

Size dependence of transfer free energies: A hard-sphere-chain-based formalism

Seishi Shimizu,^{a)} Mitsunori Ikeguchi, Shugo Nakamura, and Kentaro Shimizu
*Bioinformation Engineering Laboratory, Department of Biotechnology, The University of Tokyo,
1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan*

(Received 11 May 1998; accepted 2 November 1998)

The main purpose of this paper is to present a theoretical scheme which describes the solvation and transfer free energies of small molecules and relate them to solvent contributions in the biomolecular processes. Several proposals, based originally on Flory–Huggins theory, have been made recently that there is a non-negligible solute's volume-proportional term in solvation free energy and the term should be subtracted to obtain solute/solvent contact free energy for biochemical applications. These proposals have resulted in the revision of the magnitude of the hydrophobic effect in biomolecules. The validity has been controversial, since the existence, physical origin, and magnitude of the volume-proportional term have been model dependent. In this paper, we cleared up this problem by using an accurate fused-hard sphere model and a perturbation scheme in which the compensation between the repulsive and attractive interactions has been clarified. The solvation free energy is shown to be dependent on the solute's surface area and curvature: the volume-proportional term is shown to be negligibly small. This disproves the basic assumption of the previous theories whose purpose is to “correct” the magnitude of the solvation free energy by subtracting volume-proportional terms. The relationship of our theory to previous theories is also discussed. © 1999 American Institute of Physics. [S0021-9606(99)50106-2]

I. INTRODUCTION

Quantitative information on solvation and transfer free energies is indispensable to the understanding and modeling of biomolecular structure and binding. A large amount of data on the oil/water partitioning and gas solubility (vapor/liquid partitioning) of small molecules has already been accumulated.¹ The transfer free energy² from phase 1 to 2 can be calculated from the partitioning data as:

$$\Delta G^* = RT \ln \left(\frac{\rho^{(1)}}{\rho^{(2)}} \right), \quad (1)$$

where $\rho^{(i)}$ s are the solute's number densities in phase i .³

Given the transfer data of small solutes, how can we estimate the solvation free energy of a protein in a given conformation state, or the change in free energy when the protein unfolds^{4,5}? It has been assumed^{5,6} in biochemistry that transfer free energies are proportional to the solvent accessible surface area (ASA) of the solute. The surface area parameters $\Delta G^*/\text{ASA}$ are determined from the experimental transfer free energy data of small solutes, and applied to various simulations of protein folding and protein–protein association.⁷

In the past decade, however, there have been several proposals that there is a volume-proportional term (much larger than the almost negligible pressure–volume (PV) work upon solute insertion) in the transfer free energy⁸ as well as an ASA-dependent contribution, and that the volume-proportional term should be subtracted from transfer free en-

ergies to obtain the surface area parameters (“correction”). Such proposals have resulted in the large change in the magnitude of solvation contributions for macromolecules⁸ and have been controversial.^{8–10} We thus come to an inevitable question: Is there indeed a volume-proportional term in transfer free energy? This paper is devoted to answering this question.

The existence of the volume-proportional term in the solvation free energy was first predicted by the Flory–Huggins (FH) theory,¹¹ where entropy of mixing is predicted to be volume proportional and the nonideality from the size differences between solute and solvent has been incorporated by using the lattice solution model. The lattice solution model has recently been extended to take the free volume into account and the extended theory also predicts the existence of the volume-proportional term.^{12,13} However, several facts show that FH theory is not a satisfying model. It is shown by Hall and co-workers¹⁴ that FH significantly underestimates the fused hard-sphere compressibility factor. Moreover, the FH agrees well in the chemical potential with the simulation in the lattice system.¹³ These facts show that the lattice system differs significantly from the continuum space in the excluded-volume effects. Some authors, furthermore, have reported that FH describes experimental partitioning data. However, due to the difference between the lattice and continuum space, these explanations should be reconsidered. There are some unnatural assumptions made in the explanation of experimental data by FH^{8,15} for example: (1) considering flexible segmental benzene and cyclohexane molecules in the experiments of De Young and Dill^{9,15} and (2) assuming fully packed and zero free volume limits in the explana-

^{a)}Electronic mail: seishi@bi.a.u-tokyo.ac.jp

tion of the solubility of xenon in a series of alkanes⁸ which made the physical meaning of the parameters unclear and made xenon a flexible 5.5 mer. The description of solvation and partitioning therefore requires a continuum space theory.

The existence of the volume-proportional term has also been predicted from gas-based theories. Hildebrand's theory¹⁶ is based on the van der Waals fluid model rather than a lattice model, but it nonetheless predicts the existence of a volume-proportional term similar to that predicted by FH theory. Sharp *et al.*¹⁷ used the ideal gas to derive the volume-proportional term. In the gas-based continuum-space theories,^{17–20} the mixing of free volume distributed to each solute and solvent molecule is the origin of the volume-proportional term.²¹ It is noteworthy, however, that neither the ideal gas nor the van der Waals gas model represents the excluded-volume effects of the real molecules.²² The difference between the lattice and the gas theories is that while the gas theory predicts the existence of the volume-proportional term even for the monomer solvent molecules, the lattice theory does not. On the other hand, Ben-Naim and Lovett have shown, based on virial theory, that the volume term contained in solvation free energy is PV work (which means this term is negligibly small at 1 atm). This result apparently disproves former gas-based "rederivations." However, their theory is valid only for monomers at a low-density limit.²³ The question as to the existence of a volume-proportional term of solvation for monomeric as well as chain molecules should be answered by a model that better represents the excluded-volume effects in solution.

The scaled particle theory (SPT)^{24–26} is based on the accurate analytical model of the hard-sphere system and is a far better model of the excluded volume effects for liquids than the theories based on the ideal gas or the van der Waals fluid, although it is limited to monomeric solutes and solvents. The hard-sphere chemical potential upon which SPT is based contains a term for the PV work performed when the solute is introduced into the system. Since the pressure of the hard-sphere system is usually as high as 1000 atm²⁵ at the density of water, this PV work amounts to several kilocalories. The way in which this term should be treated has not been clarified and has been a source of confusion ("hard-sphere pressure problem") in understanding the existence of the volume-proportional term of the solvation free energy.^{27,28} The scaled particle theory of Pierotti²⁵ ignores this hard-sphere PV term under the physical insight that PV work of real solvent is negligibly small: if one follows this approach, one concludes that there is no volume-proportional term in the entropy of solvation. Sharp *et al.*²⁸ follow this theory of the hard-sphere pressure. Neff and McQuarrie,²⁶ however, made straightforward use of the hard-sphere pressure in a different perturbation scheme and reported better agreement with the experimentally obtained solubility. Thus if one follows their treatment of the hard-sphere pressure, one concludes that there is a volume-proportional term. These seemingly contradicting conclusions result from the insufficiency in the treatment of the hard-sphere pressure and the unclear mechanism that makes up the real ambient pressure.

To resolve the controversy over the existence and mag-

nitude of the volume-proportional term in monomeric and chain molecular solution, we present in this paper a new theory of solvation based on a fused hard-sphere reference system in the continuum space and a novel perturbation scheme which describes the solvation shell for incorporating attractive interactions. This treatment is free from artifacts of the lattice. Our theory, furthermore, overcomes the hard-sphere pressure problem in SPT and the hard-soft compensation is treated rigorously. We show in this theory that the volume-proportional term is negligibly small even for chain solvents; the solvation free energy is dependent on surface area and curvature. This clarifies the dependence of transfer free energy on molecular size and shape and it gives a theoretical foundation for estimating the solvation effects in biomolecules.

II. THEORY

A. Hard-sphere contribution to the solvation free energy

Scaled particle theory²⁴ is equivalent to the analytical solution of the Percus–Yevick integral equation by Lebowitz,²⁹ a solution obtained through the compressibility equation, and it has a physical consistency at the limit of large solute size. The SPT provides one of the most realistic models of solvation for spherical molecular fluids. The chemical potential $\mu_{a,\text{hard}}$ of the hard-sphere solute of diameter R_a (at the number density ρ_a) in the solution of the hard-sphere solvent of diameter R_s of the number density ρ_s is given, in the infinite dilution, by the following equation²⁹:

$$\begin{aligned} \mu_{a,\text{hard}} &= \mu_{a,\text{hard}}^* + kT \ln \rho_a \\ \mu_{a,\text{hard}}^* &= P_{\text{hard}} \Delta V^m + kT \left(\frac{R_a}{R_s} \right)^2 \frac{3}{2} \frac{\eta(\eta+2)}{(1-\eta)^2} \\ &\quad + kT \left(\frac{R_a}{R_s} \right) \frac{3\eta}{1-\eta} - kT \ln(1-\eta), \end{aligned} \quad (2)$$

where $\mu_{a,\text{hard}}^*$ is the pseudochemical potential³ which expresses the free energy change upon insertion of a solute at a fixed position. The η in this equation is the packing fraction of the liquid ($\eta = \sum_i \frac{1}{6} \pi \rho_i R_i^3$), and P_{hard} and ΔV^m are the pressure of the hard-sphere reference system (which throughout this paper is called simply "hard-sphere pressure") and the hard-core volume of the solute:

$$P_{\text{hard}} = kT \rho_s \frac{1 + \eta + \eta^2}{(1 - \eta)^3}, \quad (3)$$

and

$$\Delta V^m = \frac{\pi}{6} R_a^3. \quad (4)$$

The treatment of the product $P_{\text{hard}} \Delta V^m$, which for the hydration of methane can be as large as 2 kcal/mol is comparable to the solvation free energy.²⁷ As is pointed out in Sec. I, the treatment of this term has been confused and the volume dependency has been unclear. To clarify the treat-

ment of the $P_{\text{hard}}\Delta V^m$ term, we introduced the attractive interactions and rigorously formulated their compensation with $P_{\text{hard}}\Delta V^m$.

