT$ is the reciprocal of the thermal energy, $c(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2)$ is the direct correlation function of the unperturbed liquid, and $\Delta\rho(\mathbf{r}_2,\omega_2)$ is the deviation from the bulk density, $\langle\rho(\mathbf{r}_2,\omega_2)\rangle-\overline{\rho}$. To make progress and develop a numerically tractable theory, it is assumed that the structure of the unperturbed liquid is given by the MSA. For a liquid of dipolar hard spheres with diameter σ the direct correlation function is,[31]

$$c(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2)=\begin{cases}c^{000}(r_{12})+[\mathbf{e}_1\cdot\mathbf{e}_2]c^{110}(r_{12})+[\mathbf{e}_1\cdot\mathbf{T}(\mathbf{r}_{12})\cdot\mathbf{e}_2]c^{112}(r_{12}) & \text{for } r_{12}\leqslant\sigma\\ \beta[\mu(\omega_1)\cdot\mathbf{T}(\mathbf{r}_{12})\cdot\mu(\omega_2)](r_{12})^{-3} & \text{for } r_{12}\geqslant\sigma,\end{cases} \tag{11}$$

where $\mathbf{r}_{12}=\mathbf{r}_2-\mathbf{r}_1$, $r_{12}=|\mathbf{r}_{12}|$, $\mathbf{e}_i=\mu(\omega_i)/\mu$ is a unit vector directed along the dipole moment, and $\mathbf{T}=[3(\mathbf{r}_{12}/r_{12}):(\mathbf{r}_{12}/r_{12})-\mathbf{I}]$ is a traceless tensor of rank two ($\mathbf{I}$ is the unit tensor). In the MSA theory, the functions c^{mnl} are polynomials depending only on the distance r_{12} and vanishing identically for a distance lager than the diameter σ of a solvent particle.[31] The three functions, c^{000}, c^{110}, and c^{112}, are shown in Figs. 1 and 2. The correlation functions are discontinuous at $r=\sigma$, and are identically equal to zero for r larger than σ. The parameters of the dipolar solvent, which were chosen to best reproduce the properties of liquid water at 300 K, are given in Sec. IV and in the caption to Fig. 1.

The present model and the physical assumptions describing the average solvent structure around the solute are char-

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

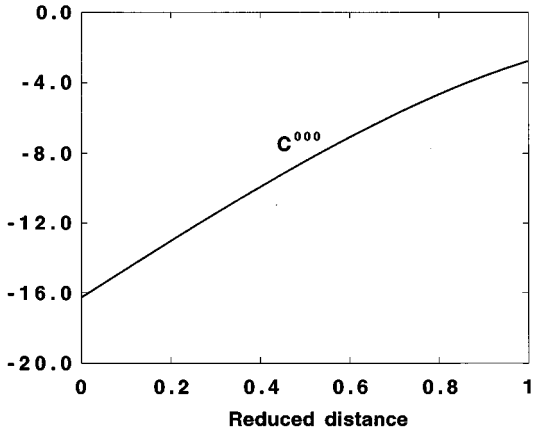

FIG. 1. Direct correlation function of of the liquid constituted of hard spheres with embedded dipole in the MSA approximation. The function $c^{000}(r)$ in Eq. (11) is shown as a function of the reduced distance r/σ, where σ is the diameter of the solvent hard sphere. The function is discontinuous at the point $r=\sigma$, and is identically equal to zero for $r\geqslant\sigma$. The parameters of the solvent were adjusted to represent liquid water at 300 K and at a density 0.0334 Å^{-3} based on the analytic solution to the MSA equation (Ref. 31). The hard sphere diameter is 2.7165 Å to reproduce the experimental value of isothermal compressibility of water. The embedded dipolar moment of the solvent molecule is 2.23943 D to yield the bulk dielectric constant.

acterized by the solute–solvent interaction U, the HNC equation, and the MSA direct correlation function. In order to reveal the influence of the dipole–dipole electrostatic interactions between the solvent particles it is useful to express the direct correlation function of the pure liquid in terms of a long range component ϕ and a short range component c^s, i.e.,[11]

$$c(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2)\equiv\phi(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2)+c^s(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2). \tag{12}$$

As long as the total direct correlation function of the pure liquid is given by Eq. (11), the separation in terms of a long and a short range component is a formal manipulation which does not affect the final result. Thus, the choice of ϕ is not unique and must be guided by convenience. Traditionally,

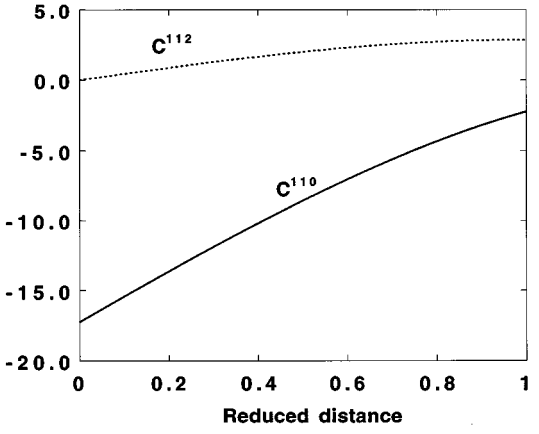

FIG. 2. The $c^{110}(r)$ and $c^{112}(r)$ correlation functions in Eq. (11) shown as a function of the reduced distance r/σ. The functions are discontinuous at the point $r=\sigma$, and are identically equal to zero for $r\geqslant\sigma$. See also Fig. 1.

the function ϕ is identified with the dipole–dipole interactions,[32,33] $\beta\mu(\omega_1)\cdot\mathbf{T}\cdot\mu(\omega_2)\ (r_{12})^{-3}$, for all distances r_{12}. However, at small distances such a choice leads to singularities that are difficult to treat accurately in numerical calculations. To avoid the singularities the following construction was adopted,

$$\phi(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2)=\begin{cases}-\beta[\mu(\omega_1)\cdot\mu(\omega_2)](4\pi/3v) & \text{for}\quad r_{12}\leqslant\sigma\\ \beta[\mu(\omega_1)\cdot\mathbf{T}(\mathbf{r}_{12})\cdot\mu(\omega_2)](r_{12})^{-3} & \text{for}\quad r_{12}\geqslant\sigma.\end{cases} \tag{13}$$

The function ϕ in Eq. (13) represents the interaction between a point dipole $\mu(\omega_1)$ and a uniformly polarized sphere of radius σ centered at $\mathbf{r}_2$. That is,

$$\phi(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2)=-\beta\mu(\omega_1)\cdot\nabla_1\left[\int d\mathbf{r}_3\,\frac{\mu(\omega_2)\cdot\mathbf{r}_{13}}{r_{13}^3}\,\frac{1}{v}\,\Theta(r_{23}-\sigma)\right], \tag{14}$$

where v is the volume of the sphere $4\pi\sigma^3/3$, and $\Theta(r_{23}-\sigma)$ is a Heaviside function equal to one for r_{23} smaller than σ, and equal to zero otherwise. The integral in the square brackets in Eq. (14) can be recognized as the electrostatic potential caused by the uniform spherical polarization density $\Theta\mu/v$ mentioned above, which can be recognized as the convolution of the Θ-function with the electrostatic potential produced by one point dipole, $\mu\cdot\mathbf{r}/r^3$. From Eqs. (11), (12), and (13), the short range direct correlation function is,

