An integral equation theory for solvation effects on the molecular structural fluctuation

Yoshihiro Matsumura and Hirofumi Sato

Articles you may be interested in

Multiple time step molecular dynamics in the optimized isokinetic ensemble steered with the molecular theory of solvation: Accelerating with advanced extrapolation of effective solvation forces
J. Chem. Phys. **139**, 244106 (2013); 10.1063/1.4848716

A modified repulsive bridge correction to accurate evaluation of solvation free energy in integral equation theory for molecular liquids
J. Chem. Phys. **137**, 024106 (2012); 10.1063/1.4733393

Molecular origin of the hydrophobic effect: Analysis using the angle-dependent integral equation theory
J. Chem. Phys. **128**, 024507 (2008); 10.1063/1.2823733

Solvation in molecular ionic liquids
J. Chem. Phys. **119**, 6411 (2003); 10.1063/1.1611875

Solvation structure and stability of peptides in aqueous solutions analyzed by the reference interaction site model theory
J. Chem. Phys. **107**, 1586 (1997); 10.1063/1.474511

An integral equation theory for solvation effects on the molecular structural fluctuation

Yoshihiro Matsumura[1] and Hirofumi Sato[1,2,a)]

[1]*Department of Molecular Engineering, Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan*
[2]*Elements Strategy Initiative for Catalysts and Batteries (ESICB), Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan*

A new integral equation theory is proposed, which enables us to efficiently compute conformational distribution of a polyatomic molecule in solution phase. The solvation effect on the intramolecular correlation function is evaluated through a self-consistent procedure. In addition, the analytical expression of solvation free energy is derived, explicitly taking into account the molecular structural fluctuation. The derived equation establishes a direct route between the structural fluctuation and free energy of the molecule. The method was successfully applied to a series of *n*-alkanes in aqueous solutions to demonstrate the superiority of the proposed theory. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4923038]

I. INTRODUCTION

Structural fluctuation is one of the fundamental aspects of a molecule. A flexible macromolecule possesses many different conformers, and the thermal fluctuation among them often plays an important role in many chemical and biological processes.[1,2] In computational approaches, molecular simulation like molecular dynamics is a straightforward and most widely used method[3–6] to deal with the fluctuation. However, in general, consideration of the full conformational space is infeasible, especially for large size molecules. Although many efficient algorithms have been proposed so far, essential understanding of the conformational space, namely, exploring the most stable conformation and accurate computation of the conformational distribution, is still very challenging even for single molecule. Computation of the molecular conformations in solution system is further time-consuming, especially for a molecule consisting of a few tens of heterogeneous atoms due to complicated intermolecular interaction from numerous surrounding molecules.

Integral equation theory (IET) is an alternative approach to tackle this difficulty.[7,8] In this theory, a closed-form equation of correlation functions is derived to characterize molecular system. For example, reference interaction site model (RISM),[8] a statistical mechanical theory for molecular liquids, is commonly utilized to deal with solvation effects. Thanks to its analytical nature, adequate statistical ensemble is expected to be obtained. While a simplified approach was adopted to describe polymer solutions,[9–12] the intramolecular correlation functions have been discussed so far to characterize the structure of polyatomic molecules.[13–22] Among them, Yokogawa *et al.* proposed a new IET,[22] in which the correlation function was divided into two parts; reference (ω^{ref}) and residual ($\Delta\omega$) parts. While the former was computed from Monte Carlo simulation on an isolated molecule, the latter part representing a change of intramolecular correlation functions attributed to many body effect in the liquid is obtained based on a new IET. Because difficulties in solution-phase computation are mainly attributed to this many body effect, so-called "sampling problem" can be largely avoided once a sufficient accuracy is achieved in the preparation of ω^{ref}. However, the theory is applicable only to pure liquid system consisting of a single component.

Here, we present a new integral equation theory to efficiently compute conformational distribution of a molecule immersed into solvent: the formula enables us to treat an infinitely dilute solution system. Although several approaches based on effective Hamiltonian have been proposed by utilizing Monte Carlo simulation,[18,19] atomistic level of theory for solute-solvent system has not been available. Another strong point of the present work is an analytical expression of solvation free energy including the effect of the structural fluctuation of the solute molecule. Unlike other works,[23–26] intramolecular correlation function is explicitly treated. The proposed method was applied to a series of *n*-alkanes in aqueous solutions to demonstrate the superiority of the present theory.

II. THEORY

A. Intramolecular correlation function

An infinitely dilute solution system is considered, where a poly-atomic, solute molecule is immersed into solvent. The intramolecular correlation function of the solute molecule is defined by

$$\omega_{ab}(r;\lambda) = \frac{\int d\mathbf{r^u} d\mathbf{r^v} \delta(|\mathbf{r}_a - \mathbf{r}_b| - r) e^{-\beta V_\lambda}}{4\pi r^2 Z_\lambda}, \quad (1)$$

a)hirofumi@moleng.kyoto-u.ac.jp

143, 014104-1

where $\beta = 1/k_B T$, k_B is Boltzmann constant. $\mathbf{r^u}$ and $\mathbf{r^v}$ are, respectively, solute (u) and solvent (v) coordinates, subscripts a and b denote atomic sites in the solute molecule. Z_λ is configuration integral of the system,

$$Z_\lambda = \int d\mathbf{r^u} d\mathbf{r^v} e^{-\beta V_\lambda}, \tag{2}$$

where V_λ is potential function of the system with coupling parameter λ ($0 \leq \lambda \leq 1$),

$$V_\lambda(\mathbf{r^u}, \mathbf{r^v}; \lambda) = V^{uu}(\mathbf{r^u}) + V^{uv}(\mathbf{r^u}, \mathbf{r^v}; \lambda) + V^{vv}(\mathbf{r^v}). \tag{3}$$

V^{uu} is intramolecular potential of the solute molecule, and V^{vv} is intermolecular interaction between solvent molecules. $V^{uv}(\lambda)$ is solute-solvent interaction, which is a function of λ, changing from $\lambda = 0$ (without interaction) to $\lambda = 1$ ("perfect" interaction). The effective (mean) solvation potential W^{solv} is

then introduced,[13,18,27–29]

$$Z_\lambda = \int d\mathbf{r^u} e^{-\beta V^{uu}} \int d\mathbf{r^v} e^{-\beta(V^{uv}(\lambda)+V^{vv})}$$

$$= \int d\mathbf{r^u} e^{-\beta(V^{uu}+W^{solv}(\lambda))} Z^v, \tag{4}$$

where

$$W^{solv}(\mathbf{r^u}; \lambda) = -\beta^{-1} \ln \frac{\int d\mathbf{r^v} e^{-\beta(V^{uv}(\lambda)+V^{vv})}}{Z^v}$$

$$= -\beta^{-1} \int_0^\lambda d\lambda' \frac{d}{d\lambda'} \ln \int d\mathbf{r^v} e^{-\beta(V^{uv}(\lambda')+V^{vv})}. \tag{5}$$