B. Contribution of attractive interactions

We thus require a quantitative treatment of the compensation between the term which originates in the hard-sphere reference system (which in this paper is called the ‘‘hard term’’) and the term due to attractive interaction (the ‘‘soft term’’). To see the way in which the hard and soft terms compensate and make up the solvation free energy, we use the perturbation approach based on a hard-sphere system developed by Henderson and co-workers.³⁰ In the perturbation scheme, we decompose the attractive interaction contribution of the solvation free energy into the local (the solvation shell) and bulk contributions, and the bulk contribution is shown to compensate with $P_{\text{hard}}\Delta V^m$.

1. The perturbation approach of the solvation shell

Consider a solution of molecules comprising two molecular species: a (solute) and s (solvent), and assume infinite dilution so that solute–solute interactions are negligible. We also require that the intersubunit potential energy U_{N_a, N_s} is pairwise additive and each pair potential is decomposed into ‘‘hard sphere repulsion’’ u^{hard} and ‘‘soft’’ part u^{soft} , which expresses the contribution such as dispersion forces and electrostatic potentials.

In the dense fluids, the molecular structure of the fluids is basically determined by the hard-sphere repulsion. From the density functional perturbation theory proposed by Henderson, Evans and co-workers,³⁰ the free energy of the pure solvent of the density $\rho(\mathbf{r})$ is expressed as

$$A[\rho(\mathbf{r})] = A^{\text{hard}}[\rho(\mathbf{r})] + A^{\text{soft}}[\rho(\mathbf{r})], \quad (5)$$

$$A^{\text{soft}}[\rho(\mathbf{r})] = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \rho(\mathbf{r}) \rho(\mathbf{r}') g_{ss}^{\text{hard}}(\mathbf{r}, \mathbf{r}') u_{ss}^{\text{soft}}(\mathbf{r}, \mathbf{r}'),$$

where g_{ss}^{hard} and u_{ss}^{soft} express solvent–solvent correlation function of the hard-sphere reference system and solvent–solvent attraction, respectively. The free energy of the system after the insertion of a solute molecule to the fixed position (at the origin) in the fluid under constant volume (the relation between constant pressure insertion and constant volume insertion will become clear in Sec. II B 3) is expressed as:

$$A^*[\rho(\mathbf{r})] = A^{*,\text{hard}}[\rho(\mathbf{r})] + A^{*,\text{soft}}[\rho(\mathbf{r})],$$

$$A^{*,\text{soft}}[\rho(\mathbf{r})] = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \rho(\mathbf{r}) \rho(\mathbf{r}') g_{ss}^{*,\text{hard}}(\mathbf{r}, \mathbf{r}') u_{ss}^{\text{soft}}(\mathbf{r}, \mathbf{r}') + \int d\mathbf{r} u_{as}^{\text{soft}}(\mathbf{r}) \rho(\mathbf{r}) g_{as}^{*,\text{hard}}(\mathbf{r}), \quad (6)$$

where $g_{ss}^{*,\text{hard}}$ and $g_{as}^{*,\text{hard}}$, respectively, express correlation function between solute and solvent molecules, and between solvent molecules in the presence of a solute; u_{as}^{soft} expresses solvent–solvent attraction.

The chemical potential μ is the difference of the free energy of the above two states, which is expressed as the

sum of the hard sphere part μ_{hard} and the soft part μ_{soft} as $\mu = \mu_{\text{hard}} + \mu_{\text{soft}}$. The modeling of μ_{hard} has been discussed in Sec. II A. Our interest here is how we can express μ_{soft} . The result can be obtained as

$$\mu_{\text{soft}} = A^{*,\text{soft}}[\rho(\mathbf{r})] - A^{\text{soft}}[\rho(\mathbf{r})]$$

$$= A^{\text{direct}}(T, V, 1, N_s) + \Delta A^{\text{solvent}}(T, V, N_s),$$

$$A^{\text{direct}}(T, V, 1, N_s) = \int d\mathbf{r} u_{as}^{\text{soft}}(\mathbf{r}) \tilde{\rho}_{as}^*(\mathbf{r}), \quad (7)$$

$$\Delta A^{\text{solvent}}(T, V, N_s) = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' u_{ss}^{\text{soft}}(\mathbf{r}, \mathbf{r}') \times [\tilde{\rho}_{ss}^*(\mathbf{r}, \mathbf{r}') - \tilde{\rho}_{ss}(\mathbf{r}, \mathbf{r}')],$$

where $A^{\text{direct}}(T, V, 1, N_s)$ and $\Delta A^{\text{solvent}}(T, V, N_s)$, respectively, express solute–solvent interaction and the change of solvent–solvent interaction upon the insertion of a solute and the solute–solvent and solvent–solvent two-body distribution functions are defined, respectively, as $\tilde{\rho}_{as}^*(\mathbf{r}) = \rho(\mathbf{r}) g_{as}^{*,\text{hard}}(\mathbf{r})$, $\tilde{\rho}_{ss}(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r}) \rho(\mathbf{r}') g_{ss}^{\text{hard}}(\mathbf{r}, \mathbf{r}')$, and $\tilde{\rho}_{ss}^*(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r}) \rho(\mathbf{r}') g_{ss}^{*,\text{hard}}(\mathbf{r}, \mathbf{r}')$.

2. Solute–solvent interaction

$\tilde{\rho}_{as}^*(\mathbf{r})$ is zero in the range of $0 \leq |\mathbf{r}| \leq [(R_a + R_s)/2]$ because of the hard core repulsion. If the potential $u_{as}^{\text{soft}}(\mathbf{r})$ is a short-ranged potential, the direct interaction is expected to be proportional to ASA:

$$A^{\text{direct}}(T, V, 1, N_s) \approx 4\pi \left(\frac{R_a + R_s}{2} \right)^2 \Delta d \times \left[u_{as}^{\text{soft}} \left(\left| \mathbf{r} \right| = \frac{R_a + R_s}{2} \right) \tilde{\rho}_{as} \left(\left| \mathbf{r} \right| = \frac{R_a + R_s}{2} \right) \right],$$

where Δd is the range of the potentials u_{as}^{soft} . $A^{\text{direct}}(T, V, N_s)$ is thus proportional to ASA probed with the monomer of the solvent.

3. Solvent–solvent interaction

Now we rewrite $\Delta A^{\text{solvent}}$ using the binding energy. The binding energy in the presence of a fixed solute is defined as

$$B^*(\mathbf{r}) = \int d\mathbf{r}' \frac{\tilde{\rho}_{ss}^*(\mathbf{r}, \mathbf{r}') u_{ss}^{\text{soft}}(\mathbf{r}, \mathbf{r}')}{\tilde{\rho}_{as}^*(\mathbf{r})}, \quad (8)$$

and the bulk binding energy of the same component is given by

$$B^0 = \int d\mathbf{r}' \frac{\tilde{\rho}_{ss}(\mathbf{r}, \mathbf{r}') u_{ss}^{\text{soft}}(\mathbf{r}, \mathbf{r}')}{\rho_s}, \quad (9)$$

where $\rho_s = N_s/V$. B^0 is a constant due to the isotropicity of the system.

Using these binding energies, we can rewrite the $\Delta A^{\text{solvent}}(T, V, N_s)$ in Eq. (7) as

$$\Delta A^{\text{solvent}}(T, V, N_s) = \frac{1}{2} \left(\int d\mathbf{r} \tilde{\rho}_{\text{as}}^*(\mathbf{r}) B^*(\mathbf{r}) - \int d\mathbf{r} \rho_s B^0 \right). \quad (10)$$

The binding energy $B^*(\mathbf{r})$ far from the solute located at the origin tends to

$$\lim_{r \rightarrow \infty} B^*(\mathbf{r}) \equiv B^{*\infty}. \quad (11)$$

Note that compression of the solution upon the insertion of the solute makes $B^{*\infty}$ different from B^0 .