$$\begin{aligned}c^s(\mathbf{r}_1,\omega_1,\mathbf{r}_2,\omega_2)=c^{000}(r_{12})&+[\mathbf{e}_1\cdot\mathbf{e}_2]c^{110}(r_{12})\\&+[\mathbf{e}_1\cdot\mathbf{T}(\mathbf{r}_{12})\cdot\mathbf{e}_2]c^{112}(r_{12})\\&+\beta[\mu(\omega_1)\cdot\mu(\omega_2)](4\pi/3v)\Theta(r_{12}-\sigma)\end{aligned} \tag{15}$$

for r_{12} smaller than σ, and zero otherwise. After substitution of Eqs. (13) and (15) in Eq. (10), the solvent distribution function is,

$$\begin{aligned}\langle\rho(\mathbf{r}_1,\omega_1)\rangle=\overline{\rho}\,&\exp[-\beta u^s(\mathbf{r}_1)+c^{000}\star\Delta\rho(\mathbf{r}_1)]\\&\times\exp[\beta\mu(\omega_1)\cdot\mathbf{A}(\mathbf{r}_1)],\end{aligned} \tag{16}$$

where the symbol $\star$ represents a three-dimensional spatial convolution, and $\Delta\rho(\mathbf{r}_1)$ is the deviation of the average density from the bulk density $\langle\rho(\mathbf{r}_1)\rangle-\overline{\rho}$. The form of the distribution function, expressed as a product of position-and orientation-dependent parts, is similar to that proposed by Liu and Ichiye.[26] In particular, the orientation-dependent distribution corresponds to the Boltzmann factor of a dipole in an effective orienting field. In the present theory, the effective orienting field $\mathbf{A}$ is a position-dependent vector defined in terms of the electrostatic potentials $\psi^{(0)}$ and $\psi^{(\mathrm{rf})}$, and of the polarization density $\mathbf{P}$,

$$\mathbf{A}(\mathbf{r}_1)=-\nabla_1\left[\psi^{(0)}(\mathbf{r}_1)+\frac{1}{v}\,\Theta\star\psi^{(\mathrm{rf})}(\mathbf{r}_1)\right]+\mathbf{C}\star\mathbf{P}(\mathbf{r}_1). \tag{17}$$

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

The quantity $\mathbf{C}(\mathbf{r}_{12})$ is a tensor of rank two constructed from a Heaviside Θ-function and the short range direct correlation functions,

$$\mathbf{C}(\mathbf{r}_{12})=\left[\frac{c^{110}(r_{12})}{\beta\mu^2}+\frac{4\pi}{3v}\,\Theta(r_{12}-\sigma)\right]\mathbf{I}+\frac{c^{112}(r_{12})}{\beta\mu^2}\,\mathbf{T}(\mathbf{r}_{12}). \tag{18}$$

Its components, shown in Fig. 2, are short range functions equal to zero for r_{12} larger than the diameter σ of a solvent particle. All the spatial convolutions appearing in the above integral equations can be easily expressed using Fourier transforms [defined as $\hat{f}(\mathbf{k})=\int d\mathbf{k} f(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r})$]. In particular, the Fourier transform of the vector $\mathbf{A}(\mathbf{r})$ is,

$$\hat{\mathbf{A}}(\mathbf{k})=-i\mathbf{k}\left[\hat{\psi}^{(0)}(\mathbf{k})+\frac{1}{v}\,\hat{\Theta}(\mathbf{k})\hat{\psi}^{(\mathrm{rf})}(\mathbf{k})\right]+\hat{\mathbf{C}}(\mathbf{k})\hat{\mathbf{P}}(\mathbf{k}), \tag{19}$$

where,

$$\hat{\psi}^{(0)}(\mathbf{k})=\frac{4\pi}{k^2}\sum_n q_n\exp(-i\mathbf{k}\cdot\mathbf{r}_n), \tag{20}$$

from Eq. (3), and,

$$\hat{\psi}^{(\mathrm{rf})}(\mathbf{k})=-\frac{4\pi}{k^2}\,i\mathbf{k}\cdot\hat{\mathbf{P}}(\mathbf{k}), \tag{21}$$

from Eq. (7).

The solvent structure can also be expressed in terms of the average density and the polarization density since the functions in Eq. (16) involve only these two reduced distributions. Integration of Eq. (16) over ω_1 yields the averaged density,

$$\langle\rho(\mathbf{r}_1)\rangle=\overline{\rho}\exp[-\beta u^s(\mathbf{r}_1)+c^{000}\star\Delta\rho(\mathbf{r}_1)]\,\frac{\sinh(\beta|\mathbf{A}|\mu)}{\beta|\mathbf{A}|\mu} \tag{22}$$

and the average polarization density,

$$\mathbf{P}(\mathbf{r}_1)=\mu\langle\rho(\mathbf{r}_1)\rangle\mathscr{L}(\beta|\mathbf{A}|\mu)\,\frac{\mathbf{A}}{|\mathbf{A}|}, \tag{23}$$

where $\mathscr{L}(x)=1/\tanh(x)-1/x$ is the Langevin response function.[8,34] Equations (16), (22), and (23), which determine completely the average solvent structure in the presence of the solute, are the main result of this paper.

III. CONTINUUM MACROSCOPIC ELECTROSTATICS LIMIT

It is interesting to examine under what conditions the familiar law of continuum macroscopic electrostatic is recovered by the present integral equation theory.[30] The constitutive relation of macroscopic electrostatics is,

$$\mathbf{P}(\mathbf{r})=\left[\frac{\epsilon(\mathbf{r})-1}{4\pi}\right]\mathbf{E}^{\mathrm{tot}}(\mathbf{r}), \tag{24}$$

where $\mathbf{E}^{\mathrm{tot}}=-\nabla\psi^{\mathrm{tot}}$ is the total electrostatic field, (the label on $\mathbf{r}_1$ will be omitted in this section for simplicity). According to Eq. (24), the polarization density at point $\mathbf{r}$ obeys a *linear* relation involving the *local* electrostatic field. In contrast, the polarization density in the integral equation theory is given, through Eq. (23), by a *nonlinear* Langevin response function with *nonlocal* spatial convolutions. Assuming that the strength of the electrostatic field is small relative to the thermal energy and that saturation effects are negligible, it is possible to linearize Eqs. (22) and (23) with respect to $\beta|\mathbf{A}|\mu$ when $\beta|\mathbf{A}|\mu\ll 1$. The linearized equations are,

$$\langle\rho(\mathbf{r})\rangle=\overline{\rho}\exp[-\beta u^s(\mathbf{r})+c^{000}\star\Delta\rho(\mathbf{r})], \tag{25}$$

and

$$\mathbf{P}(\mathbf{r})=\langle\rho(\mathbf{r})\rangle\,\frac{\beta\mu^2}{3}\,\mathbf{A}(\mathbf{r}), \tag{26}$$

where $\mathbf{A}$ is given by Eq. (17). Electrostriction effects are neglected in the linear approximation since the solvent density, given by Eq. (25), is independent of the electrostatic potential. The above equations have a nonlocal character because they involve the convolutions $\Theta\star\psi^{(\mathrm{rf})}(\mathbf{r})$ and $\mathbf{C}\star\mathbf{P}(\mathbf{r})$. The convolution $\Theta\star\psi^{(\mathrm{rf})}(\mathbf{r})/v$ in Eq. (17) results in a spatial averaging of the electrostatic potential over a small region, an operation that is reminiscent of the spatial coarse graining that is often invoked in the derivation of the macroscopic constitutive relation.[30,34] Assuming that the variations of the solvent electrostatic potential and polarization density are negligible on the length scale of the diameter σ of a solvent particle (corresponding to the spatial range of the Θ-function and the tensor $\mathbf{C}$), the convolutions can be approximated by local relations. That is,