Z^v is configuration integral of the pure solvent system. When the V^{uv} and V^{vv} are site-pairwise and additive, the integration over λ' of Eq. (5) is analytically performed based on RISM/HNC (hypernetted chain) theory,[8,13,18,27,28]

$$W^{solv}(\mathbf{r^u}; \lambda) = -(2\beta)^{-1} \sum_{a,b} \sum_{SS'} \rho^v \int d\mathbf{r} \int d\mathbf{r}' c_{aS}^{uv}(|\mathbf{r_a} - \mathbf{r}|; \lambda) \chi_{SS'}^{vv}(|\mathbf{r} - \mathbf{r}'|) c_{bS'}^{uv}(|\mathbf{r}' - \mathbf{r_b}|; \lambda)$$

$$- \beta^{-1} \sum_{a,S} \rho^v \int d\mathbf{r} \{ c_{aS}^{uv}(r; \lambda) - \frac{1}{2}(h_{aS}^{uv}(r; \lambda))^2 \} \tag{6}$$

$$= \frac{1}{2} \sum_{a,b} \upsilon_{ab}(r_{ab}; \lambda) - \beta^{-1} \sum_{a,S} \rho^v \int d\mathbf{r} \{ c_{aS}^{uv}(r; \lambda) - \frac{1}{2}(h_{aS}^{uv}(r; \lambda))^2 \}. \tag{7}$$

h^{uv} and c^{uv} are the total and direct correlation functions between solute and solvent with the number density, ρ^v. Subscript "S" denotes an atomic site in the solvent molecule, $\chi_{SS'}^{vv}$ is density-density correlation function of the pure solvent,

$$\chi_{SS'}^{vv} = \omega_{SS'}^v + \rho^v h_{SS'}^{vv}, \tag{8}$$

where $h_{SS'}^{vv}$ and $\omega_{SS'}^v$ are total and intramolecular correlation functions of the pure solvent. While the second term of Eq. (7) represents the solute-solvent interaction, the first term (υ_{ab}) describes the solvent-mediated interactions between two sites (a and b) in the solute molecule, namely, solvent induced self-consistent pair interaction.[28] Based on these equations and the relation $\upsilon_{ab} = \upsilon_{ba}$, Eq. (1) is rewritten as follows:

$$\omega_{ab}(r; \lambda) = \omega_{ab}(r; 0) e^{-\beta \upsilon_{ab}(r; \lambda) - \beta \xi_{ab}(r; \lambda)}. \tag{9}$$

Here, ξ_{ab} is introduced to represent the indirect contribution,

$$\xi_{ab}(r; \lambda) = -\beta^{-1} \ln \frac{\omega_{ab}(r; \lambda)}{\omega_{ab}(r; 0)} - \upsilon_{ab}(r; \lambda). \tag{10}$$

Because of complicated many-body effect, the explicit form of this function is unknown. Alternatively, the function is expanded with respect to $\{\omega_{ab}\}$, corresponding to a HNC approximation,[4,7,24–26]

$$\xi_A(r; \lambda) \approx \sum_B \int dr' \frac{\delta \xi_A(r; \lambda)}{\delta \omega_B(r'; \lambda)} \bigg|_{\lambda=0} [\omega_B(r'; \lambda) - \omega_B(r'; 0)]$$

$$= \sum_B \int dr' \left\{ -\beta^{-1} \frac{\delta(r - r')}{\omega_B(r; 0)} \delta_{AB} - \frac{\delta \upsilon_A(r; \lambda)}{\delta \omega_B(r'; \lambda)} \bigg|_{\lambda=0} \right\} [\omega_B(r'; \lambda) - \omega_B(r'; 0)]$$

$$= -\beta^{-1} \left\{ \frac{\omega_A(r; \lambda)}{\omega_A(r; 0)} - 1 \right\} - \sum_B \int dr' \frac{\delta \upsilon_A(r; \lambda)}{\delta \omega_B(r'; \lambda)} \bigg|_{\lambda=0} [\omega_B(r'; \lambda) - \omega_B(r'; 0)]$$

$$= -\beta^{-1} \left\{ \frac{\omega_A(r; \lambda)}{\omega_A(r; 0)} - 1 \right\} + \beta^{-1} \sum_B \int dr' \frac{\chi_{A,B}^{-1}(r, r'; 0)}{4\pi r^2} [\omega_B(r'; \lambda) - \omega_B(r'; 0)]$$

$$= -\beta^{-1} \left\{ \frac{\omega_A(r; \lambda)}{\omega_A(r; 0)} - 1 \right\} + \beta^{-1} \gamma_A(r; \lambda). \tag{11}$$

Here, capitals (A and B) are used to denote a pair of atomic sites in the solute molecule a, b ($a \neq b$) to make the following derivation clear. Please note that only halves of pairs a, b ($a < b$) should be considered for A, B because of the symmetry, $\omega_{ab} = \omega_{ba}$, $\xi_{ab} = \xi_{ba}$, and so on. $\chi_{A,B}$ is the intramolecular density-density correlation function defined by

$$\begin{aligned}
\chi_{A,B}(r, r''; \lambda) &= \frac{1}{4\pi r''^2} \frac{\delta \omega_A(r; \lambda)}{\delta(-\beta v_B(r''; \lambda))} \\
&= \left\langle \frac{\delta(r_A - r)\delta(r_B - r'')}{(4\pi r^2)(4\pi r''^2)} \right\rangle_\lambda \\
&\quad - \left\langle \frac{\delta(r_A - r)}{4\pi r^2} \right\rangle_\lambda \left\langle \frac{\delta(r_B - r'')}{4\pi r''^2} \right\rangle_\lambda,
\end{aligned} \quad (12)$$

where $\langle \ldots \rangle_\lambda$ denotes an ensemble average at a given value of λ and

$$\begin{aligned}
\sum_C \int dr'' &\chi_{A,C}(r, r''; \lambda) \chi_{C,B}^{-1}(r'', r'; \lambda) \\
&= \sum_C \int dr'' \frac{\delta \omega_A(r; \lambda)}{\delta v_C(r''; \lambda)} \frac{\delta v_C(r''; \lambda)}{\delta \omega_B(r'; \lambda)} \\
&= \delta_{AB}\delta(r - r').
\end{aligned} \quad (13)$$