By using $B^{*\infty}$, we can rewrite $\Delta A^{\text{solvent}}(T, V, N_s)$ as

$$\Delta A^{\text{solvent}}(T, V, N_s) = \frac{1}{2} \int d\mathbf{r} \tilde{\rho}_{\text{as}}^*(\mathbf{r}) [B^*(\mathbf{r}) - B^{*\infty}] + \frac{1}{2} \int d\mathbf{r} [\tilde{\rho}_{\text{as}}^*(\mathbf{r}) B^{*\infty} - \rho_s B^0]. \quad (12)$$

Since the number of solvent molecules is constant upon insertion of the solute, it holds that

$$\int d\mathbf{r} \tilde{\rho}_{\text{as}}^*(\mathbf{r}) = \int d\mathbf{r} \rho_s, \quad (13)$$

and the fact that $B^{*\infty}$ is a constant can be used to rewrite Eq. (12) into a simple form:

$$\Delta A^{\text{solvent}}(T, V, N_s) = \frac{1}{2} \int d\mathbf{r} \tilde{\rho}_{\text{as}}^*(\mathbf{r}) [B^*(\mathbf{r}) - B^{*\infty}] + \frac{1}{2} \int d\mathbf{r} \rho_s (B^{*\infty} - B^0). \quad (14)$$

We show that the second term of Eq. (14) is equivalent to $P_{\text{soft}} \Delta V$ as follows: Matubayasi *et al.*³¹ by using the work of Lebowitz and Percus³² have shown that $\tilde{\rho}_{\text{as}}^*(\mathbf{r})$ and $\tilde{\rho}_{\text{ss}}^*(\mathbf{r}, \mathbf{r}')$ at \mathbf{r} and \mathbf{r}' far from the solute at the origin tend to:

$$\begin{aligned} \tilde{\rho}_{\text{as}}^*(\mathbf{r}) &\rightarrow \rho_s - \Delta V \frac{\partial \rho_s}{\partial V}, \\ \tilde{\rho}_{\text{ss}}^*(\mathbf{r}, \mathbf{r}') &\rightarrow \tilde{\rho}_{\text{ss}}(\mathbf{r}, \mathbf{r}') - \Delta V \frac{\partial \tilde{\rho}_{\text{ss}}(\mathbf{r}, \mathbf{r}')}{\partial V}, \end{aligned} \quad (15)$$

where $\Delta V = \partial \mu_{\text{hard}}^* / \partial P_{\text{hard}}$ is the hard-sphere excess volume of solvation. It follows from these relationships that

$$\frac{\tilde{\rho}_{\text{ss}}^*(\mathbf{r}, \mathbf{r}')}{\tilde{\rho}_{\text{as}}^*(\mathbf{r})} \rightarrow \frac{\tilde{\rho}_{\text{ss}}(\mathbf{r}, \mathbf{r}')}{\rho_s(\mathbf{r})} - \Delta V \frac{\partial}{\partial V} \left(\frac{\tilde{\rho}_{\text{ss}}(\mathbf{r}, \mathbf{r}')}{\rho_s(\mathbf{r})} \right) + o\left(\frac{1}{V}\right), \quad (16)$$

and the binding energy $B^{*\infty}$ then becomes

$$B^{*\infty} = B^0 - \Delta V \frac{\partial B^0}{\partial V}. \quad (17)$$

Thus the binding energy $B^{*\infty}$ is shown to be equal to the binding energy of the solution of the density $N_s / (V - \Delta V)$.

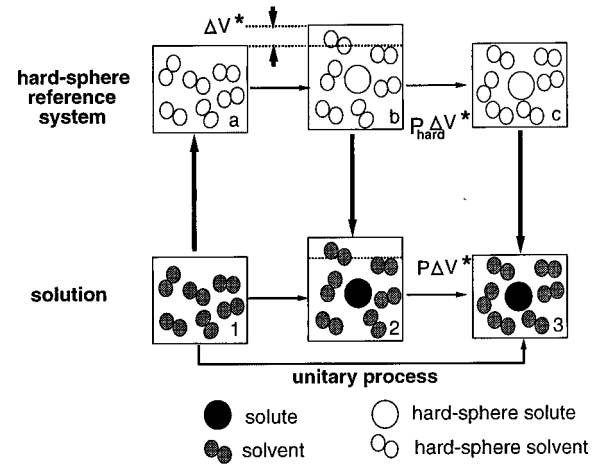


FIG. 1. Schematic summary of the perturbation scheme. Ben-Naim's unitary process (Ref. 3) 1→3, insertion of a solute at a fixed position in the solvent, is decomposed in the following thermodynamic cycle in terms of the fused-hard sphere (FHS) reference systems. 1→a: turn off attractive interactions with (T, V) fixed. a→b: insert a hard-sphere solute into the system of hard-sphere solution. The volume increment is ΔV , the hard-sphere excess volume of solvation. b→c: compress the FHS system to the volume of a. c→3: recover attractive interactions.

By using Eq. (17), we can rewrite the second term of Eq. (14) as

$$\begin{aligned} \frac{1}{2} \int d\mathbf{r} \rho_s (B^{*\infty} - B^0) &= -\Delta V \frac{1}{2} \int d\mathbf{r} \rho_s \frac{\partial B^0}{\partial V} \\ &= -\Delta V \frac{\partial}{\partial V} \frac{1}{2} \int d\mathbf{r} \rho_s B^0 \\ &= -\Delta V \frac{\partial A_{\text{soft}}}{\partial V} \\ &= P_{\text{soft}} \Delta V. \end{aligned} \quad (18)$$

The second term in Eq. (14), $P_{\text{soft}} \Delta V$, almost cancels out $P_{\text{hard}} \Delta V$ for liquid solutions, because $P = P_{\text{hard}} + P_{\text{soft}}$ and $P \Delta V$ is negligibly small. The first term in Eq. (14) ($\equiv \tilde{\mu}_{\text{solvent}}$) is then the leading term and is localized in the shell: $B^*(\mathbf{r})$ converges to $B^{*\infty}$ at a certain distance λ . When λ is short enough, it is proportional to the solute's coordination number and thus is dependent on surface area and curvature.

C. Extracting shell thermodynamical quantities from the hard-sphere reference system

Here we present the scheme for the compensation between hard and soft contributions based on the formulation of $P_{\text{soft}} \Delta V$ in Sec. II C. The results of this perturbation scheme are summarized schematically in Fig. 1, where μ^* ($= \mu_{\text{hard}}^* + \mu_{\text{soft}}$), μ_{hard}^* , and μ_{soft} , respectively, correspond to the free energy changes upon processes (1→3), (a→c), and (1→a) + (c→3). The main result of Sec. II C is $\mu_{\text{soft}} = \tilde{\mu}_{\text{soft}} + P_{\text{soft}} \Delta V$, where $\Delta V = (\partial \mu_{\text{hard}}^* / \partial P_{\text{hard}})_{T, N_s}$ is the

hard-sphere volume of solvation and $P_{\text{soft}}\Delta V$ is equal to $(2 \rightarrow b) + (c \rightarrow 3)$. The free energy change upon $a \rightarrow b$ is the noncompensating part calculated as

$$\tilde{\mu}_{\text{hard}} = \mu_{\text{hard}}^* - P_{\text{hard}}\Delta V, \quad (19)$$

and the terms $P_{\text{hard}}\Delta V$ and $P_{\text{soft}}\Delta V$ make up the ambient pressure term of $P\Delta V$ ($2 \rightarrow 3$). The volume of solvation ΔV in the hard-sphere reference system is calculated as

$$\begin{aligned} \Delta V &= \left(\frac{\partial}{\partial P_{\text{hard}}} (\mu_{a,\text{hard}} - kT \ln \rho_a) \right)_{T, N_s, N_a} \\ &= \Delta V^m + \Delta V^{\text{free}}, \end{aligned}$$

where ΔV^m is the molecular volume of the solute and ΔV^{free} is the free volume in the hard-sphere reference system:

$$\begin{aligned} \Delta V^{\text{free}} &= \frac{1}{kT} \frac{\frac{\pi}{6} R_s^3}{(1 + 4\eta + 4\eta^2)} \left[\left(\frac{R_a}{R_s} \right)^2 3(1 + 2\eta)(1 - \eta) \right. \\ &\quad \left. + \left(\frac{R_a}{R_s} \right) 3(1 - \eta)^2 + (1 - \eta)^3 \right]. \quad (20) \end{aligned}$$

Note that the free volume given by Eq. (20) is on the order of R_a^2 and is negligible in the $R_a \rightarrow \infty$ limit relative to ΔV^m , which is on the order of R_a^3 , and is physically plausible.

$\tilde{\mu}_{a,\text{hard}}$ is then calculated as

$$\begin{aligned} \tilde{\mu}_{a,\text{hard}} &= \mu_{a,\text{hard}}^* - P_{\text{hard}}\Delta V = kT \left(\frac{R_a}{R_s} \right)^2 \left(\frac{3}{2} \frac{\eta(\eta + 2)}{(1 - \eta)^2} \right) \\ &\quad + kT \left(\frac{R_a}{R_s} \right) \frac{3\eta}{1 - \eta} - kT \ln(1 - \eta) \\ &\quad - kT \eta \frac{1 + \eta + \eta^2}{(1 - \eta)^3} \left(\frac{R_a}{R_s} \right)^2 [3(1 + 2\eta)(1 - \eta)] \\ &\quad + \left(\frac{R_a}{R_s} \right) 3(1 - \eta)^2 + (1 - \eta)^3, \quad (21) \end{aligned}$$

where the last term comes from $-P_{\text{hard}}\Delta V^{\text{free}}$. It is noteworthy that the $\tilde{\mu}_{a,\text{hard}}$ obtained does not contain any $(R_a/R_s)^3$ term. This shows that there exists no volume-proportional terms in the solvation free energy, as will also be discussed in Sec. III.