$$\lim_{\sigma\to\text{small}}\ \psi^{(0)}(\mathbf{r})+\frac{1}{v}\,\Theta\star\psi^{(\mathrm{rf})}(\mathbf{r})\approx\psi^{\mathrm{tot}}(\mathbf{r}), \tag{27}$$

and,

$$\lim_{\sigma\to\text{small}}\ \mathbf{C}\star\mathbf{P}(\mathbf{r})\approx\hat{\mathbf{C}}(0)\mathbf{P}(\mathbf{r})=[\alpha\mathbf{I}]\mathbf{P}(\mathbf{r})=\alpha\mathbf{P}(\mathbf{r}), \tag{28}$$

where the constant α is,

$$\alpha=\frac{4\pi}{3}+\frac{\hat{c}^{110}(0)}{\beta\mu^2}, \tag{29}$$

from Eq. (18). The limiting behavior of the components of the tensor $\hat{\mathbf{C}}$ as $k\to 0$ is illustrated in Fig. 3; in this limit the tensor is proportional to $\alpha\mathbf{I}$, with $\alpha=0.539$. A linear and local relation exists, then, between the polarization density and the total electrostatic field,

$$\mathbf{P}(\mathbf{r})=\langle\rho(\mathbf{r})\rangle\,\frac{\beta\mu^2}{3}\,[\mathbf{E}^{\mathrm{tot}}(\mathbf{r})+\alpha\mathbf{P}(\mathbf{r})]. \tag{30}$$

By analogy with the constitutive relation of macroscopic electrostatics,[30] it is possible to identify a position-dependent dielectric constant $\epsilon(\mathbf{r})$. Since,

$$\mathbf{P}(\mathbf{r})=\left[\frac{\langle\rho(\mathbf{r})\rangle\beta\mu^2}{3-\langle\rho(\mathbf{r})\rangle\beta\mu^2\alpha}\right]\mathbf{E}^{\mathrm{tot}}(\mathbf{r}), \tag{31}$$

from Eq. (30), it follows that,

$$\epsilon(\mathbf{r})=1+\frac{4\pi\langle\rho(\mathbf{r})\rangle\beta\mu^2}{3-\langle\rho(\mathbf{r})\rangle\beta\mu^2\alpha}. \tag{32}$$

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

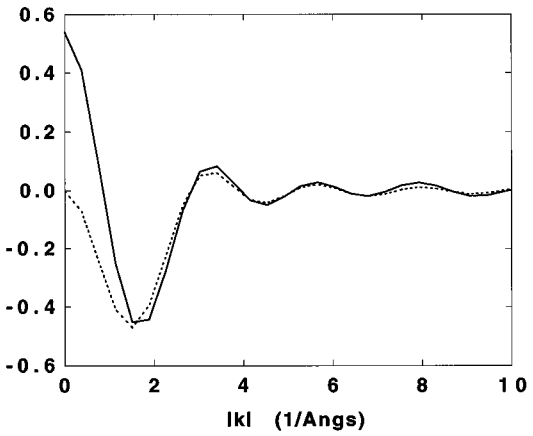

FIG. 3. The diagonal (solid line) and nondiagonal (dashed line) components of the tensor $\hat{\mathbf{C}}(\mathbf{k})$ from Eq. (19) as a function of the amplitude of Fourier vector $|\mathbf{k}|$, for equivalent components $k_x=k_y=k_z$. In the $k\to 0$ limit the diagonal components are $\alpha=0.539$ and the nondiagonal components are zero for any $\mathbf{k}$. By construction, the tensor is always diagonal when two of the components of the vector $\mathbf{k}$ are equal to zero.

It should be stressed that the position-dependent dielectric constant in Eq. (32) may not always be valid. In particular, if the quantity $\langle\rho(\mathbf{r})\rangle\beta\mu^2\alpha>3$ at some position $\mathbf{r}$ the system is beyond the Curie point and is in a ferroelectric state.[34] This situation may occur at the solute–solvent contact, where the solvent density $\langle\rho(\mathbf{r})\rangle$ is usually larger. In the case of biomolecules solvated by water, such ferroelectric state does not reflect the results obtained from detailed simulations and is most certainly unphysical.[1] Using Eq. (29), the dielectric constant of the bulk liquid can be expressed as,

$$\epsilon_{\text{bulk}}=1+\frac{4\pi\overline{\rho}\beta\mu^2}{3-\overline{\rho}\beta\mu^2\alpha}=\frac{1-\overline{\rho}\hat{c}^{110}(0)/3+2y}{1-\overline{\rho}\hat{c}^{110}(0)/3-y},\tag{33}$$

where $y=4\pi\beta\overline{\rho}\overline{\mu}^2/9$ is a dimensionless constant. This expression for the dielectric constant of a polar liquid is in agreement with previous statistical mechanical results.[23,30,32,33] Since the solvent density $\langle\rho(\mathbf{r})\rangle$ is usually dominated by the short range repulsive potential $u^s(\mathbf{r})$, the simple two-valued dielectric function $\epsilon(\mathbf{r})$,

$$\epsilon(\mathbf{r})=\begin{cases}1 & \text{if } u^s(\mathbf{r})=\infty\\ \epsilon_{\text{bulk}} & \text{if } u^s(\mathbf{r})=0,\end{cases}\tag{34}$$

is often used to describe the solvation of biomolecular solutes in water.[3] This convention can be understood on the basis of Eqs. (31) and (32). Finally, following from Eqs. (9), (24), (31), and (32),

$$\nabla^2\psi^{\text{tot}}(\mathbf{r})=-4\pi\rho^{(0)}_{\text{elec}}(\mathbf{r})-\nabla\cdot\{[\epsilon(\mathbf{r})-1]\nabla\psi^{\text{tot}}(\mathbf{r})\}=-4\pi\rho^{(0)}_{\text{elec}}(\mathbf{r})-\nabla\cdot[\epsilon(\mathbf{r})\nabla\psi^{\text{tot}}(\mathbf{r})]+\nabla^2\psi^{\text{tot}}(\mathbf{r}),\tag{35}$$

and the total electrostatic potential obeys the Poisson equation of macroscopic electrostatics for continuum media,[30]

$$\nabla\cdot[\epsilon(\mathbf{r})\nabla\psi^{\text{tot}}(\mathbf{r})]=-4\pi\rho^{(0)}_{\text{elec}}(\mathbf{r}).\tag{36}$$

The present analysis can be related to previous discussions of continuum macroscopic electrostatics.[28,30,34,35] Traditionally, the dielectric response of polar liquids has been described in terms of an isolated solvent dipole under the influence of an effective orienting field $\mathbf{E}$.[28,35] A fundamental step in those theories is the determination of the orienting field from the total macroscopic electrostatic field $\mathbf{E}^{\text{tot}}$. In his theory of dielectrics, Debye approximated the orienting field as the total electrostatic field augmented by the Clausius–Mosotti internal field $(4\pi/3)\mathbf{P}$ corresponding to the self-field present at the center of a spherical region of uniformly polarized material.[35] Onsager showed that this approximation is incorrect because the orienting field should be separated into a reaction field, pointing in the instantaneous direction of the solvent dipole, and a cavity field, acting in the center of spherical empty cavity cut out of a uniform dielectric continuum.[28] Since the reaction field cannot exert a net torque on the solvent molecule, only the cavity field contributes to the orienting field. In the present theory, the effective orienting field acting on a solvent molecule is the position-dependent vector $\mathbf{A}$, which is defined in terms of $\psi^{(0)}$, $\psi^{(\text{rf})}$, and $\mathbf{P}$, from a nonlocal relation involving spatial convolutions. The different orienting fields $\mathbf{E}'$ may be summarized as,

$$\mathbf{E}'=\begin{cases}\mathbf{A} & \text{nonlocal theory}\\ \mathbf{E}^{\text{tot}}+[(4\pi/3)+\hat{c}^{110}(0)/(\beta\mu^2)]\mathbf{P} & \text{local approximation}\\ \mathbf{E}^{\text{tot}}+(4\pi/3)\mathbf{P} & \text{Clausius–Mosotti}\\ \mathbf{E}^{\text{tot}}[3\epsilon/(2\epsilon+1)] & \text{Onsager.}\end{cases}\tag{37}$$