In the last equation in Eq. (11), the function γ_A is introduced, which has a similar role to direct correlation function

in the standard framework of integral equation theory. From Eqs. (10), (11), (13), OZ-type equation (14) and HNC-like closure equation (15) for the intramolecular correlation functions are, respectively, obtained,

$$\omega_{ab}(r; 0)\widetilde{\Delta\omega_{ab}}(r) = \frac{1}{2} \sum_{c,d} \int d\mathbf{r}' \gamma_{cd}(r'; 1) \chi_{cd,ab}(r', r; 0), \quad (14)$$

$$\gamma_{ab}(r; 1) = \widetilde{\Delta\omega_{ab}}(r) - \ln\left\{\widetilde{\Delta\omega_{ab}}(r) + 1\right\} - \beta v_{ab}(r; 1), \quad (15)$$

where

$$\begin{aligned}
\widetilde{\Delta\omega_{ab}}(r) &= \frac{\Delta\omega_{ab}(r)}{\omega_{ab}(r; 0)}, \\
\Delta\omega_{ab}(r) &= \omega_{ab}(r; 1) - \omega_{ab}(r; 0).
\end{aligned} \quad (16)$$

Here, the indices of atomic-sites pair are put back, and we used the relations, $\chi_{ab,cd}(r, r') = \chi_{cd,ab}(r', r)$, $\gamma_{cd} = \gamma_{dc}$, $\chi_{cd,ab} = \chi_{dc,ab}$, and $\chi_{cc,ad} = 0$. Note that the same equations are derived using variational procedure (see Appendix A).

The partial linearized HNC (PLHNC) approximation is also derived from Eq. (15),

$$\gamma_{ab}(r; 1) = \begin{cases} \widetilde{\Delta\omega_{ab}}(r) - \ln\left\{\widetilde{\Delta\omega_{ab}}(r) + 1\right\} - \beta v_{ab}(r) & \text{for} \quad \widetilde{\Delta\omega_{ab}}(r) \leq 0, \\ -\beta v_{ab}(r) & \text{for} \quad \widetilde{\Delta\omega_{ab}}(r) > 0. \end{cases} \quad (17)$$

In this study, we used this closure equation. In practice, $\omega_{ab}(r; 0)$ and $\chi_{ab,cd}(r, r'; 0)$ were obtained using Monte Carlo simulation of an isolated solute molecule, and these equations are iteratively solved to obtain the intramolecular correlation function $\omega_{ab}(r; 1)$. One might think if the splitting of ω into the reference and residual parts is always possible or not. The present procedure utilizes a self-consistent treatment, which is expected to be more robust than the perturbative approach. Actually, the procedure works well in all the present systems, and adequate solutions are always obtained in our experience.[30]

B. Expression of free energy

Based on these equations, solvation free energy ΔF involving the effects of solute structural fluctuation is analytically derived. ΔF is defined as the difference between the free energy of the solute-solvent system ($F_{\lambda=1}$) and that without solute-solvent interaction ($F_{\lambda=0}$). The "charging formula"[23] is employed,

$$\begin{aligned}
\Delta F &= F_{\lambda=1} - F_{\lambda=0} \\
&= -\beta^{-1} \int_0^1 d\lambda \frac{d\ln Z_\lambda}{d\lambda} = \int_0^1 d\lambda \left\langle \frac{dW^{\text{solv}}(\lambda)}{d\lambda} \right\rangle_\lambda \\
&= \frac{1}{2} \sum_{a,b} \int_0^1 d\lambda \int d\mathbf{r} \frac{dv_{ab}(r; \lambda)}{d\lambda} \omega_{ab}(r; \lambda) - \beta^{-1} \sum_{a,S} \rho^{\text{v}} \int d\mathbf{r} \int_0^1 d\lambda \frac{d}{d\lambda} \left\{ c_{aS}^{\text{uv}}(r; \lambda) - \frac{1}{2}(h_{aS}^{\text{uv}}(r; \lambda))^2 \right\}.
\end{aligned} \quad (18)$$

Using the following relationship derived from Eq. (10) in the first term,

$$\frac{dv_{ab}(r; \lambda)}{d\lambda} \omega_{ab}(r; \lambda) = -\beta^{-1} \frac{d\omega_{ab}(r; \lambda)}{d\lambda} - \omega_{ab}(r; \lambda) \frac{d\xi_{ab}(r; \lambda)}{d\lambda}, \quad (19)$$

the integration over λ is carried out using Eq. (15) or Eq. (17). If we choose the former, the following HNC-like expression is obtained:

$$\begin{aligned}
\Delta F^{\text{HNC}} &= (2\beta)^{-1} \sum_{a,b} \int d\mathbf{r} \, \omega_{ab}(r; 0) \left\{ \frac{1}{2} \widetilde{\Delta\omega_{ab}}^2(r) - \gamma_{ab}(r; 1) - \frac{1}{2} \widetilde{\Delta\omega_{ab}}(r) \gamma_{ab}(r; 1) \right\} \\
&\quad - \beta^{-1} \sum_{a,S} \rho^{\text{v}} \int d\mathbf{r} \left\{ c_{aS}^{\text{uv}}(r; 1) - \frac{1}{2}(h_{aS}^{\text{uv}}(r; 1))^2 \right\}.
\end{aligned} \quad (20)$$

Similarly, PLHNC-like expression is obtained,

$$\Delta F^{\text{PLHNC}} = (2\beta)^{-1} \sum_{a,b} \int d\mathbf{r}\, \omega_{ab}(r;0) \left\{ \frac{1}{2} \widetilde{\Delta\omega_{ab}}^2(r) \Theta(-\widetilde{\Delta\omega_{ab}}) - \gamma_{ab}(r;1) - \frac{1}{2} \widetilde{\Delta\omega_{ab}}(r)\gamma_{ab}(r;1) \right\}$$
$$- \beta^{-1} \sum_{a,S} \rho^{\text{v}} \int d\mathbf{r} \left\{ c_{aS}^{\text{uv}}(r;1) - \frac{1}{2}(h_{aS}^{\text{uv}}(r;1))^2 \right\}, \tag{21}$$

where Θ is the Heaviside step function, *i.e.*, $\Theta(x) = 1$ ($x \geq 0$) or $\Theta(x) = 0$ ($x < 0$). These analytical expressions allow us to readily calculate ΔF.