D. Extension to chain molecular solvents

1. Soft term: Extension of the perturbation scheme

The solvation shell formalism of the soft term presented in Sec. II B can be extended to the chain molecular solvents. Appendix A generalizes the perturbation scheme into the chain molecular solvents. Equation (15) has served as the basis of our hard soft compensation scheme, and the generalization of this equation to chain molecular solvents is given

in Appendix B. It is shown in Appendix B that the fundamental relationship for the soft term, $\tilde{\mu}_{\text{soft}} = \mu_{\text{soft}} - P_{\text{soft}}\Delta V$ holds even for chain molecular solvents: we can dissect μ_{soft} into soft PV work and the solvation shell contributions ($\tilde{\mu}_{\text{soft}}$).

2. Hard term: A fused hard-sphere model of solvents

The thermodynamic quantities of the fused hard-sphere (FHS) system can be calculated by Wertheim's thermodynamic perturbation theory (TPT),³³ which gives an accurate analytical expression of the equation of state of FHS in comparison with simulations,³⁴ which has not been realized by Flory-Huggins. In TPT, a monomeric hard-sphere system is chosen as the reference system and intramolecular bonds are taken into account by first-order perturbation from the reference system as $A_{\text{FHS}} = A_{\text{HS}} + A_{\text{bond}}$, where A_{HS} is the free energy of the hard-sphere system and A_{bond} is the change of free energy upon the introduction of intramolecular bonds. This theory is shown to give an accurate description of the compressibility factor.³⁴ The pseudochemical potential of the monomer solute a in the FHS solvent s can be calculated from

$$\mu_{\text{hard}}^* = \mu_{\text{monomer}}^* + \mu_{\text{bond}}, \quad (22)$$

where μ_{bond} expresses the contribution of intramolecular bonds, which is given by $\mu_{\text{bond}} = (\partial A_{\text{bond}} / \partial N_a)_{T, V, N_s}$. Here we are interested in the polymerization of the solvent, which is given as

$$\frac{A_{\text{bond}}}{kT} = -N_s^c (m_s - 1) \ln \rho_s^c g_{\text{ss}}^{\text{HS}}(R_s) + N_s^c (m_s - 1), \quad (23)$$

where m_s is the number of monomers in a solvent molecule (which is assumed here to be a homopolymer for simplicity), $g_{\text{ss}}(R_s)$ is the two-body correlation function for monomers, and N_s^c and ρ_s^c are, respectively, the number and the number density of the solvent chain molecules. We have used the Lebowitz's compressibility solution result²⁹ for the two-body correlation functions. μ_{bond} in infinite dilution limit is given by,

$$\begin{aligned} \mu_{\text{bond}} &= -kT \left(\frac{R_a}{R_s} \right)^2 \frac{m_s - 1}{m_s} \frac{\frac{3}{2}\eta}{(1 - \eta)(1 - \frac{1}{2}\eta + \frac{1}{4}\eta^2)} \\ &\quad + P^{\text{bond}} \left(\frac{\pi}{6} R_a^3 \right), \quad (24) \end{aligned}$$

where

$$P^{\text{bond}} = -kT (m_s - 1) \rho_s^c \frac{1 + \eta + \frac{1}{4}\eta^2}{(1 - \eta)(1 - \frac{1}{2}\eta + \frac{1}{4}\eta^2)}.$$

Then we extract the noncompensating part from the chemical potential of a solute in the chain molecular solvents. From the discussion of Sec. II C, we know that the noncompensating hard term can be extracted as

$$\begin{aligned}
\tilde{\mu}_{\text{hard}} &= \mu_{\text{hard}}^* - P_{\text{hard}} \Delta V \\
&= kT \left(\frac{R_a}{R_s} \right)^2 \left(\frac{3}{2} \frac{\eta(\eta+2)}{(1-\eta)^2} - \frac{m_s-1}{m_s} \frac{3\eta}{2(1-\eta)(1-\frac{1}{2}\eta) + \frac{1}{4}\eta^2} \right) + kT \left(\frac{R_a}{R_s} \right) \frac{3\eta}{1-\eta} - kT \ln(1-\eta) \\
&\quad - kT \eta \frac{\frac{1+\eta+\eta^2}{(1-\eta)^3} - \frac{m_s-1}{m_s} \frac{1+\eta+\frac{1}{4}\eta^2}{(1-\eta)(1-\frac{1}{2}\eta+\frac{1}{4}\eta^2)}}{(1+4\eta+4\eta^2) - \frac{m_s-1}{m_s} (1-\eta)^2 \frac{1+2\eta-\frac{3}{2}\eta^2+\frac{1}{4}\eta^3+\frac{7}{16}\eta^4}{(1-\frac{1}{2}\eta+\frac{1}{4}\eta^2)^2}} \\
&\quad \times \left[\left(\frac{R_a}{R_s} \right)^2 \left(3(1+2\eta)(1-\eta) - \frac{m_s-1}{m_s} \frac{(1-\eta)^2(\frac{3}{2}-\frac{9}{8}\eta^2+\frac{3}{4}\eta^3)}{(1-\frac{1}{2}\eta+\frac{1}{4}\eta^2)^2} \right) + \left(\frac{R_a}{R_s} \right) 3(1-\eta)^2 + (1-\eta)^3 \right] \quad (25)
\end{aligned}$$

and, as in the case of the monomer solvent, is dependent on the surface area and curvature. The consequence of this generalization will be discussed in Sec. II E.

The theoretical scheme presented here can be generalized to the chain molecular solutes under the assumption that $\tilde{\mu}_{\text{soft}} = \mu_{\text{soft}} - P_{\text{soft}} \Delta V$ holds for chain solutes. The expression of $\tilde{\mu}_{\text{hard}}$ is given in Appendix C. In this case, it also appears likely that there are no solute-volume proportional terms.

E. Comparison with Flory χ

Here we briefly mention the relationship between our soft term and the Flory χ . The Flory χ term¹² is expressed in the following form: if we remove the assumption of zero free volume $\phi_a + \phi_s = 1$ and the assumption that the partial free volume is expressed as $\Delta V = m_a z$ as assumed in Ref. 12,

$$\begin{aligned}
\tilde{\mu}_\chi &= m_a z (\epsilon_{aa} \phi_a + \epsilon_{as} \phi_s) \\
&\quad - \frac{1}{2} z \Delta V (\epsilon_{aa} \phi_a^2 + 2\epsilon_{as} \phi_a \phi_s + \epsilon_{ss} \phi_s^2), \quad (26)
\end{aligned}$$

where ϵ_{ij} is the intersite contact interaction between molecular species i and j , and ϕ_i is the packing fraction of specie i . The first term comes from direct solute-solution interaction, which corresponds to A^{direct} in the continuum-space theory. The second term comes from the breaking of the intermolecular interactions by the introduction of a solute, and it corresponds to our ΔA^{inter} . But in the Flory χ , the same term also appears in $P_{\text{soft}} \Delta V$. The lattice theory has an artifact in that the ΔA^{inter} has the same form as $P_{\text{soft}} \Delta V$, the bulk quantity, and the shell localization of the interaction energy is not taken into account.

III. DISCUSSION

In this section, we used a rigorous perturbation scheme and showed that the solvation free energy of the spherical solute does not contain any volume-proportional terms (other than a small PV term) and is dependent on the solute's surface area and curvature. An influence of curvature on solvation free energy, which had been proposed previously,^{35,36} is supported by the present theoretical scheme, but we have not come up with a volume-proportional term like that expected for both monomer and chain solvents in the theories such as

FH, Hildebrand, and ideal gas. In Sec. III A, we clarify, in the light of our rigorous treatment of hard-soft compensation in FHS, the theoretical basis and assumptions in the previous gas models that lead one to expect a volume-proportional term to contribute to the solvation free energy. The way to overcome the hard-sphere pressure problem of SPT is discussed in Sec. III B. We then point out some deficiencies of the FH in Sec. III C and comment on the validity of the FH "correction" of the solvation free energy in Sec. III D.

A. Volume-proportional terms of solvation in gas models

1. Hildebrand's theory

Hildebrand "rederived" the FH type of volume-proportional term from van der Waals fluid model.¹⁶ But as pointed out by Chan and Dill,¹² Hildebrand's theory is based on an assumption that the molar volume, and thus the free volume, is proportional to the molecular volume. This leads to the volume-dependent entropy of solvation, but it fails for large solutes. In our theory, ΔV^{free} is dependent on surface area and curvature, and is not proportional to the molecular volume ΔV^m .

2. The theory of Sharp *et al.*

The origin of the volume-dependent term in the ideal gas theory of Sharp *et al.*¹⁷ is understood as follows: In taking the ideal gas limit of the "entropical part" of the formula $\tilde{\mu}_{\text{hard}} = \mu_{\text{hard}} - P_{\text{hard}} \Delta V$, instead of μ_{hard} and ΔV going to zero as expected, the volume ΔV remains constant and the partial molar volume determined experimentally is used. In this case, $\tilde{\mu}_{\text{hard}}$ goes to $-P_{\text{ideal}} \Delta V^{\text{molar}}$. This is the origin of the volume dependent term in the solvation free energy. The treatment violates our hard-soft compensation which takes place for the process of the volume increment by ΔV .