In local theories, the orienting field can be expressed as the total electrostatic field multiplied by a prefactor. For liquid water, the Clausius–Mosotti and Onsager prefactor are 27.3 and 1.49, respectively, while it is 4.67 in the local approximation to the present theory. For instance, it can be seen that the Clausius–Mosotti approximation corresponds to the complete neglect of the short range direct correlation function c^{110}. In the present analysis, the term $(4\pi/3)\mathbf{P}$ arises directly from the fact that the long range function ϕ was subtracted for r smaller than σ in the short range direct correlation function c^s [see Eqs. (11), (12), and (14)]. Recently, the concept of cavity field developed by Onsager was extended by assuming the existence of a distance-dependent dielectric constant $\epsilon(r)$ to describe the distribution of water molecules in the neighborhood of a chloride ion.[27] Based on the present analysis, the significance of the function $\epsilon(r)$ is unclear since the effective orienting field $\mathbf{A}$ is a complex function involving non local convolutions. It can be seen from Eq. (31) that replacing the non local convolutions by local relations implies that the solvent polarization is directed parallel to the local electrostatic field. This could be particularly important in the neighborhood of the solute where the orienting field $\mathbf{A}$ cannot be approximated by local relations and should be determined through the nonlocal Eq. (17).

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

It is possible to relate the present theory to previous approaches. For example, approximating the convolutions, as in Eqs. (27) and (28), by local relations, while the Langevin response function is kept in Eq. (23), yields a local nonlinear equation for the polarization density. Such a theory corresponds to the PDLD approach of Warshel.[8] However, in a nonlinear theory the average solvent density should also be coupled to the field through Eq. (22) according to the present derivation. In contrast, the PDLD is concerned only with the solvent polarization and the average density is independent of electrostatic field. The nonlinear coupling between the electrostatic field and the solvent density gives rise to electrostriction effects. Such variations in the average density may be negligible if the solvent is almost incompressible. This physical situation is realized when the quantity $c^{000}(0)$ is very large and negative, since the isothermal compressibility χ_T is equal to $\beta/[\overline{\rho}-\overline{\rho}^2 c^{000}(0)]$.[11] Lastly, it is possible to establish a link with the mean-field theory of Gruen and Marčelja.[9,10] Their mean-field equations are recovered from the present theory if the nonlocal convolutions are approximated, to second order, through a gradient expansion of the polarization density. The microscopic correlation length appearing in their theory is then related to moments of the short range direct correlation function.

IV. APPLICATIONS AND COMPUTATIONAL DETAILS

To calculate the density and polarization around the solutes, two nonlinear coupled integral Eqs. (22) and (23) must be solved self-consistently. Based on previous experience with the 3D-HNC integral equations,[29] numerical solution of such nonlinear coupled integral is expected to be very difficult. As a first attempt, a linearized version of the theory was considered. In the linearized theory, the two integral Eqs. (25) and (26) are decoupled. Thus, the density can be calculated, in a first step, from the short range solute–solvent repulsive potential $u^s(\mathbf{r}_1)$, and the polarization can be calculated, in a second step, from the solute charge distribution and the solvent density. Equation (25) corresponds to a HNC integral equation for the density. Further linearization with respect to $c^{000}\star\Delta\rho(\mathbf{r}_1)$ yields a Percus–Yevick (PY) integral equation,[11]

$$\langle\rho(\mathbf{r}_1)\rangle=\overline{\rho}\exp[-\beta u^s(\mathbf{r}_1)]\times[1+c^{000}\star\Delta\rho(\mathbf{r}_1)]. \quad (38)$$

Both the HNC Eq. (25), or the PY Eq. (38), can be solved to calculate the solvent density, which is then used subsequently in Eq. (26) to calculate the polarization density.

The completely linearized theory based on Eqs. (26) and (38), was chosen in the present calculations. This linearized theory corresponds to a three dimensional extension of the familiar MSA theory for a solute of arbitrary shape at infinite dilution immersed in a polar liquid, i.e., a 3D-MSA. Of course, there are well-known limitations with such a theory. In particular, the criterion of the linearization can be violated in the case of strongly charged solutes. Nevertheless, attempts to solve the linearized 3D-MSA integral equation are worthwhile. Despite its imperfections, the MSA theory usually provides meaningful results, and for this reason, has played an extremely important role in clarifying the essential concepts used to understand the structure of polar liquids.[11,31] Moreover, the present efforts provide a first exploration of the difficulties that are encountered in attempting to calculate the average solvent structure around a molecular solute with irregular shape using such an integral equation.

Three different solutes were considered to illustrate the theory; a univalent positive ion, a water molecule, and a *N*-methylacetamide (NMA) molecule in the *trans* conformation. The geometry, the Lennard-Jones (LJ) parameters and the partial charges of the solute NMA molecule were the same as used previously by Jorgensen for the *trans* conformation.[36] The geometry, and partial charges of the solute water molecule were taken from the TIP3P model.[37] The LJ parameters, $\sigma=2.7165$ Å and $\epsilon=0.152\,073$ kcal/mol, were used for the sphere, representing the ion and the solute water and for the solvent spheres to calculate the solute–solvent short range potential $u^s(\mathbf{r}_1)$ in Eq. (2). The parameters of the solvent were adjusted to represent liquid water at 300 K and at density 0.0334 Å^{-3} based on the analytic solution to the MSA equation.[31] The hard sphere diameter was adjusted to reproduce the experimental value of the isothermal compressibility $\chi_T=-(\partial \ln V/\partial p)_T$ of 45×10^{-6} atm^{-1} and the dipolar moment of the solvent molecule was adjusted to reproduce the bulk dielectric constant of 80 (see Ref. 38 for experimental values). These empirical adjustments yield a hard sphere of 2.7165 Å and a dipolar moment of 2.23943 D.[31] A radius of 2.7165/2 Å was used with the standard combination rule to calculate the solvent–solute Lennard-Jones interactions.