III. COMPUTATIONAL DETAILS

A. Models for *n*-alkanes in aqueous solutions

In this study, three models are employed for *n*-butane. One is a model which modified CHARMM96[31,32] (denoted as Model 1) and the second is OPLS-AA[33] (Model 2). In these all-atom type models, bond lengths of C–H and C–C are, respectively, fixed at 1.09 and 1.53 Å. Bond angles of H–C–H, C–C–H, and C–C–C are all fixed at 109.5°, and the torsional part in V^{uu} is described with the potential V^{Torsion} for dihedral angle ϕ of C–C–C–C,

$$V^{\text{Torsion}}(\phi) = V_0 + \sum_{n=1}^{5} V_n[1 - (-1)^n \cos(n\phi)], \tag{22}$$

where the parameter sets, $V_0 = 0.000$, $V_1 = 2.348$, $V_2 = -1.568$, $V_3 = 2.535$, $V_4 = -0.784$, $V_5 = 0.470$ kcal mol^{-1}, which correspond to Scott Scheraga potential,[34] are employed for "Model 1," and OPLS-UA torsional potential[35] is used for "Model 2." The third model (Model 3) is OPLS-UA (united atom).[35] The bond length of C–C is fixed at 1.53 Å, and the bond angle C–C–C is fixed at 112.0°. For other *n*-alkanes, from pentane to decane, Model 3 (OPLS-UA) is used. Pentane and heptane are also treated with Model 2 (OPLS-AA). TIP3P-type water model is chosen as a solvent. All the Lennard Jones parameters and atomic charges are taken from the literature[31–33,35] and listed in Table I. Non-bonded interaction between groups separated by more than three bonds was added into the intramolecular interaction potential.

B. Computational procedures

Procedures of the present method are shown in Scheme 1. First of all, $\omega_{ab}(r;0)$ and $\chi_{ab,cd}(r,r';0)$ are prepared using the Monte Carlo simulation of an isolated solute molecule. $\omega_{ab}(r;0)$ is used as input for solving the RISM/HNC equations to obtain h^{uv} and c^{uv}, which represent the solvation structure. All of h^{uv}, c^{uv}, and $\chi_{ab,cd}(r,r';0)$ are input to compute $\omega_{ab}(r;1)$. The obtained function is then used for evaluating new h^{uv} and c^{uv}, and $\omega_{ab}(r;1)$ is also updated. This procedure is repeated until the converged $\omega_{ab}(r;1)$ is obtained.

In the RISM/HNC calculation, the number of grid points and the grid width are 4096 and 0.010 bohr, respectively. In

the calculation of the integral equations for $\omega_{\lambda=1}$, the number of grid points and the grid width are 1024 and 0.016 bohr for Model 1. 2048 and 0.016 bohr are employed in Model 3 (OPLS-UA) and 512 and 0.032 bohr in Model 2 (OPLS-AA). The details of grid parameters are discussed in the supplementary material.[36]
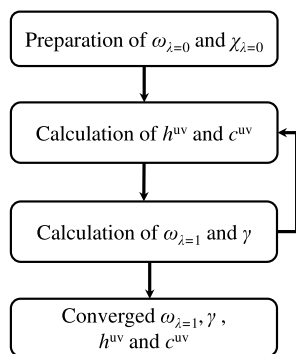
IV. RESULTS AND DISCUSSION

A. Conformational equilibria of *n*-alkanes

First, we examined *n*-butane in aqueous solutions. Because the molecule has only one conformational degree of freedom, the hydration effects on *n*-butane have been investigated by many researchers.[15,31,37–43]

Fig. 1 displays the intramolecular correlation functions between terminal carbon atoms. Results of the three different models for an isolated molecule and aqueous solutions are shown. (The same functions are plotted with respect to the dihedral angle. Please see the supplementary material.[36]) The temperature is set at 298.15 K with a standard density of solvent water (1.00 g cm^{-3}). In each model, two peaks are found around 3.0 and 3.8 Å. The former corresponds to the *gauche* conformer and the latter to the *trans*. The *gauche* peak is broader than the *trans*, meaning that the structural fluctuation of the *gauche* conformer is larger. In aqueous solutions, the *gauche* peak becomes higher, which is consistent with a conventional picture of hydrophobic interactions.

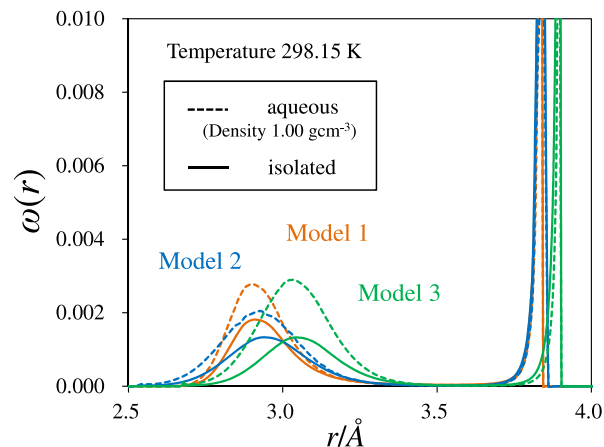TABLE I. Lennard-Jones parameters of each atom.[31–33,35]

Site	σ/Å	ϵ/kcal mol^{-1}	Z/e
Model 1			
C(CH$_3$)	2.001	0.078	−0.249
H(CH$_3$)	0.835	0.024	0.066
C(CH$_2$)	2.001	0.056	−0.121
H(CH$_2$)	0.835	0.028	0.086
Model 2			
C(CH$_3$)	3.500	0.066	−0.180
C(CH$_2$)	3.500	0.066	−0.120
H	2.500	0.030	0.060
Model 3			
CH$_3$	3.905	0.175	0.000
CH$_2$	3.905	0.118	0.000
TIP3P-type water			
O	3.150	0.152	−0.834
H	0.400	0.046	0.417

```
┌─────────────────────────────────┐
│ Preparation of ω_{λ=0} and χ_{λ=0} │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│ Calculation of h^{uv} and c^{uv}  │◄─┐
└─────────────────────────────────┘  │
                 │                    │
                 ▼                    │
┌─────────────────────────────────┐  │
│ Calculation of ω_{λ=1} and γ      │──┘
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│ Converged ω_{λ=1}, γ,             │
│ h^{uv} and c^{uv}                 │
└─────────────────────────────────┘
```

SCHEME 1.