B. Volume dependency and attractive interactions in scaled particle theory

The volume dependency of the solvation free energy has not been clarified even in the framework of SPT, and is recognized^{27,28} as a critical problem in clarifying the

molecular-volume dependence of the solvation free energy. As mentioned in Sec. I, if one follows Pierotti's approach,²⁵ one concludes that there is no volume-dependent term in the entropy of solvation, but if one follows the approach of Neff and McQuarrie,²⁶ one concludes that a volume-dependent term exists. The two theories adopted different ways of taking attractive potentials into account. We have concluded Sec. III A that no volume proportional term contributes to the solvation free energy of the monomer solutes, except for a usually negligible PV term. What are the relationships between these three conclusions? They are summarized in the following compact form.

$$\mu = \mu_{\text{hard}} + \mu_{\text{soft}}(\text{Neff \& McQuarrie}) \quad (27)$$

$$= \tilde{\mu}_{\text{hard}} + \tilde{\mu}_{\text{soft}} + P\Delta V \approx \tilde{\mu}_{\text{hard}} + \tilde{\mu}_{\text{soft}}(\text{our theory}) \quad (28)$$

$$= \tilde{\mu}_{\text{hard}} + \tilde{\mu}_{\text{soft}} + \tilde{\mu}_{\text{res}} + P\Delta V \approx \tilde{\mu}_{\text{hard}} + \tilde{\mu}_{\text{soft}}(\text{Pierotti}), \quad (29)$$

where newly introduced quantities are defined as $\tilde{\mu}_{\text{hard}} = \mu_{\text{hard}} - P_{\text{hard}}\Delta V^m$, $\tilde{\mu}_{\text{soft}} = A^{\text{direct}}(T, V, 1, N_s)$ and $\tilde{\mu}_{\text{res}} = \Delta A^{\text{solvent}}(T, V, N_s) - P_{\text{hard}}\Delta V^{\text{free}}$ (definitions of each of the terms are given in Sec. II).

(i) The formalism of Neff and McQuarrie: Neff and McQuarrie²⁶ criticized Pierotti's way of simply ignoring $P_{\text{hard}}\Delta V^m$: they proposed instead to use the calculated value of the hard-sphere pressure and proposed further to introduce a term corresponding to our $\Delta A^{\text{solvent}}$. The only difference is that they have used the *partial derivative* $\Delta A^{\text{solvent}} = \partial A(T, V, N_s, N_a) / \partial N_a$ instead of our *difference* $\Delta A^{\text{solvent}} = A(T, V, N_s, N_a + 1) - A(T, V, N_s, N_a)$. In the partial derivative approach,²⁶ the solvation shell has not been treated. The volume-proportional term remained in the Neff and McQuarrie formalism only because the solvation shell of the soft term has not been treated in their formalism and thus the compensation between the soft and hard terms had not been treated. If we consider a solvation shell as we have done in Sec. II, the hard-soft compensation enables us to conclude that solvation free energy does not contain any volume terms, as shown in Eq. (35), which was beyond their scope.

(ii) The formalism of Pierotti: Pierotti's SPT²⁵ is based on an assumption that $P_{\text{hard}}\Delta V^m$ always compensates the corresponding soft part. The attractive interactions are taken into account by perturbation, but in a manner different from the formalism of Neff and McQuarrie: only direct solute-solvent interaction A^{direct} is taken into account. As shown in Eq. (29), Pierotti's formalism can be derived from ours by assuming $\tilde{\mu}_{\text{res}} \approx 0$, which is equivalent to $\Delta A^{\text{solvent}}(T, V, N_s) \approx P_{\text{hard}}\Delta V^{\text{free}}$. This means that upon the insertion of a solute, the attractive part of the change of solvent-solvent interaction always compensates the free volume expansion by hard-sphere pressure. The validity of this assumption will be discussed in a forthcoming paper,³⁷ but the success of Pierotti's formalism suggests that $\tilde{\mu}_{\text{res}} \approx 0$ is not a bad approximation.

(iii) The theory of Sharp *et al.*: The hard-sphere theory of Sharp *et al.*²⁸ had presented no theoretical treatment of the soft term, but their scheme of solvation free energy is principally equivalent to Pierotti's. They have made an assumption that the volume change upon solvation, namely the sol-

ute's molar volume, is equal to $\Delta V^m / \eta$. This, however, is equivalent to assuming molar volume to be molecular-volume proportional and has the same deficiency as that in Hildebrand's theory.

C. Comparison to Flory-Huggins theory

1. Volume dependency in the Flory-Huggins theory

Although FH theory predicts the existence of a volume-proportional term in water \rightarrow oil transfer due to the flexibility of the oil solvents, this flexibility does not lead (in our hard-sphere chain theory) to a molecular-volume-proportional term. The inconsistency between FH and our theory is discussed here in detail.

The configurational part of the FH chemical potential is expressed as¹²

$$\begin{aligned} \tilde{\mu}_{\text{hard}}^{\text{FH}} = kT \frac{\Delta V_{\text{FH}}}{V} [N_a(m_a - 1) + N_s(m_s - 1)] \\ + kT \Delta V_{\text{FH}}^{\text{free}} \ln(1 - \eta), \end{aligned} \quad (30)$$

where ΔV_{FH} and $\Delta V_{\text{FH}}^{\text{free}}$ express the excess volume and the free volume in the FH model, and η is the packing fraction. The FH theory imposes several assumptions on Eq. (30), whose validity will be examined later in this subsection: it assumes that $\Delta V_{\text{FH}} = m_a$, equivalently $\Delta V_{\text{FH}}^{\text{free}} = 0$ and $\eta = 1$. This equation then becomes, in the infinite dilution limit,

$$\tilde{\mu}_{\text{hard}}^{\text{FH}} = kT m_a \left(1 - \frac{1}{m_s} \right). \quad (31)$$

It is this equation that has been used by many authors^{12,13,15} to model solvation and transfer experiments.

The validity of concluding the existence of a volume-proportional term from Eq. (31) of the FH theory should be questioned. Our theory contains the $1 - (1/m_s)$ term, which is the same as FH in the case $m_a = 1$ (if we can further assume that the solvation shell formalism presented in Sec. IID can also be generalized to the chain molecular solutes, we have the term $m_a[1 - (1/m_s)]$ in $\tilde{\mu}_{a,\text{hard}}$), and these terms are in the curvature- and surface area-proportional terms, not in the volume-proportional terms. The lattice FH theory, on the other hand, is based on a prerequisite that all subunits are of the same size, which is equivalent to requiring $R_a/R_s = 1$, and in this case there is no distinction between R_a/R_s , $(R_a/R_s)^2$, and $(R_a/R_s)^3$. Thus it is insufficient to conclude that $\tilde{\mu}_{\text{hard}}$ in FH is a volume-proportional term.

2. Assumptions in the Flory-Huggins theory

We then examine the validity of the assumptions made to derive Eq. (31) from (30). In Fig. 2, the solvent chain length dependence of $\tilde{\mu}_{\text{hard}}$ for monomer solutes is plotted. The packing fraction was kept constant at $\eta = 0.5$ for the fused-hard-sphere model. The monomer radius ratio $r = R_a/R_s$ has also varied in the case of the fused-hard-sphere model. Very small chain-length dependence is observed in $\tilde{\mu}_{a,\text{hard}}$ for the fused-hard-sphere model; there is only a slight tendency of $\tilde{\mu}_{a,\text{hard}}$ decreasing with chain length m . This is intuitively plausible since when the packing fraction is con-

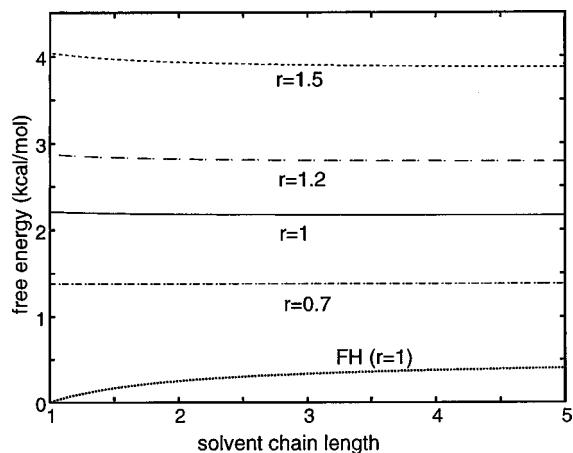


FIG. 2. $\tilde{\mu}_{\text{hard}}$ from fused hard-sphere theory vs solvent chain length for monomer solutes with various monomer solutes at several monomer diameter ratios ($r=R_o/R_s$) when $\eta=0.5$. See the text for an explanation of the opposite tendency for the fused hard-sphere and Flory–Huggins theories.

stant, the connectivity of the solvents contributes to cavity opening. But the FH theory predicts the opposite. The wrong chain-length dependency originates in the assumptions made to derive Eq. (31) from Eq. (30). If $\Delta V_{\text{FH}}^{\text{free}}$ increases with the solvent chain length, FH is then expected to have the same tendency as that of the fused hard spheres.

Figure 3 plots $\tilde{\mu}_{\text{hard}}$ upon insertion of methane in liquid alkanes from hexane to hexadecane, the case where the radius of the solute and monomer of the solvents are the same as required in the derivation of FH, the experimental packing fractions of each solvent.³⁸ Although the FH theory [Eq. (31)] contains unnatural assumptions, as pointed out above, it predicts the same tendency of the chain-length-dependence of $\tilde{\mu}_{\text{hard}}$ as the fused hard-sphere theory.