The solvent density was calculated from Eq. (38) using the method described previously in the case of neutral substrates immersed in a liquid of spherical particles.[29] Briefly, the numerical solution of Eq. (38) in Cartesian coordinates involves a mapping of the average density $\langle\rho(x,y,z)\rangle$ and of the potential $u^s(x,y,z)$ on a discrete three-dimensional cubic grid, e.g., $\langle\rho(x,y,z)\rangle\rightarrow\langle\rho(i,j,k)\rangle$ and $u^s(x,y,z)\rightarrow u^s(i,j,k)$. A discrete grid of 128 points with a spacing of 0.225 Å corresponding to an elementary cell of 28.8 Å was used in the calculations. The spatial convolution $c^{000}\star\Delta\rho(\mathbf{r})$ in Eq. (38) was calculated using a numerical three-dimensional fast Fourier transform (FFT) procedure.[39] The convolution was calculated directly, without zero-padding. This corresponds to a periodic system in the x, y, and z directions. An iterative scheme in terms of the deviation from the bulk density Δp with simple mixing was used to solve the Eq. (38). The nth iteration is obtained from,

$$\begin{aligned}\Delta\rho^{(n+1)}(\mathbf{r}_1)=\lambda\{&\overline{\rho}\exp[-\beta u^s(\mathbf{r}_1)]\\&\times[1+c^{000}\star\Delta\rho^{(n)}(\mathbf{r}_1)]-1\}\\&+(1-\lambda)\Delta\rho^{(n)}(\mathbf{r}_1),\end{aligned} \quad (39)$$

where λ is a mixing factor. Each iteration cycle involves the calculation of two three-dimensional FFT (one forward and one backward). A uniform bulk density $\overline{\rho}$ was taken as the initial guess for the iteration. To stabilize the solution and

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

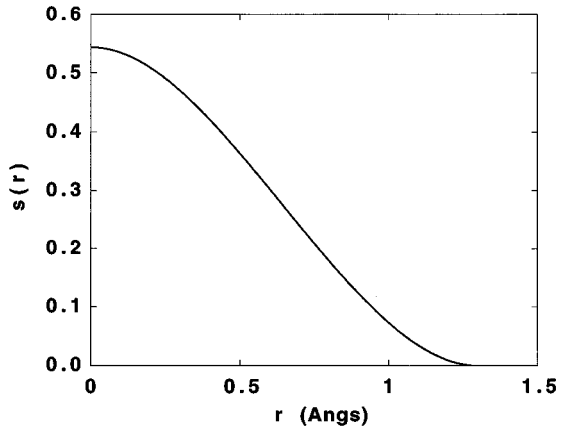

FIG. 4. The function $s(\mathbf{r})$ used in the calculation of the solute charge density in Eq. (41).

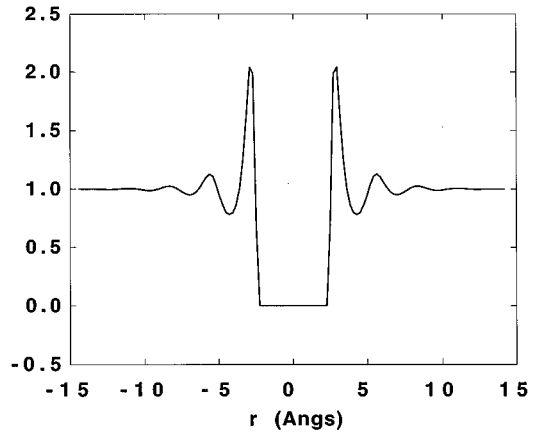

FIG. 5. Normalized density distribution $\langle\rho(\mathbf{r})\rangle/\overline{\rho}$ around the spherical LJ particle representing the positively charged ion or the solute TIP3P water molecule.

reach convergence, it was necessary to initiate the cycle by performing 10 iterations with a mixing factor of 0.05 using Eq. (39). Then, 50 to 70 iterations with a mixing factor of 0.1 were used to reach convergence. It was observed that the calculations would diverge with a larger mixing factor in some cases. The iteration cycle was repeated until the maximum difference of $|\rho^{(n+1)}-\rho^{(n)}|/\lambda\overline{\rho}$ over the whole grid was less than 4×10^{-4}.

After the solvent density has been calculated, the polarization was obtained from Eq. (26). The three components of polarization density, i.e., $P_x(x,y,z)$, $P_y(x,y,z)$, and $P_x(x,y,z)$, were mapped onto a cubic grid. The Heaviside $\hat{\Theta}$-function, as well as the six components of the $\hat{\mathbf{C}}$ tensor; $\hat{\mathbf{C}}_{xx}$, $\hat{\mathbf{C}}_{yy}$, $\hat{\mathbf{C}}_{zz}$, $\hat{\mathbf{C}}_{xy}$, $\hat{\mathbf{C}}_{xz}$, and $\hat{\mathbf{C}}_{yz}$, appearing in Eq. (19) were mapped onto the grid in term of their Fourier modes. A special treatment was used for the Fourier transform of the bare Coulomb potential in Eq. (20). Due to the finite truncation to a maximum wave vector in $\mathbf{k}$-space, the plane waves corresponding to the Fourier transform of the discrete point charges lead to undesirable aliasing in the FFT procedure. The aliasing can be avoided, without affecting the physical results, by using a smoother charge distribution with smaller Fourier components at large values of $\mathbf{k}$. The point charge distribution was smoothed using the polynomial function $s(r)$ given by,

$$s(r)=\frac{3}{4\pi\sigma_s^3}\left[10\left(\frac{r}{\sigma_s}\right)^3-15\left(\frac{r}{\sigma_s}\right)^2+5\right], \tag{40}$$

if r smaller than σ_s, and zero otherwise. The value of σ_s was chosen to be 1.3 Å, which is smaller than the radius of all particles in the system. The function $s(r)$, shown in Fig. 4, is normalized such that $\hat{s}(\mathbf{k}=0)=1$. The bare potential is then,

$$\hat{\psi}^{(0)}(\mathbf{k})=\frac{4\pi}{k^2}\sum_n q_n\hat{s}(\mathbf{k})\exp(-i\mathbf{k}\cdot\mathbf{r}_n). \tag{41}$$

The Fourier transform $\hat{s}(\mathbf{k})$ was calculated using FFT. To solve Eq. (26), a simple iterative scheme with mixing was again used to obtain self-consistency. Starting from the polarization components at the nth iteration, the components of the vector $\hat{\mathbf{A}}(\mathbf{k})$ are calculated using Eq. (19). The polarization density for the $(n+1)$th interaction is then calculated from Eq. (26),

$$\mathbf{P}^{(n+1)}(\mathbf{r})=\lambda\left\{\langle\rho(\mathbf{r})\rangle\frac{\beta\mu^2}{3}\mathbf{A}^{(n)}(\mathbf{r})\right\}+(1-\lambda)\mathbf{P}^{(n)}(\mathbf{r}). \tag{42}$$

Each cycle involves the calculation of six three-dimensional FFT (three forward and three backward). Zero values for all polarization components were taken as the initial guess for the iterational cycle. To stabilize the solution and reach convergence, it was necessary to initiate the cycle by performing 40 iterations with a mixing factor of 0.01 using Eq. (42). Then, about 300 iterations were used to reach convergence. To accelerate convergence, each iteration with a mixing factor of 0.1 was followed by three iterations with a mixing factor of 0.01. The cycle was repeated until the the change in the polarization components P_x, P_y, and P_z summed over the whole grid, $\Sigma_{i=1}^{N^3}|\mathbf{P}^{(n+1)}-\mathbf{P}^{(n)}|$ was less than $\lambda N^3\times10^{-8}$. Equations (20) and (21) diverge at $\mathbf{k}=0$ in Fourier space. To avoid the singularity, all Fourier modes corresponding to $\mathbf{k}=0$ were assumed to be zero. This corresponds to the solution of Eqs. (20) and (21) with spherical conducting boundary conditions at infinity.[40] After convergence was reached, the electrostatic potential $\psi^{(\mathrm{rf})}$ caused by the polarization density was calculated from Eq. (21). The total memory requirements for the calculation of polarization are on the order of 250 megabytes (RAM). Each calculation needed about 10–15 h to converge on an SGI Power Challenge using the iterative scheme described above.