FIG. 1. Intramolecular correlation function between terminal carbon atoms of n-butane.

Equilibrium constant for the *trans-gauche* isomerization ($K = x_{gauche}/x_{trans}$) is estimated from the populations of *gauche* (x_{gauche}) and *trans* (x_{trans}). The results are listed in Table II together with previous works.[15,31,39–46] This molecule has been used as a typical example to study the solvation effects on the structural fluctuation. Though numerous potential parameters are available, the equilibrium constants of isolated molecule evaluated are similar, and the population of the *trans* conformer is about twice as that of the *gauche* form. In aqueous environment, the equilibrium constant is increased in almost all the studies. In the present calculation, both of the all atom models (Model 1 and 2) show the similar degree of hydration effects on K value, while the result of united atom model (Model 3) becomes further large. As shown in the table, the computed K's are scattered from 0.5 to 4, and the present results fall within these values in previous studies. The interaction potential of Model 1 is the same as that in the work by Cui and Smith[31] demonstrating the reliability of the present theory.

The intramolecular correlation functions between two terminal carbon atoms in longer-chain molecules are plotted in Fig. 2. The temperature is set at 298.15 K with a standard density of solvent water (1.00 g cm^{-3}). As the number of carbon atoms is increased from n-pentane to n-decane, the distribution becomes broader and shifts to longer. This is quan-

titatively evaluated using a mean distance between terminal carbon atoms (Table III),

$$\langle r \rangle = \int d\mathbf{r} r \omega(r). \tag{23}$$

The mean distances of Model 3 and 2 are, respectively, 4.72 and 4.58 Å (n-pentane), 5.74 and 5.57 Å (n-heptane). Although former model gives slightly longer distance, both of the united- and all-atom models exhibit the similar trend. The hydration effects shrink the solute structure, and $\langle r \rangle$'s are slightly decreased in aqueous solutions. The difference due to solvation looks to follow a regular curve with a minimum at octane-nonane except for butane (Model 3). The gap between butane and other alkanes may be explained in terms of pentane effect, which does not exist only in butane. A conformation with the shortest end-to-end distance, which is allowed geometrically, can be taken in the case of butane. But the pentane effect prohibits longer alkanes to take such conformations in reality. In solution phase, butane is surrounded by solvent molecules, which is significantly different from isolated butane. For other alkanes, their (effective) conformational space is reduced by pentane

TABLE II. Equilibrium constants for the trans-gauche isomerization of n-butane at ambient condition.

Method	Model	$K = x_{gauche}/x_{trans}$		Reference
		Isolated	Aqueous	
MD	United	0.54	0.54	Tobias and Brooks[40]
	All	0.59	0.85	Tobias and Brooks[40]
MC	United	0.54	2.3	Rosenberg *et al.*[41]
RISM/HNC	All	0.54	2.08	Cui and Smith[31]
RISM/HNC	United	0.54	~4[a]	Zichi and Rossky[15]
IET[b]	United	0.50	1.2	Pratt and Chandler[39]
MC	United	0.47	1.0	Jorgensen[42]
MC	United	0.47	0.79	Jorgensen and Buckner[43]
This work				
	All	0.52	0.96	Model 1
	All	0.48	0.90	Model 2
	United	0.48	2.03	Model 3
Exp.		0.40-0.50		[44–46]

[a]Reference 40, estimated from Figure 2 of Reference 15.
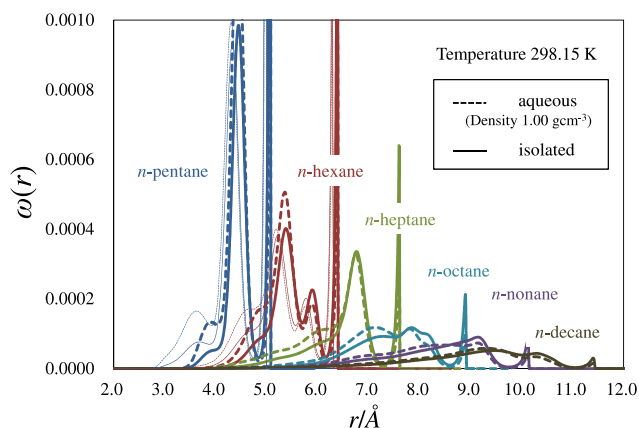[b]Theory based on an integral equation for the pair correlation functions.

FIG. 2. Intramolecular correlation functions between terminal carbon atoms of n-alkanes from pentane to decane. Thick and thin lines denote Models 3 and 2, respectively.

effect, even in the isolated molecule. By closely looking at the individual intramolecular correlation functions, *gauche* populations of each dihedral angle are almost unchanged by hydration except for both ends. The change of the distribution in the figure is mainly attributed to the terminus carbons. A recent replica exchange molecular dynamics, REMD, simulation by Ferguson *et al.* shows that a water-induced chain collapse is not observed for n-alkanes smaller than n-eicosane.[6] The present calculations up to n-decane are in qualitative agreement with this simulation. The results of potential of mean force plots, root-mean-square end to end distance,[30] $\sqrt{\langle r^2 \rangle}$, and conformational population, $4\pi r^2 \Delta\omega(r)$, of n-alkanes are also discussed in the supplementary material.[36]

B. Hydration free energy

The hydration free energy of the n-butane is fundamental interest from the viewpoint of hydrophobic effects. We calculated the ΔF (Eq. (21)) of n-butane in aqueous solutions at the four thermodynamic conditions, for which experimental data are available.[37,38] The temperature and solvent density of these conditions are follows: (**I**) 283.15 K and 0.9997 g cm^{-3}, (**II**) 298.15 K and 0.9970 g cm^{-3}, (**III**) 313.15 K and 0.9922 g cm^{-3}, and (**IV**) 328.15 K and 0.9875 g cm^{-3}. The results are listed in Table IV, where the condition I is chosen as a reference point ($\Delta\Delta F$). The calculated $\Delta\Delta F$ using three models resembles the experimental data. As described in Sec. IV A, computed K depends sensitively on the model of n-butane, but the agree-

TABLE III. Mean distance between terminal carbon atoms of n-alkane, $\langle r \rangle$/Å.