3. Entropy of coupling and volume proportionality

Chan and Dill¹² have discussed that the contact free energy can be extracted from solvation free energy by subtract-

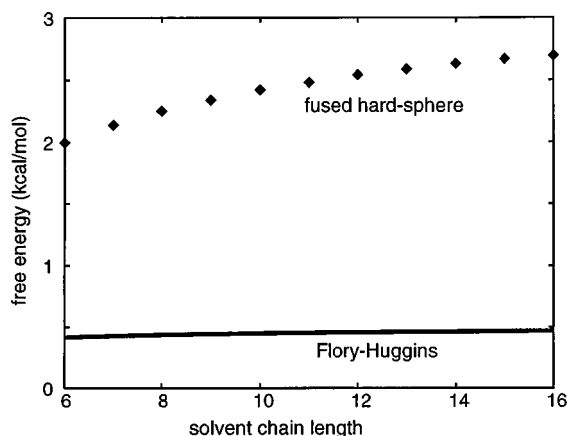


FIG. 3. Solvent (=liquid alkane) chain length dependence of the $\tilde{\mu}_{\text{hard}}$ calculated from fused hard-sphere and Flory–Huggins (FH) theories. In spite of the unrealistic assumptions of the Flory–Huggins theory, it predicts at least the same tendency on the chain-length dependence as the fused hard-sphere theory.

ing the contributions of the coupling of translational degrees of freedom to excluded volume, or rotational or internal degrees of freedom (“entropy of coupling”). However, what they have actually calculated was the *whole* entropic term upon insertion of a solute (see Ref. 22 and the discussion in Sec. III D) which includes the entropy of coupling. Since FHS agrees well with the Monte Carlo simulation for hard-sphere chain fluids,³⁴ our $\tilde{\mu}_{\text{hard}}$ already contains the coupling effect for the flexible polymers giving a better description of the excluded volume effects than FH. Since our $\tilde{\mu}_{\text{hard}}$ does not contain any volume-proportional term, it is unlikely that the entropy of coupling for the flexible polymers introduces the volume-proportional term. The case of rigid polymers will be discussed in the forthcoming paper.

D. Definition of solvation free energy

Here we will briefly comment on the definition of solvation free energy. Our theory does not support the “correction” of the solvation free energy or the procedures based on FH theory to obtain solute/solvent contact interactions for the following two reasons: first, the necessity of the “correction” is based on the preposition that there *is* a volume-dependent term in the solvation free energy and that to obtain “transfer free energy that depends only on solute/solvent contact area,” the volume-proportional term should be subtracted.¹³ However, the hard-sphere-chain-based theory presented in this paper did not lead to the existence of the volume-proportional term (except the small PV term). Second, the “correction” would lose solute-induced change in solvents from the obtained solvation free energy. The entropic term due to solute-induced change in the solvent’s configuration should be included in the contact free energy¹² but is lost in the “corrected” free energy.

IV. CONCLUSION

To clarify whether a solute’s volume-proportional term exists in the solvation free energy, we have presented a fused-hard-sphere-based perturbation theory of solvation. It is based on an accurate model of chain molecular fluids in the continuum space,³³ and it clarifies the compensation of contributions from the hard sphere and attractive interactions, which have not been clarified in previous scaled-particle theories. Flory–Huggins theory had predicted the existence of the volume-proportional term from which the “correction” of the solvation free energy [i.e., to subtract volume-proportional terms from ΔG^* in Eq. (1)] has been proposed.¹³ This has caused a serious controversy for a decade, but our theory has shown that the solvation free energy is dependent on surface area and curvature, and that no volume-proportional terms contribute to solvation free energy (except for an ambient pressure PV term that is negligible for liquids). This disproves the proposition of subtracting the volume-proportional terms. We think that this model is useful in modeling solvation and transfer processes, and a forthcoming paper³⁷ is devoted to the application of this theory to the modeling of transfer experiments.

ACKNOWLEDGMENTS

The authors thank Professor Kuniaki Nagayama, Professor Tatsuo Ooi, Professor Ken A. Dill, Professor Richard Friesner, Professor Kim Sharp, and Dr. B. K. Lee for valuable discussions. One of the authors (S.S.) is grateful to Kaori Sakurai for continuous encouragement.

APPENDIX

In this Appendix we extend the solvation shell formalism of the soft term presented in Sec. II B to chain molecular solvent. We first briefly present the extension of the perturbation scheme of Sec. II B 1 to the chain molecule. We then extend Eq. (15), which has served as the basis of solvation shell formalism, to intersubunit distribution functions of the chain molecules. This is done by extending the discussion of Matubayasi *et al.*,³¹ which has been focused on monomer solvents, to chain solvents. The discussion presented here serves as the foundation for extending the theory of size dependence to chain solvents as is done in Sec. III C.

1. Perturbation scheme for chain solvents

Consider a fluid of N_s solvent molecules which consists of m_s distinguishable subunits. The system is homogeneous and isotropic. The solvent is confined in volume V at temperature T . With V and T kept constant, we introduce a spherical solute a into the solution, fixed at an arbitrary chosen origin. Since the subunits are assumed to be distinguishable, we must deal with (for two-body distribution functions) all the possible pairs of subunits, i.e., intramolecular and intermolecular pairs. The perturbation scheme is readily extended to this case, under the assumption that the reference can be taken as a hard-sphere system, which as a result gives the following form for $\Delta A^{\text{solv}}(T, V, N_s)$:

$$\Delta A^{\text{solv}}(T, V, N_s) = \frac{1}{2} \sum_{i,j} \sum_{\kappa=1}^2 \int d\mathbf{r}_1 d\mathbf{r}_2 u_{1i,\kappa j}^{\text{soft}}(\mathbf{r}_1, \mathbf{r}_2) \times [\tilde{\rho}^{*(1i,\kappa j)}(\mathbf{r}_1, \mathbf{r}_2) - \tilde{\rho}^{(1i,\kappa j)}(\mathbf{r}_1, \mathbf{r}_2)], \quad (\text{A1})$$

where $\tilde{\rho}^{*(1i,\kappa j)}(\mathbf{r}_1, \mathbf{r}_2)$ and $\tilde{\rho}^{(1i,\kappa j)}(\mathbf{r}_1, \mathbf{r}_2)$ express the two-body distribution function between the i th subunit of molecule 1 and the j th subunit of subunit in κ , with and without the solute, respectively.

We then give a chain solvent version of $\Delta A^{\text{solv}}(T, V, N_s)$ expressed in terms of the binding energy B . This time the concept of binding energy is extended as follows: we consider a binding energy of subunit i in the ‘‘sea’’ of subunit j . Here j may be either on the molecule (the same as i) or on the other molecule. We distinguish between these two cases by using the parameter κ introduced above. The binding energies are defined as

$$B^{*(1i,\kappa j)}(\mathbf{r}_1) = \int d\mathbf{r}_2 \frac{\tilde{\rho}^{*(1i,\kappa j)}(\mathbf{r}_1, \mathbf{r}_2) u_{1i,\kappa j}^{\text{soft}}(\mathbf{r}_1, \mathbf{r}_2)}{\tilde{\rho}^{*(1i)}(\mathbf{r}_1)} \quad (\text{A2})$$

and

$$B^{0(1i,\kappa j)} = \int d\mathbf{r}_2 \frac{\tilde{\rho}^{(1i,\kappa j)} u_{1i,\kappa j}^{\text{soft}}(\mathbf{r}_1, \mathbf{r}_2)}{\rho^{(1i)}}. \quad (\text{A3})$$

$\Delta A^{\text{solv}}(T, V, N_s)$ can then be expressed in terms of the binding energies as

$$\Delta A^{\text{solv}}(T, V, N_s) = \frac{1}{2} \sum_{i,j} \sum_{\kappa=1}^2 \int d\mathbf{r}_1 \tilde{\rho}^{*,1i}(\mathbf{r}_1) \times [B^{*(1i,\kappa j)}(\mathbf{r}_1) - B^{*(1i,\kappa j)\infty}] + \frac{1}{2} \sum_{i,j} \sum_{\kappa=1}^2 \int d\mathbf{r}_1 \tilde{\rho}^{1i} \times [B^{*(1i,\kappa j)\infty} - B^{0(1i,\kappa j)}], \quad (\text{A4})$$

and the first term is surface area and curvature proportional when B^* covers to $B^{*\infty}$ in a short distance. To show that the second term is equal to $P_{\text{soft}} \Delta V$, we need to generalize Eq. (15) to the chain molecular solvent. The next section does this.