V. RESULTS AND DISCUSSION

To determine the solvent structure around the positive ion, the average solvent density and the polarization density must be calculated. In a first step, the average solvent density around a neutral LJ particle was calculated using Eq. (38). The result is shown in Fig. 5 along one of the grid axis through the center of the particle. Similar calculations have

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

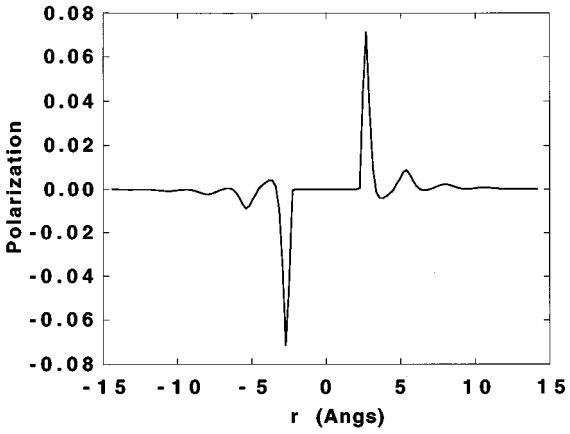

FIG. 6. The projection of the radial polarization density to the axis going through the positive ion. The polarization is radial directed and shown in atomic units $e/\text{Å}^2$ where e is a unit charge.

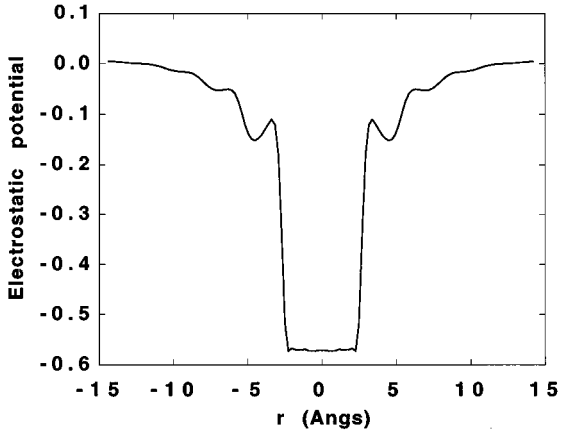

FIG. 7. The electrostatic potential caused by the solvent polarization $\psi^{(\mathrm{rf})}$ defined by Eq. (7) around the positive ion. The potential is in atomic units $e/\text{Å}$ where e is a unit charge.

been obtained previously with the 3D-HNC theory.[29] These calculations have shown that the periodic boundary conditions introduced by the three-dimensional FFT have only a small influence on the average solvent density calculated with the integral equation at a large distance from the solute.[29] The maximum average density is about $2.0\overline{\rho}$ at the contact with the particle. In the bulk region, the density is an oscillating function decreasing in amplitude. The density quickly goes to zero inside the sphere due to the core repulsion.

In a second step, the calculated solvent density was used in Eq. (26) to find the polarization density around the positive spherical ion. The radial component of the polarization density is shown in Fig. 6. The polarization density is an oscillating function with the highest peak about 0.07 $e/\text{Å}^2$ at the distance σ from the ion. Such a polarization density exceeds the maximum possible value according to the solvent density, i.e., $2.0\overline{\rho\mu}=0.031$ e/Å^2. This can be explained by the linearization of the Langevin function in Eq. (23). Examination of the polarization over the whole space shows that the criterion ($\beta|\mathbf{A}|\mu<0.5$) is violated at 2.83% of grid points, where the normalized solvent density is larger than 0.01. The condition $\beta|\mathbf{A}|\mu>2$ is fulfilled at 0.29% of grid points, located exactly at the surface of the ion. To avoid the overestimated polarization density, it is necessary to solve the system of nonlinearized coupled Eqs. (22) and (23). Although this is possible, it is more difficult numerically and efforts are in progress to develop more efficient algorithms for such solution. The character of the solvent structure is not strongly affected by the periodic boundary conditions. The solvent structure around the ion is very similar, but not identical, to previous results based on the MSA.[41] It is likely that the differences are due to the spherical conducting boundary conditions used in Eqs. (20) and (21).[40] Those boundary conditions are necessary because the system is periodic and carries a net charge in each unit cell. Similar considerations are necessary in simulations of periodic charged systems using Ewald summation.[2] The reaction field arising from the polarization, calculated from Eq. (21), is shown in Fig. 7. In the bulk region the reaction field is a complex function exhibiting strong oscillations. In contrast, the reaction field inside the ion core is constant, as it would be according to the continuum macroscopic approximation. This follows from the fact that the solvent is excluded from this region and the system has perfect spherical symmetry. However, the magnitude of the reaction field is modified due to the spherical conducting boundary conditions at infinity.

Using the same method, the solvent structure around a solute water molecule was determined. The average solvent density around the neutral LJ particle, unchanged from the previous calculation, was used to calculate the polarization density around the water molecule. The reaction field due to the solvent polarization was calculated from Eq. (21). In Fig. 8, the calculated reaction field in the plane of the water molecule is shown in a contour plot representation. It is observed

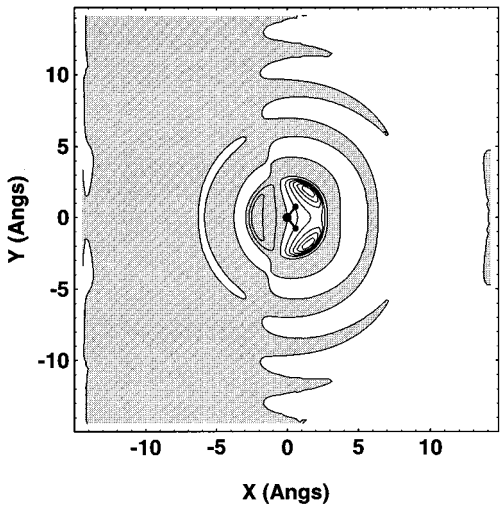

FIG. 8. The electrostatic potential caused by the solvent polarization $\psi^{(\mathrm{rf})}$ defined by Eq. (7) in the plane of the water molecule. The potential in the plane of the water molecule is shown in atomic units $e/\text{Å}$. The contours are plotted in the range from -0.6 to 0.3 with increments of 0.1. Shaded and white areas correspond to regions in which the potential is respectively positive or negative.

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

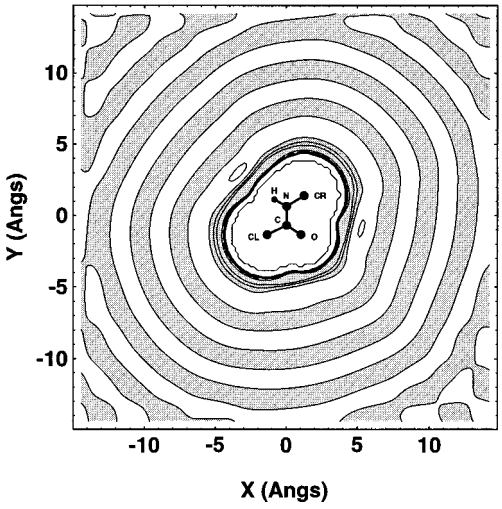

FIG. 9. The contour plot of the normalized density distribution $\langle\rho(\mathbf{r})\rangle/\overline{\rho}$ in the plane of the NMA molecule. The contours are plotted in the range of the normalized density from 0.0 to 2.0 with increments of 0.5. The shaded areas correspond to regions where the density is larger than $\overline{\rho}$. The white areas correspond to regions where the density is less than $\overline{\rho}$; the next white region corresponds to a decrease of the density to a value less than $0.5\overline{\rho}$.