Solute	Model 3			Model 2		
	Isolated	Aqueous	Diff.	Isolated	Aqueous	Diff.
n-butane	3.63	3.34	−0.29	3.56	3.41	−0.15
n-pentane	4.72	4.54	−0.18	4.58	4.45	−0.13
n-hexane	5.74	5.51	−0.23	5.57	5.44	−0.13
n-heptane	6.71	6.47	−0.24			
n-octane	7.66	7.39	−0.27			
n-nonane	8.54	8.27	−0.27			
n-decane	9.36	9.12	−0.24			

TABLE IV. Hydration free energy of n-butane.

	$\Delta\Delta F$/kcal mol^{-1}				
	Model 1	Model 2	Model 3	Cui and Smith[31]	Expt.[a]
I[b]	0.000	0.000	0.000	0.000	0.000
II	0.675	1.313	0.898	0.689	0.424
III	1.266	2.410	1.617	1.293	0.781
IV	1.837	3.478	2.315	1.821	1.073

[a]Experimental data are taken from References 31, 37, and 38.
[b]Condition **I** is chosen as a reference point, so the values equal to zero.

ment in $\Delta\Delta F$ is semi-quantitative. As shown in Appendix B, the contributions from effective solvation potential (W^{solv}) are dominant in ΔF (Eq. (B4)). In the end, we note that the hydration free energy of n-alkanes becomes increasing as the chain length becomes longer. However, in order to more accurately evaluate ΔF, we have to use the repulsive bridge correction[47] for RISM/HNC theory, which is beyond the scope of this paper.

V. CONCLUSION

In the present study, we developed a new theory to efficiently calculate molecular structural fluctuation in solute-solvent system at an infinite dilution. In the methodology, the molecular structural fluctuation is described by a set of the intramolecular correlation functions. Monte Carlo method is employed to evaluate the fluctuation in an isolated solute molecule, and then the solvation part is calculated based on a new integral equation theory. An analytical expression of the corresponding solvation free energy is also derived.

We applied the theory to a series of n-alkanes in the aqueous solutions and examined the structural fluctuations. The distribution of n-butane was computed, where the structural fluctuation of *gauche* conformer is larger than that of *trans*, and the population of *gauche* becomes larger by hydration. The thermodynamic dependence of hydration free energy of n-butane semi-quantitatively reproduces the experimental value. The contributions of effective solvation potential are dominant due to the compensation of other terms. The conformations of n-alkanes up to n-decane in aqueous solutions are also examined. The distribution of compact structures is slightly populated due to hydration.

ACKNOWLEDGMENTS

APPENDIX A: VARIATIONAL PRINCIPLE

The same set of equations can be derived based on the variational principle. The procedure is very similar to that for the standard RISM/HNC equations and begun with the trial function, $\mathcal{L}$,

$$
\begin{aligned}
\beta\mathcal{L} = &-\frac{1}{2}\sum_{a,b}\int d\mathbf{r}\,\omega_{ab}(r;0)\{e^{-\beta\upsilon_{ab}(r;1)-\beta\xi_{ab}(r;1)} - 1 + \beta\xi_{ab}(r;1) + \beta\widetilde{\Delta\omega_{ab}}(r)\xi_{ab}(r;1)\\
&+\frac{1}{2}\widetilde{\Delta\omega_{ab}}^2(r)\} - \frac{1}{8}\sum_{a,b,c,d}\int d\mathbf{r}\int d\mathbf{r}'\gamma_{ab}(r;1)\chi_{ab,cd}(r,r';0)\gamma_{cd}(r';1)\\
&-\frac{1}{2}\sum_{a,b}\int d\mathbf{r}\,\omega_{ab}(r;0)\{-\widetilde{\Delta\omega_{ab}}(r)\gamma_{ab}(r;1)\}\\
&-\sum_{a,S}\rho^{\mathrm{v}}\int d\mathbf{r}\{e^{-\beta U_{aS}^{\mathrm{LJ+ES}}(r)+t_{aS}(r)} - 1 - t_{aS}(r) - h_{aS}^{\mathrm{uv}}(r;1)t_{aS}(r) + \frac{1}{2}(h_{aS}^{\mathrm{uv}}(r;1))^2\}\\
&-\sum_{a,S}\rho^{\mathrm{v}}\int d\mathbf{r}\{-h_{aS}^{\mathrm{uv}}(r;1)c_{aS}^{\mathrm{uv}}(r;1)\},
\end{aligned}
\tag{A1}
$$

$$
\upsilon_{ab}(r;1) = -\beta^{-1}\sum_{S,S'}\rho^{\mathrm{v}}\int d\mathbf{r}''\int d\mathbf{r}'c_{aS}^{\mathrm{uv}}(|\mathbf{r}-\mathbf{r}''|;1)\chi_{SS'}^{\mathrm{vv}}(|\mathbf{r}''-\mathbf{r}'|)c_{bS'}^{\mathrm{uv}}(|\mathbf{r}'|;1).
\tag{A2}
$$

Here, $t_{ab} = h_{aS}^{\mathrm{uv}} - c_{aS}^{\mathrm{uv}}$, $U_{aS}^{\mathrm{LJ+ES}}$ is the bare solute-solvent interaction potential. It is noted that, although the formal expression is the same as that in ΔF (Eq. (20)), all (variable) functions appearing in $\mathcal{L}$ are freely chosen. In other words, $\mathcal{L}$ can be regarded as a functional of the correlation functions ξ_{ab}, $\widetilde{\Delta\omega_{ab}}$, γ_{ab}, t_{aS}, h_{aS}^{uv}, and c_{aS}^{uv}.