2. Solvation shell formalism for chain solvents

Here we investigate the asymptotic behavior of the distribution functions far from the solute. We will omit the sign s (=solvent) from the indices below. Three distributions are of special interests: (i) one-body, (ii) intermolecular two-body, and (iii) intramolecular two body correlation functions. The Ursell correlation functions are given as

$$F_i(\mathbf{r}_{1\alpha}) = \frac{\delta}{\delta a(\mathbf{r}_{1\alpha})} \ln \left(\int D_{\text{all}} e^{-\beta U_{N_s}} \prod_{i=1}^{N_s} \prod_{j=1}^{m_s} [1 + a(\mathbf{r}_{ij})] \right)_{a(\mathbf{r}=0)},$$

$$F_{i,j}(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) = \frac{\delta^2}{\delta a(\mathbf{r}_{1\alpha}) \delta a(\mathbf{r}_{2\beta})} \times \ln \left(\int D_{\text{all}} e^{-\beta U_{N_s}} \prod_{i=1}^{N_s} \prod_{j=1}^{m_s} [1 + a(\mathbf{r}_{ij})] \right)_{a(\mathbf{r}=0)}, \quad (\text{A5})$$

$$F_{\{i,j\}}(\mathbf{r}_{1\alpha}, \mathbf{r}_{1j}) = \frac{\delta^2}{\delta a(\mathbf{r}_{1\alpha}) \delta a(\mathbf{r}_{1j})} \times \ln \left(\int D_{\text{all}} e^{-\beta U_{N_s}} \prod_{i=1}^{N_s} \prod_{j=1}^{m_s} [1 + a(\mathbf{r}_{ij})] \right)_{a(\mathbf{r}=0)},$$

where U_{N_s} is the sum of the solvent–solvent interaction. The correlation functions tend, in the presence of a solute, to

$$\begin{aligned}\bar{F}_i(\mathbf{r}_{1\alpha}) &= \frac{\delta}{\delta a(\mathbf{r}_{1\alpha})} \ln \left(\int D_{\text{all}} e^{-\beta(U_{N_s} + U_{uv})} \prod_{i=1}^{N_s} \prod_{j=1}^{m_s} [1 + a(\mathbf{r}_{ij})] \right)_{a(\mathbf{r}=0)} \\ \bar{F}_{i,j}(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) &= \frac{\delta^2}{\delta a(\mathbf{r}_{1\alpha}) \delta a(\mathbf{r}_{2\beta})} \ln \left(\int D_{\text{all}} e^{-\beta(U_{N_s} + U_{uv})} \prod_{i=1}^{N_s} \prod_{j=1}^{m_s} [1 + a(\mathbf{r}_{ij})] \right)_{a(\mathbf{r}=0)} \\ \bar{F}_{\{i,j\}}(\mathbf{r}_{1\alpha}, \mathbf{r}_{1j}) &= \frac{\delta^2}{\delta a(\mathbf{r}_{1\alpha}) \delta a(\mathbf{r}_{1j})} \ln \left(\int D_{\text{all}} e^{-\beta(U_{N_s} + U_{uv})} \prod_{i=1}^{N_s} \prod_{j=1}^{m_s} [1 + a(\mathbf{r}_{ij})] \right)_{a(\mathbf{r}=0)},\end{aligned}\quad (\text{A6})$$

where U_{uv} is the sum of the solute–solvent interaction. The relationship between the distribution function in the presence and that in the absence of the solute is the main concern of this Appendix. We first show in detail how the solvent solvent distribution function *in the solution* $\bar{F}_{i,j}(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta})$ can be related to the quantity *in the solvent* $F_{i,j}(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta})$ when both $\mathbf{r}_{1\alpha}$ and $\mathbf{r}_{2\beta}$ are sufficiently far from the solute. First we manipulate $\bar{F}_{i,j}(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta})$ as follows:

$$\begin{aligned}\bar{F}_{i,j}(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) &= \frac{\delta^2}{\delta a(\mathbf{r}_{1\alpha}) \delta a(\mathbf{r}_{2\beta})} \ln \left(\int D_{\text{all}} e^{-\beta(U_{N_s} + U_{uv})} \prod_{i=1}^{N_s} \prod_{j=1}^{m_s} [1 + a(\mathbf{r}_{ij})] \right)_{a(\mathbf{r}=0)} \\ &= e^{-\beta[u(\mathbf{r}_{1\alpha}) + u(\mathbf{r}_{2\beta})]} \frac{\delta^2}{\delta a(\mathbf{r}_{1\alpha}) \delta a(\mathbf{r}_{2\beta})} \ln \left(\int D_{\text{all}} e^{-\beta U_{N_s}} \prod_{i=1}^{N_s} \prod_{j=1}^{m_s} [1 + a(\mathbf{r}_{ij})] \right)_{a(\mathbf{r})=f(\mathbf{r})} \\ &= e^{-\beta[u(\mathbf{r}_{1\alpha}) + u(\mathbf{r}_{2\beta})]} \frac{\delta^2}{\delta a(\mathbf{r}_{1\alpha}) \delta a(\mathbf{r}_{2\beta})} \\ &\quad \times \left(\sum_{n=1}^{\infty} \sum'_{\{i_2, i_3, \dots, i_n\}} \sum_{j_1} \cdots \sum_{j_n} \int d\mathbf{r}_{1j_1} d\mathbf{r}_{2j_2} \cdots d\mathbf{r}_{i_n j_n} a(\mathbf{r}_{1j_1}) \cdots a(\mathbf{r}_{i_n j_n}) F(\mathbf{r}_{1j_1}, \dots, \mathbf{r}_{i_n j_n}) \right)_{a(\mathbf{r})=f(\mathbf{r})}.\end{aligned}\quad (\text{A7})$$

Here $\sum'_{i_2, i_3, \dots, i_n}$ is defined as follows: First let ν_1 indices $i_1, i_2, \dots, \Omega_1 \equiv i_{\nu_1}$ equal 1 and the next $\Omega_2 \equiv \nu_2 - \nu_1$ indices equal 2 and so on. Ω_α expresses the number of indices $\{i_k\}$ equal to α . The sum is taken over a set of $\{i_k\}$ under the condition that $\Omega_1 \geq \Omega_2 \geq \cdots \geq \Omega_{N_s}$. If $\Omega_1 = \Omega_2 = \cdots = \Omega_k$, the factor $1/k!$ will automatically be multiplied.

Carrying out the above functional derivative leads to

$$\begin{aligned}\bar{F}_{i,j}(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) &= e^{-\beta[u(\mathbf{r}_{1\alpha}) + u(\mathbf{r}_{2\beta})]} F(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) + e^{-\beta[u(\mathbf{r}_{1\alpha}) + u(\mathbf{r}_{2\beta})]} \\ &\quad \times \sum_{n=1}^{\infty} \sum'_{\{I_1, I_2, \dots, I_n\}} \sum_{j_1} \cdots \sum_{j_n} \int d\mathbf{x}_{I_1 j_1} \cdots d\mathbf{x}_{I_n j_n} f(\mathbf{x}_{I_1 j_1}) \cdots f(\mathbf{x}_{I_n j_n}) F(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}, \mathbf{x}_{I_1 j_1}, \dots, \mathbf{x}_{I_n j_n}).\end{aligned}\quad (\text{A8})$$

When $\mathbf{r}_{1\alpha}$ and $\mathbf{r}_{2\beta}$ are far from the solute, then $e^{-\beta[u(\mathbf{r}_{1\alpha}) + u(\mathbf{r}_{2\beta})]} = 1$ holds. And the contribution from the integral part of Eq. (A8) is confined in the region around the solute, where $f[(\mathbf{x})]$ s have a nonzero value. So, the two sets of points $\{\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}\}$ and $\{\mathbf{x}_{1j_1}, \dots, \mathbf{x}_{i_n j_n}\}$ are far apart and $(I_k)_{k=1, \dots, n} \neq 1, 2$. It is in this condition that the finite volume correction formula of Lebowitz and Percus³² can be used (it is clear from the derivation that it is not limited to the monomer solution). The set $\{I_1, \dots, I_n\}$ does not coincide with 1 and 2. At this time $\Sigma'' = \Sigma'$. $\bar{F}_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta})$ then becomes

$$\begin{aligned}\bar{F}_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) &\rightarrow F_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) + kT \frac{\partial V}{\partial P} \frac{\partial F_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta})}{\partial V} \\ &\quad \times \sum_{n=1}^{\infty} \sum'_{\{i_2, \dots, i_n\}} \sum_{j_1} \cdots \sum_{j_n} \int d\mathbf{x}_{1j_1} \cdots d\mathbf{x}_{i_n j_n} f(\mathbf{x}_{1j_1}) \cdots f(\mathbf{x}_{i_n j_n}) \frac{\partial F(\mathbf{x}_{1j_1}, \dots, \mathbf{x}_{i_n j_n})}{\partial V} \\ &= F_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) - \frac{\partial V}{\partial P} \frac{\partial \mu^*}{\partial V} \frac{\partial F_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta})}{\partial V},\end{aligned}\quad (\text{A9})$$

where the last line comes from the cluster expansion form of the chemical potential. Manipulating partial derivatives in the last line leads to the conclusion

$$\bar{F}_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) \rightarrow F_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta}) - \Delta V \frac{\partial F_i(\mathbf{r}_{1\alpha}, \mathbf{r}_{2\beta})}{\partial V}. \quad (\text{A10})$$

This is in the same form as in the case of monomer solvents. Repeating the same procedure we obtain similar results for

other Ursell function, and simple tedious calculation leads to a relation the same as Eq. (15).