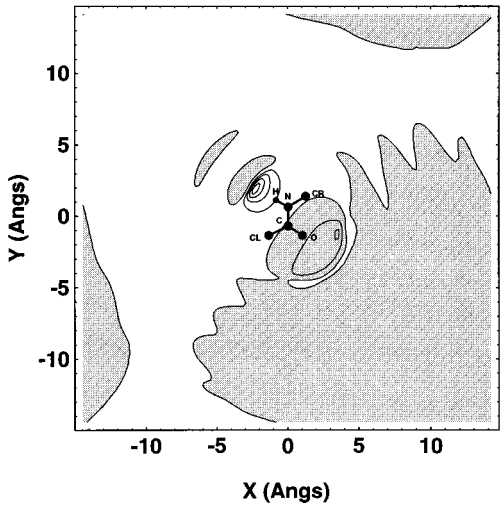

FIG. 10. The electrostatic potential caused by the solvent polarization $\psi^{(\mathrm{rf})}$ defined by Eq. (7) in the plane of the NMA molecule. The potential was calculated in atomic units $e/\text{Å}$. The contours are plotted in the range from -0.6 to 0.3 with increments of 0.1. The regions, where the potential is positive are dashed.

that the potential is positive in the region on the side of the water oxygen, and is negative on the side of the water hydrogens, in accord with the dipolar character of the water molecule charge distribution. However, it is clear that the solvent reaction exhibit strong oscillations that are significantly more complex than in the continuum approximation.

The character of the average solvent structure around the ion and the water molecule results from the spherical symmetry of the solvent-excluded region. The average solvent density is calculated using Eq. (38) around a neutral LJ particle and the result is then used subsequently in Eq. (26) to calculate the polarization density. To illustrate the integral equation in the case of a complex solute with a marked irregular shape, the solvent structure around NMA was calculated. The NMA molecule, taken from the work of Jorgensen, is modeled with 6 LJ functions and 6 partial charges.[36] In a first step, the average solvent density was calculated using Eq. (38). The density in the plane of the molecule is shown in a contour plot representation in Fig. 9. The presence of strong density oscillations spreading far away reflects the solvent structuration due to the solute perturbation. In the plane of the NMA molecule, the density reaches a maximum value of $3.95\overline{\rho}$ near the solute; the maximum density over the whole space around the NMA is $4.98\overline{\rho}$.

The average solvent density was used to calculate the polarization density around the solute NMA molecule based on Eq. (26). The reaction field due to the solvent polarization was calculated from Eq. (21). In Fig. 10, a contour plot of the calculated reaction field in the plane of the NMA molecule is shown. Shaded and white regions correspond respectively to positive or negative values of the potential. In accord with the dipolar character of the charge distribution of the NMA molecule, the potential is positive on the side of the carbonyl oxygen, and is negative on the side of the amide hydrogen. Nevertheless, the characteristic pattern of a solvated dipolar solute is significantly altered by the presence of strong oscillations in the potential due to the solvent structure.

It is interesting to examine one important assumption of continuum electrostatics, that is, whether the polarization density $\mathbf{P}(\mathbf{r})$ is oriented parallel to the total electrostatic field $\mathbf{E}^{\mathrm{tot}}(\mathbf{r})$. In Fig. 11 the probability distribution of the cosine between $\mathbf{P}$ and $\mathbf{E}^{\mathrm{tot}}$ is shown for the various solutes. It is clear that the relative orientation of $\mathbf{P}$ and $\mathbf{E}^{\mathrm{tot}}$ is not unique, as prescribed by continuum electrostatics. Large peaks at $\cos=1$ and $\cos=-1$ are observed, corresponding respectively to parallel and antiparallel orientation of the polarization density relative to the total electrostatic field. In the case

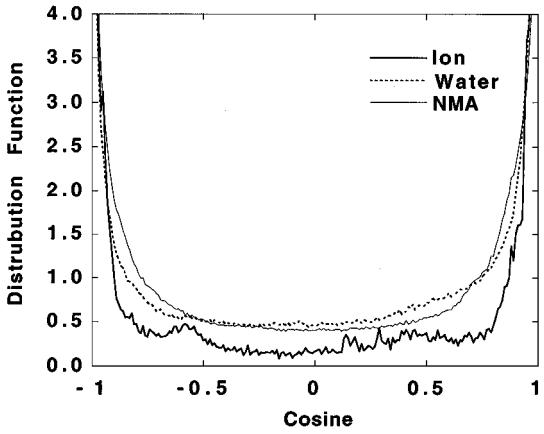

FIG. 11. Distribution histogram of the cosine between the polarization density $\mathbf{P}$ and the electrostatic field $\mathbf{E}$ at all point $\mathbf{r}$ where the polarization is nonzero. The cases of the ion (bold solid line), water molecule (dashed line), and NMA molecule (thin solid line) solutes are shown.

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

of the ion solute, the relationship between the direction of the polarization density and of the electrostatic field can be understood using a simple argument. For a radially symmetric system, direct integration of the Poisson equation yields that $E(r)=e/r^2-4\pi P(r)$ in the radial direction $\hat{\mathbf{r}}$. Therefore, the vectors **E** and **P** point in the same direction only if the magnitude $P(r)$ is smaller than e/r^2 while values of P larger than e/r^2 necessarily result in an electrostatic field that is pointing in the direction opposite to **E**. At the ion-solvent contact the polarization density is expected to reach a value corresponding almost to the maximum saturation of 0.03 (in the present calculation based on a linearized theory, the value is 0.07 at contact). Since such a value of the polarization multiplied by the factor 4π is larger than e/r^2 (equal to 0.1355) at the ion-solvent contact, the polarization and the electrostatic field do not point in the same direction. In continuum electrostatics, the factor $4\pi P(r)$ is smaller than e/r^2 at all distances. If the system had perfect spherical symmetry, **P** and $\mathbf{E}^{\mathrm{tot}}$ would be necessarily parallel or antiparallel. As observed in Fig. 11, intermediate orientations are also present in the system due to the periodic boundary conditions. In the case of the water solute, the cosine between the polarization relative and the electrostatic field differs from 1 and -1 at more than 60% of all grid points. In the present theory, the polarization density **P** is pointing in the direction of the effective orienting field **A**, which is obtained on the basis of Eq. (17) as a nonlocal function from $\psi^{(0)}$, $\psi^{(\mathrm{rf})}$, and **P**. The existence of regions in which the polarization density and the total electrostatic field do not point in the same direction is thus a direct consequence of the nonlocal convolution used to construct the orienting field **A**. Such important structural feature cannot be reproduced when it is assumed that the effective field orienting the solvent dipoles is equal to the electrostatic field modulated by a position-dependent prefactor.[8,26,27]

VI. SUMMARY

An integral equation theory was formulated to describe the average structure of a polar liquid around a complex molecular solute of irregular shape. The integral equation was developed from the HNC equation for a solute at infinite dilution immersed in a solvent constituted of hard spheres with embedded dipoles. The direct correlation function from the MSA theory was used for the unperturbed solvent. A formal separation of the direct correlation function into long and short range components was used to show that the properties of the system can be expressed in terms of average density, polarization density, and electrostatic potential. Equations (16), (22), and (23) are the main formal results of the paper.