Now, variations with respect to these functions yield

$$
\begin{aligned}
\delta(\beta\mathcal{L}) = &\frac{\beta}{2}\sum_{a,b}\int d\mathbf{r}\,\omega_{ab}(r;0)\{e^{-\beta\upsilon_{ab}(r;1)-\beta\xi_{ab}(r;1)} - 1 - \widetilde{\Delta\omega_{ab}}(r)\}\delta\xi_{ab}(r;1)\\
&-\frac{1}{2}\sum_{a,b}\int d\mathbf{r}\,\omega_{ab}(r;0)\{\beta\xi_{ab}(r;1) - \gamma_{ab}(r;1) + \widetilde{\Delta\omega_{ab}}(r)\}\delta\widetilde{\Delta\omega_{ab}}(r)\\
&-\frac{1}{2}\sum_{a,b}\int d\mathbf{r}\{\frac{1}{2}\sum_{c,d}\int d\mathbf{r}'\gamma_{cd}(r';1)\chi_{cd,ab}(r',r;0) - \omega_{ab}(r;0)\widetilde{\Delta\omega_{ab}}(r)\}\delta\gamma_{ab}(r;1)\\
&-\sum_{a,S}\rho^{\mathrm{v}}\int d\mathbf{r}\{e^{-\beta U_{aS}^{\mathrm{LJ+ES}}(r)+t_{aS}(r)} - 1 - h_{aS}^{\mathrm{uv}}(r;1)\}\delta t_{aS}(r)\\
&-\sum_{a,S}\rho^{\mathrm{v}}\int d\mathbf{r}\{\sum_{b,S'}\int d\mathbf{r}''\omega_{ab}(|\mathbf{r}''|;1)\int d\mathbf{r}'c_{bS'}^{\mathrm{uv}}(|\mathbf{r}''-\mathbf{r}'|;1)\chi_{S'S}^{\mathrm{vv}}(|\mathbf{r}'-\mathbf{r}|)\\
&- h_{aS}^{\mathrm{uv}}(r;1)\}\delta c_{aS}^{\mathrm{uv}}(r;1)\\
&-\sum_{a,S}\rho^{\mathrm{v}}\int d\mathbf{r}\{-t_{aS}(r) - c_{aS}^{\mathrm{uv}}(r;1) + h_{aS}^{\mathrm{uv}}(r;1)\}\delta h_{aS}^{\mathrm{uv}}(r;1).
\end{aligned}
\tag{A3}
$$

The curly brackets in the first and second terms give the HNC-like closure for the intramolecular correlation functions (Eq. (15)). The third one is the OZ-type equation (Eq. (14)). The fourth and fifth ones give the HNC closure and the RISM equation for the solute-solvent intermolecular correlation functions, respectively. Finally, the sixth one defines t_{aS}.

In the present methodology, these closed equations for the correlation functions, ξ_{ab}, $\widetilde{\Delta\omega_{ab}}$, γ_{ab}, t_{aS}, h_{aS}^{uv}, and c_{aS}^{uv} are iteratively solved to attain mutual convergence; namely, the converged correlation functions are consistent to each other.

APPENDIX B: COMPONENTS IN FREE ENERGY

In this section, we will consider the compensation relation between energy and entropy terms in free energy change (ΔF).[48] Energy part (ΔE) and entropy part ($T\Delta S$) are, respectively, defined by

$$
\Delta E = \left(\frac{\partial(\beta\Delta F)}{\partial\beta}\right)_{N,V}, \quad T\Delta S = \beta\left(\frac{\partial\Delta F}{\partial\beta}\right)_{N,V}.
\tag{B1}
$$

Using Eq. (18) for the expression of ΔF, ΔE is rewritten as follows:

$$
\Delta E = \left\langle\frac{\partial(\beta W^{\mathrm{solv}}(1))}{\partial\beta}\right\rangle_{\lambda=1} + \beta\int_0^1 d\lambda\left\{\left\langle\frac{dW^{\mathrm{solv}}(\lambda)}{d\lambda}\right\rangle_\lambda\langle V^{\mathrm{uu}}\rangle_\lambda - \left\langle\frac{dW^{\mathrm{solv}}(\lambda)}{d\lambda}V^{\mathrm{uu}}\right\rangle_\lambda\right\}.
\tag{B2}
$$

This is essentially a similar to the well-known expression, but the effective solvation potential (W^{solv}) also depends on temperature, β. $T\Delta S$ is also rewritten as follows:

$$
\begin{aligned}
T\Delta S &= \beta \int_0^1 d\lambda \, \frac{\partial}{\partial \beta} \left\langle \frac{dW^{\text{solv}}(\lambda)}{d\lambda} \right\rangle_\lambda \\
&= \beta \int_0^1 d\lambda \left[\left\langle \frac{d}{d\lambda} \frac{\partial W^{\text{solv}}(\lambda)}{\partial \beta} \right\rangle_\lambda - \left\langle \frac{dW^{\text{solv}}(\lambda)}{d\lambda}(V^{\text{uu}} + W^{\text{solv}}(\lambda)) \right\rangle_\lambda - \beta \left\langle \frac{dW^{\text{solv}}(\lambda)}{d\lambda} \cdot \frac{\partial W^{\text{solv}}(\lambda)}{\partial \beta} \right\rangle_\lambda \right. \\
&\quad \left. + \left\langle \frac{dW^{\text{solv}}(\lambda)}{d\lambda} \right\rangle_\lambda \left\langle V^{\text{uu}} + W^{\text{solv}}(\lambda) + \beta \frac{\partial W^{\text{solv}}(\lambda)}{\partial \beta} \right\rangle_\lambda \right].
\end{aligned}
\tag{B3}
$$

Thus, the last two terms in Eq. (B2) are exactly canceled out in $T\Delta S$ and do not contribute to the total free energy change, $\Delta F = \Delta E - T\Delta S$. Consequently, ΔF is written as contributions from W^{solv},

$$
\begin{aligned}
\Delta F &= \left\langle W^{\text{solv}}(1) \right\rangle_{\lambda=1} \\
&\quad - \beta \int_0^1 d\lambda \left\{ \left\langle \frac{dW^{\text{solv}}(\lambda)}{d\lambda} \right\rangle_\lambda \left\langle W^{\text{solv}}(\lambda) \right\rangle_\lambda - \left\langle \frac{dW^{\text{solv}}(\lambda)}{d\lambda} \cdot W^{\text{solv}}(\lambda) \right\rangle_\lambda \right\}.
\end{aligned}
\tag{B4}
$$

[1] H. Nakagawa and M. Kataoka, J. Phys. Soc. Jpn. **79**, 083801 (2010).

[2] N. Hirota and K. Mizuno, J. Mol. Biol. **275**, 365 (1998).

[3] M. Karplus and J. N. Kushick, Macromolecules **14**, 325 (1981).

[4] D. Yokogawa and T. Ikegami, J. Phys. Chem. **131**, 221101 (2009).

[5] S.-H. Chong and S. Ham, Chem. Phys. Lett. **504**, 225 (2011).

[6] A. L. Ferguson, P. G. Debenedetti, and A. Z. Panagiotopoulos, J. Phys. Chem. B **113**, 6405 (2009).

[7] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Academic, New York, 1986).

[8] *Molecular Theory of Solvation*, Understanding Chemical Reactivity, edited by F. Hirata (Springer, 2003).

[9] K. S. Schweizer and J. G. Curro, Chem. Phys. **149**, 105 (1990).

[10] K. S. Schweizer and J. G. Curro, J. Chem. Phys. **94**, 3986 (1991).