3. The case of chain molecular solutes

Here we give, under the assumption that the solvation shell formalism of the soft contribution $\tilde{\mu}_{\text{soft}} = \mu_{\text{soft}} - P_{\text{soft}} \Delta V$ holds even for the chain molecular solutes, the noncompensating hard contribution for the hard term of chain molecular solutes composed of m_a subunits. Repeating the similar procedure as before, we obtain

$$\begin{aligned} \tilde{\mu}_{\text{hard}} &= \mu_{a,\text{FHS}} - P_{\text{FHS}} \Delta V \\ &= kT m_a \left(\frac{R_a}{R_s} \right)^2 \left(\frac{3}{2} \frac{\eta(\eta+2)}{(1-\eta)^2} - \frac{m_s-1}{m_s} \frac{3\eta}{2(1-\eta)(1-\frac{1}{2}\eta) + \frac{1}{4}\eta^2} \right) \\ &\quad + kT m_a \left(\frac{R_a}{R_s} \right) \frac{3\eta}{1-\eta} - kT \ln(1-\eta) - kT(m_a-1) \ln g_{aa}^{\text{HS}} \\ &\quad - kT \eta \frac{\frac{1+\eta+\eta^2}{(1-\eta)^3} - \frac{m_s-1}{m_s} \frac{1+\eta+\frac{1}{4}\eta^2}{(1-\eta)(1-\frac{1}{2}\eta+\frac{1}{4}\eta^2)}}{(1+4\eta+4\eta^2) - \frac{m_s-1}{m_s} (1-\eta)^2 \frac{1+2\eta-\frac{3}{2}\eta^2+\frac{1}{4}\eta^3+\frac{7}{16}\eta^4}{(1-\frac{1}{2}\eta+\frac{1}{4}\eta^2)^2}} \\ &\quad \times \left[m_a \left(\frac{R_a}{R_s} \right)^2 \left(3(1+2\eta)(1-\eta) - \frac{m_s-1}{m_s} \frac{(1-\eta)^2(\frac{3}{2}-\frac{9}{8}\eta^2+\frac{3}{4}\eta^3)}{(1-\frac{1}{2}\eta+\frac{1}{4}\eta^2)^2} \right) + m_a \left(\frac{R_a}{R_s} \right) 3(1-\eta)^2 \right. \\ &\quad \left. + (1-\eta)^3 \left(1 - (m_a-1) \frac{2(1-\eta)^2 + 3\left(\frac{R_a}{R_s}\right)(1+\eta)(1-\eta) + \frac{3}{2}\left(\frac{R_a}{R_s}\right)^2 \eta(2+\eta)}{2(1-\eta)^2 + 3\left(\frac{R_a}{R_s}\right)\eta(1-\eta) + \frac{3}{2}\left(\frac{R_a}{R_s}\right)^2 \eta^2} \right) \right]. \quad (\text{A11}) \end{aligned}$$

As before, $\tilde{\mu}_{a,\text{FHS}}$ is composed of surface and curvature contributions.

¹A. Leo, C. Hansch, and D. Elkins, *Chem. Rev.* **71**, 525 (1971); J.-L. Fauchere and V. Pliska, *Eur. J. Med. Chem.* **18**, 369 (1983); A. Radzicka and R. Wolfenden, *Biochem. J.* **27**, 1664 (1988); M. H. Abraham, H. S. Chadka, G. S. Whiting, and C. Mitchell, *J. Pharm. Sci.* **83**, 1085 (1994).

²The transfer of a solute from oil to water is equivalent to the difference of the solubility of the gas solute into water and oil, or the difference in *solvation free energy* of a solute in two phases. Thus in this paper, we concentrate on solvation processes, from which the transfer free energies between the two solvent phases are calculated.

³A. Ben-Naim, *Statistical Thermodynamics for Chemists and Biochemists* (Plenum, New York, 1992).

⁴Preliminary results have been published in (a) S. Shimizu, M. Ikeguchi, and K. Shimizu, *Chem. Phys. Lett.* **282**, 79 (1998); S. Shimizu, M. Ikeguchi, S. Nakamura, and K. Shimizu, *ibid.* **284**, 235 (1998).

⁵D. Eisenberg and A. D. McLachlan, *Nature (London)* **319**, 199 (1986).

⁶T. Ooi, M. Oobatake, G. Nemethy, and H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **84**, 30836 (1987).

⁷D. Eisenberg, M. Wesson, and M. Yamashita, *Chem. Scr.* **29A**, 217 (1989); N. Horton and M. Lewis, *Protein Sci.* **1**, 169 (1992); B. von

Freyberg, T. J. Richmond, and W. Braun, *J. Mol. Biol.* **233**, 275 (1993); M. D. Cummings, T. N. Hart, and R. J. Read, *Protein Sci.* **4**, 2087 (1996).

⁸Reviews of this topic are (a) H. S. Chan and K. A. Dill, *Annu. Rev. Biophys. Biomol. Struct.* **26**, 425 (1997); (b) J. M. Sanchez-Ruiz, *Eur. Biophys. J.* **24**, 261 (1996); (c) A. Ben-Naim, *Curr. Sci.* **4**, 264 (1994).

⁹A. Holtzer, *Biopolymers* **32**, 711 (1992); **34**, 315 (1994).

¹⁰A. Ben-Naim and R. M. Mazo, *J. Phys. Chem.* **97**, 10829 (1993); *J. Phys. Chem.* **101**, 11221 (1997).

¹¹P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, NY, 1953).

¹²H. S. Chan and K. A. Dill, *J. Chem. Phys.* **101**, 7007 (1994).

¹³S. Kumar, I. Szleifer, K. A. Sharp, P. J. Rossky, R. Friedman, and B. Honig, *J. Phys. Chem.* **99**, 8382 (1995).

¹⁴R. Dickman and C. K. Hall, *J. Chem. Phys.* **85**, 4108 (1996).

¹⁵L. R. De Young and K. A. Dill, *Biochem. J.* **27**, 5281 (1998); L. R. De Young and K. A. Dill, *J. Phys. Chem.* **94**, 801 (1990).

¹⁶J. H. Hildebrand, *J. Chem. Phys.* **15**, 225 (1947).

¹⁷K. A. Sharp, A. Nicholls, R. Friedman, and B. Honig, *Biochem. J.* **30**, 9686 (1991).

¹⁸I. Tunon, E. Silla, and J. L. Pascual-Ahuir, *J. Phys. Chem.* **98**, 377 (1994).

¹⁹D. J. Giesen, C. J. Cramer, and D. G. Truhlar, *J. Phys. Chem.* **98**, 4141 (1994).

²⁰J. M. Sanchez-Ruiz, *J. Phys. Chem.* **99**, 12076 (1995).

- ²¹ There has even been confusion in the ideal gas theory on the definition of solvation free energy, which subtracts volume-dependent terms to obtain solvation free energy. They are shown to be the result of choosing different standard state, and the obtained "solvation free energy" is different from Ben-Naim's unitary free energy. (a) B. Lee, *Biophys. Chem.* **51**, 263 (1994); (b) K. Soda and N. Tsuruta, *J. Phys. Soc. Jpn.* **63**, 814 (1994).
- ²² S. Shimizu, M. Ikeguchi, and K. Shimizu, *Chem. Phys. Lett.* **268**, 93 (1997).
- ²³ A. Ben-Naim and R. Lovett, *J. Phys. Chem. B* **101**, 10535 (1997).
- ²⁴ H. Reiss, H. L. Frisch, and E. Helfand, *J. Chem. Phys.* **31**, 369 (1959); E. Helfand, H. Reiss, and H. L. Frisch, *ibid.* **33**, 1379 (1960); H. Reiss, *Adv. Chem. Phys.* **69**, 281 (1965).
- ²⁵ R. A. Pierotti, *J. Phys. Chem.* **69**, 281 (1965).
- ²⁶ R. O. Neff and D. McQuarrie, *J. Phys. Chem.* **77**, 413 (1973).
- ²⁷ D. Sitkoff, K. A. Sharp, and B. Honig, *Biophys. Chem.* **51**, 397 (1994).
- ²⁸ K. A. Sharp, S. Kumar, P. J. Rosky, R. A. Friedman, and B. Honig, *J. Phys. Chem.* **100**, 14166 (1996).
- ²⁹ J. L. Lebowitz, *Phys. Rev.* **133**, 895A (1963).
- ³⁰ R. Evans, in *Fundamentals of Inhomogeneous Fluids*, edited by D. Henderson (Marcel Dekker, New York, 1992), p. 85; J. A. Barker and D. Henderson, *Rev. Mod. Phys.* **48**, 587 (1976).
- ³¹ N. Matubayasi, L. H. Reeds, and R. M. Levy, *J. Phys. Chem.* **98**, 10640 (1994).
- ³² J. L. Lebowitz and J. K. Percus, *Phys. Rev.* **122**, 1675 (1961).
- ³³ M. S. Wertheim, *J. Chem. Phys.* **87**, 7323 (1987).
- ³⁴ W. G. Chapman, G. Jackson, and K. E. Gubbins, *Mol. Phys.* **65**, 191 (1988).
- ³⁵ K. A. Sharp, A. Nicolls, R. F. Fine, and B. Honig, *Science* **252**, 106 (1991).
- ³⁶ O. Sinanoglu, *J. Chem. Phys.* **75**, 463 (1981).
- ³⁷ S. Shimizu, M. Ikeguchi, S. Nakamura, and K. Shimizu (unpublished).
- ³⁸ The packing fraction of liquid alkanes have been calculated by the liquid density from (a) and hard-sphere radius of (b). (a) *CRC Handbook in Chemistry and Physics*, 77th ed., edited by D. R. Lide (Chemical Rubber Corp., Boca Raton, FL, 1996); C. Vega, L. G. MacDowell, and P. Padilla, *J. Chem. Phys.* **104**, 701 (1996).