It was demonstrated rigorously that the present integral equation is consistent with the equation of continuum macroscopic electrostatics. Therefore, the present theory provides a generalized alternative to the current calculations of electrostatic properties based on the Poisson equation.[3] Furthermore, the present analysis clarified the fundamental assumptions involved in the continuum approximation. In particular, the equations are linearized with respect to $\beta\mu|\mathbf{A}|$ and nonlinear effects, such as saturation of the solvent polarization and coupling of the solvent density to the electrostatic potential (electrostriction), are neglected. In addition, convolutions involving the solvent polarization and the solvent electrostatic potential with short range functions are replaced by local relations and variations in the solvent electrostatic potential and the polarization density on the lengthscale of the diameter of a solvent particle are neglected. As a direct consequence, the solvent polarization must point in the direction of the local electrostatic field.

It was shown that a linearized version of the integral equation corresponding to a three-dimensional extension of the MSA equation, the 3D-MSA, can be solved numerically to obtain the solvent structure around complex molecules of irregular shape. The method was illustrated with numerical applications to the case of an ion, a water molecule and NMA. The solvent–solute interaction potential was constructed from LJ centers and point charges as used in current molecular dynamics simulations of biomolecular systems.[1] The numerical solution, obtained directly in Cartesian coordinates using a discrete three-dimensional cubic grid, yields the average solvent density and polarization density at all the points x,y,z around the solute. All spatial convolutions appearing in the theory were calculated using three-dimensional numerical FFT. To our knowledge, equivalent solutions to the 3D-MSA equation have never been reported.

The approach described here is less computationally extensive than molecular dynamics simulations to obtain the average solvent structure around a complex biomolecule. Although the results obtained with the 3D-MSA equation are encouraging, further work will be necessary to improve the quantitative accuracy of the calculations. In particular, for the charged solute in a highly polar solvent, as it was shown in the case of the ion solute, the criterion of linearization is violated at some grid points and MSA model is not likely to be very accurate. As a result the polarization can exceed the physical limit of perfectly oriented dipoles in the linearized theory based on Eqs. (26) and (38). The purpose of the calculations was to illustrate the advantages and problems calculating the average solvent structure around a molecular solute with irregular shape using integral equations. Clearly, more work will be needed before such theory can be used for quantitative prediction and this is the object of current research. The most obvious problems could be avoided by the numerical solution of Eqs. (22) and (23) to account for nonlinear effects. Furthermore, the accuracy of the nonlinear HNC Eq. (10) could be improved, and made formally exact *in principle*, by incorporating the bridge function diagrams in the theory.[11] Although the exact bridge function is unknown, effective approximations could be constructed from the results of a reference hard sphere system, as in the R-HNC theory.[21] Finally, the present theory can be extended by treating the solvent as a mixture of dipolar spheres and ions to account for the presence of salt in the solution. This theory could be useful to consider highly charged polyionic biomolecules such as nucleic acids, which are difficult to treat with the current Poisson–Boltzmann equation. As suggested by

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp

the present analysis, it can be expected that such theory will be consistent with the Poisson–Boltzmann equation in the limit of small solvent and counter ions. Further work is currently in progress in these directions.

ACKNOWLEDGMENTS

This work was supported by NSERC. The initial ideas were elaborated during a stay of B.R. in the laboratory of J. C. Smith at the CEA in Saclay. We would like to express our gratitude for his support during this period. Useful discussions with Dongqing Wei are gratefully acknowledged.

[1] C. L. Brooks III, Martin Karplus, and B. M. Pettitt, in *Advances in Chemical Physics*, edited by I. Prigogine and S. A. Rice (Wiley, New York, 1988), Vol. LXXI.
[2] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1989).
[3] K. A. Sharp and B. Honig, Annu. Rev. Biophys. Biophys. Chem. **19**, 301 (1990).
[4] M. Born, Z. Phys. **1**, 45 (1920).
[5] J. G. Kirkwood, J. Chem. Phys. **2**, 351 (1934).
[6] A. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T. Hendrickson, and W. C. Still, J. Am. Chem. Soc. **113**, 1454 (1991).
[7] B. Roux, H.-A. Yu, and M. Karplus, J. Phys. Chem. **94**, 4683 (1990).
[8] A. Warshel and J. Åqvist, Annu. Rev. Biophys. Biophys. Chem. **20**, 267 (1991).
[9] D. W. R. Gruen and S. Marčelja, J. Chem. Soc. Faraday Trans. **79**, 211 (1983).
[10] D. W. R. Gruen and S. Marčelja, J. Chem. Soc. Faraday Trans. **79**, 225 (1983).
[11] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, London, 1976).
[12] D. Chandler and H. C. Andersen, J. Chem. Phys. **57**, 1930 (1972).
[13] D. Chandler, in *The Liquid State of Matter: Fluids, Simple and Complex*, edited by E. W. Montroll and J. L. Lebowitz (North-Holland, Amsterdam, 1982), Vol. VIII.
[14] F. Hirata and P. J. Rossky, Chem. Phys. Lett. **83**, 329 (1981).
[15] D. A. Zichi and P. J. Rossky, J. Chem. Phys. **84**, 1712 (1986).
[16] M. B. Pettitt and M. Karplus, Chem. Phys. Lett. **121**, 194 (1985).
[17] M. B. Pettitt, M. Karplus, and P. J. Rossky, J. Phys. Chem. **90**, 6335 (1986).
[18] M. B. Pettitt and M. Karplus, Chem. Phys. Lett. **136**, 383 (1987).
[19] W. F. Lau and B. M. Pettitt, Biopol. **26**, 1817 (1987).
[20] G. Ramé, W. F. Lau, and B. M. Pettitt, Int. J. Peptide Protein Res. **35**, 315 (1990).
[21] P. H. Fries and G. N. Patey, J. Chem. Phys. **82**, 429 (1985).
[22] G. M. Torrie, P. G. Kusalik, and G. N. Patey, J. Chem. Phys. **88**, 7826 (1988).
[23] P. G. Kusalik and G. N. Patey, J. Chem. Phys. **88**, 7715 (1988).
[24] P. H. Fries, W. Kunz, P. Calmettes, and P. Turq, J. Chem. Phys. **101**, 554 (1994).
[25] F. Lado, E. Lomba, and M. Lombardero, J. Chem. Phys. **103**, 481 (1995).
[26] Y. Liu and T. Ichiye, Chem. Phys. Lett. **231**, 380 (1994).
[27] J-K Hyun, C. S. Babu, and T. Ichiye, J. Chem. Phys. **99**, 5187 (1995).
[28] L. Onsager, J. Chem. Phys. **58**, 1468 (1936).
[29] D. Beglov and B. Roux, J. Chem. Phys. **103**, 360 (1995).
[30] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).
[31] M. S. Wertheim, J. Chem. Phys. **55**, 4291 (1971).
[32] J. D. Ramshaw, J. Chem. Phys. **66**, 3134 (1977).
[33] G. Nienhuis and J. M. Deutch, J. Chem. Phys. **55**, 4213 (1971).
[34] C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976).
[35] P. Debye, Phys. Z. **13**, 97 (1912).
[36] W. L. Jorgensen and J. Gao, J. Am. Chem. Soc. **110**, 4212 (1988).
[37] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[38] *CRC Handbook of Chemistry and Physics*, 72nd ed., edited by D. R. Lide (Chemical Rubber, Boston, 1992).
[39] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University, Cambridge, 1990).
[40] M. Neumann, Mol. Phys. **50**, 841 (1983).
[41] D. Y. C. Chan, D. J. Mitchell, and B. W. Ninham, J. Chem. Phys. **70**, 2946 (1979).

Downloaded¬05¬Oct¬2004¬to¬128.104.70.228.¬Redistribution¬subject¬to¬AIP¬license¬or¬copyright,¬see¬http://jcp.aip.org/jcp/copyright.jsp