[11] K. S. Schweizer, Macromolecules **26**, 6033 (1993).

[12] M. Fuchs, Z. Phys. B: Condens. Matter **103**, 512 (1997).

[13] D. Chandler and L. R. Pratt, J. Chem. Phys. **65**, 2925 (1976).

[14] (a) L. R. Pratt, C. S. Hsu, and D. Chandler, J. Chem. Phys. **68**, 4202 (1978); (b) C. S. Hsu, L. R. Pratt, and D. Chandler, *ibid.* **68**, 4213 (1978).

[15] D. A. Zichi and P. J. Rossky, J. Chem. Phys. **84**, 1712 (1986).

[16] K. Dyer, J. Perkyns, B. Pettitt, and G. Stell, Condens. Matter Phys. **10**, 331 (2007).

[17] B. Kim, S.-H. Chong, R. Ishizuka, and F. Hirata, Condens. Matter Phys. **11**, 179 (2008).

[18] C. J. Grayce and K. S. Schweizer, J. Chem. Phys. **100**, 6846 (1994).

[19] T. Sumi and H. Sekino, J. Chem. Phys. **122**, 194910 (2005).

[20] T. Munakata, S. Yoshida, and F. Hirata, Phys. Rev. E **54**, 3687 (1996).

[21] R. Ishizuka and F. Hirata, Chem. Phys. Lett. **420**, 135 (2006).

[22] D. Yokogawa, H. Sato, and S. Sakaki, Chem. Phys. Lett. **487**, 241 (2010).

[23] S. J. Singer and D. Chandler, Mol. Phys. **55**, 621 (1985).

[24] N. Matubayasi and M. Nakahara, J. Chem. Phys. **113**, 6070 (2000).

[25] N. Matubayasi and M. Nakahara, J. Chem. Phys. **117**, 3605 (2002).

[26] N. Matubayasi and M. Nakahara, J. Chem. Phys. **119**, 9686 (2003).

[27] L. R. Pratt and D. Chandler, J. Chem. Phys. **66**, 147 (1977).

[28] D. Chandler, Y. Singh, and D. M. Richardson, J. Chem. Phys. **81**, 1975 (1984).

[29] The effective (mean) solvation potential W^{solv} is related to the solute intramolecular distribution function s_λ,[13,18,27,28]

$$
W^{\text{solv}}(\mathbf{r}^{\mathbf{u}}; \lambda) = -\beta^{-1} \ln \frac{\int d\mathbf{r}^{\mathbf{v}} e^{-\beta(V^{\text{uv}}(\lambda) + V^{\text{vv}})}}{Z^{\text{v}}} = -\beta^{-1} \ln \left(\frac{s_\lambda(\mathbf{r}^{\mathbf{u}})}{s_0(\mathbf{r}^{\mathbf{u}})} \cdot \frac{Z_\lambda}{Z_0} \right),
$$

where

$$
s_\lambda(\mathbf{r}^{\mathbf{u}}) = \frac{e^{-\beta V^{\text{uu}}(\mathbf{r}^{\mathbf{u}})} \int d\mathbf{r}^{\mathbf{v}} e^{-\beta(V^{\text{uv}}(\lambda) + V^{\text{vv}})}}{Z_\lambda}
$$

and

$$
s_0(\mathbf{r}^{\mathbf{u}}) = \frac{e^{-\beta V^{\text{uu}}(\mathbf{r}^{\mathbf{u}})} Z^{\text{v}}}{Z_0}
$$

or

$$
s_\lambda(\mathbf{r}^{\mathbf{u}}) = \frac{e^{-\beta(V^{\text{uu}}(\mathbf{r}^{\mathbf{u}}) + W^{\text{solv}}(\mathbf{r}^{\mathbf{u}}; \lambda))}}{\int d\mathbf{r}^{\mathbf{u}} e^{-\beta(V^{\text{uu}}(\mathbf{r}^{\mathbf{u}}) + W^{\text{solv}}(\mathbf{r}^{\mathbf{u}}; \lambda))}}.
$$

Here, we used Eqs. (4) and (5).

[30] A possibility cannot be excluded that the procedure is failed—for example, $\Delta\omega$ has multiple numerical solutions—when it is applied to a more complicated system such as polymer. Further modification of the theory may be necessary to treat such a system; for example, solvent effects on the scaling laws and Flory exponents governing the radius of gyration are analysed.

[31] Q. Cui and V. H. Smith, Jr., J. Phys. Chem. B **106**, 6554 (2002).

[32] D. Yin and A. D. Mackerell, Jr., J. Comput. Chem. **19**, 334 (1998).

[33] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).

[34] R. A. Scott and H. A. Scheraga, J. Chem. Phys. **44**, 3054 (1966).

[35] W. L. Jorgensen, J. D. Madura, and C. J. Swenson, J. Am. Chem. Soc. **106**, 813 (1984).

[36] See supplementary material at http://dx.doi.org/10.1063/1.4923038 for the details of grid parameters, root-mean-square end-to-end distance of *n*-alkanes, potential of mean force with respect to dihedral angle of *n*-butane, and conformational population of *n*-alkanes.

[37] E. Wilhelm, R. Battino, and R. J. Wilcock, J. Chem. Rev. **77**, 219 (1977).

[38] A. Ben-Naim and Y. Marcus, J. Chem. Phys. **81**, 2016 (1984).

[39] L. R. Pratt and D. Chandler, J. Chem. Phys. **67**, 3683 (1977).

[40] D. J. Tobias and C. L. Brooks, J. Chem. Phys. **92**, 2582 (1990).

[41] R. O. Rosenberg, R. Mikkilineni, and B. J. Berne, J. Am. Soc. **104**, 7647 (1982).

[42] W. R. Jorgensen, J. Chem. Phys. **77**, 5757 (1982).

[43] W. R. Jorgensen and J. K. Buckner, J. Chem. Phys. **91**, 6083 (1987).

[44] A. L. Verma, W. F. Murphy, and H. J. Bernstein, J. Chem. Phys. **60**, 1540 (1974).

[45] J. R. Durig and D. A. C. Compton, J. Phys. Chem. **83**, 265 (1979).

[46] D. A. C. Compton, S. Montero, and W. F. Murphy, J. Phys. Chem. **84**, 3587 (1980).

[47] A. Kovalenko and F. Hirata, J. Chem. Phys. **113**, 2793 (2000).

[48] H.-A. Yu and M. Karplus, J. Chem. Phys. **89**, 2366 (1988).